Every paper on every conference

2018

CVPR, AAAI, NIPS, ECCV

201810957 염지현

overview

1. Face & Facial
2. Head
3. Portrait
4. Style & Stylization

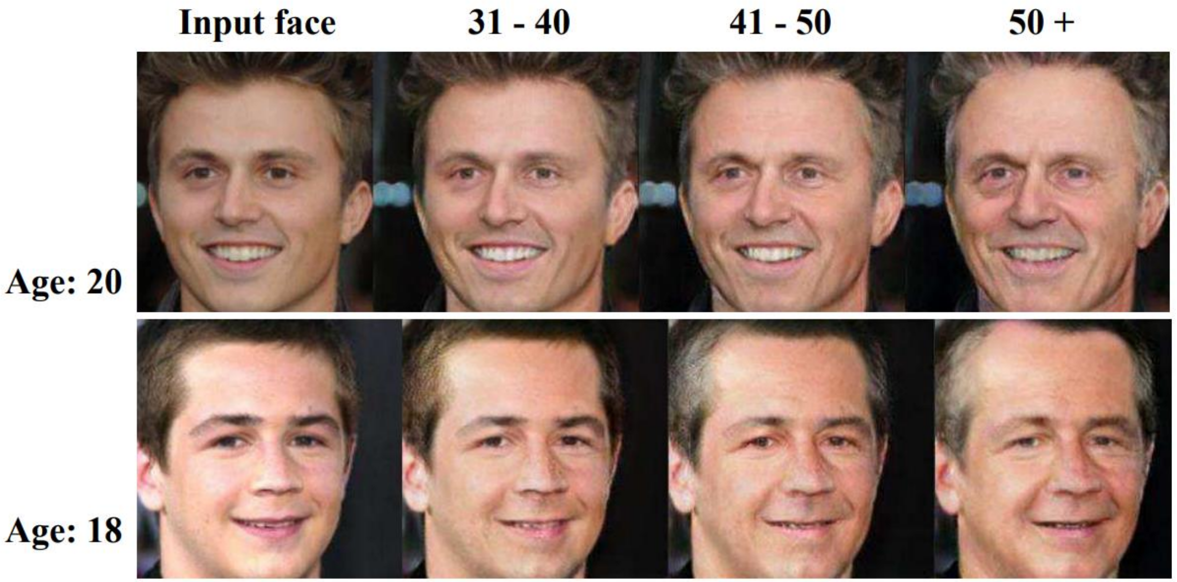
Face & Facial

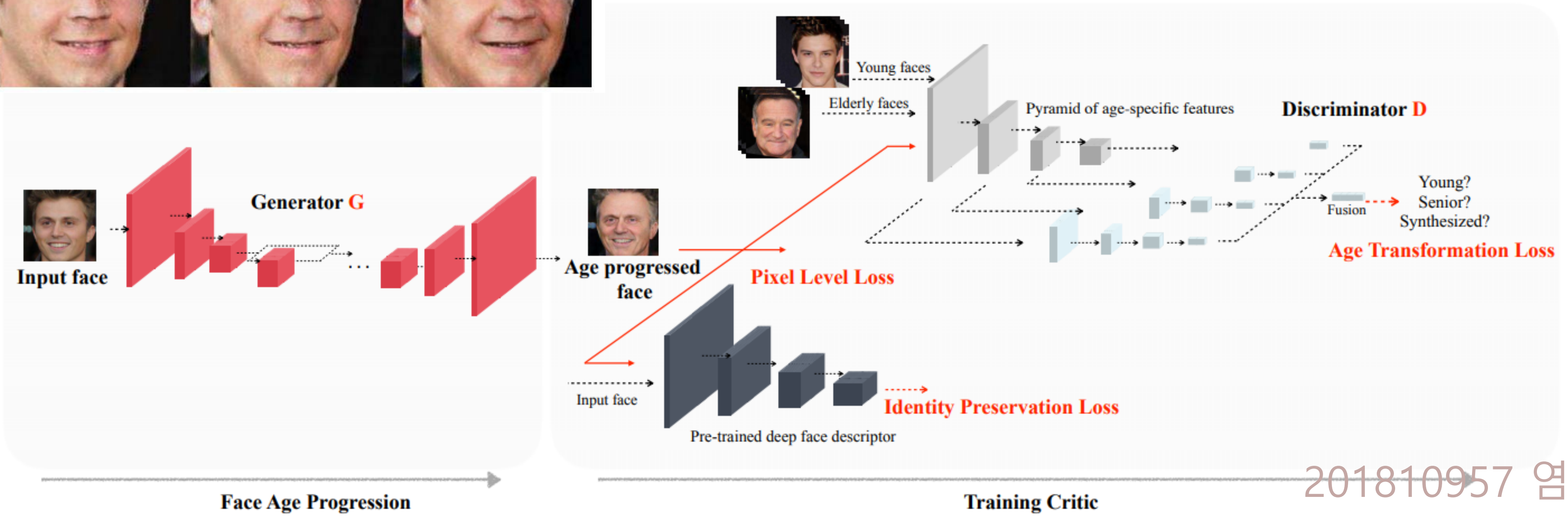
CVPR

201810957 염지현

[1] Learning Face Age Progression: A Pyramid Architecture of GANs

Hongyu Yang[1] Di Huang[1*] Yunhong Wang[1] Anil K. Jain[2]





201810957 염지현

[1] Learning Face Age Progression: A Pyramid Architecture of GANs

Hongyu Yang[1] Di Huang[1*] Yunhong Wang[1] Anil K. Jain[2]

Problem

: Face age progression

Strategy

- age-related GAN loss, individual-dependent critic을 사용하여 age transformation을 향상시키고 identity를 안정화 시킨다.
- Generator :
  - input-young face
  - encoder & decoder
  - 4개의 residual block → input & output face 연결
  - 3개의 fractionally-strided convolutional layers
- Discriminator :
  - 생성된 이미지와 나이 든 이미지가 동일해야 함
  - binary cross entropy

Background technique

: GAN, CNN

Dataset

: CACD, MORPH

201810957 염지현

[2] Learning Deep Models for Face Anti-Spoofing: Binary or Auxillary Supervision

Yaojie Liu, Amin Jourabloo, Xiaoming Liu

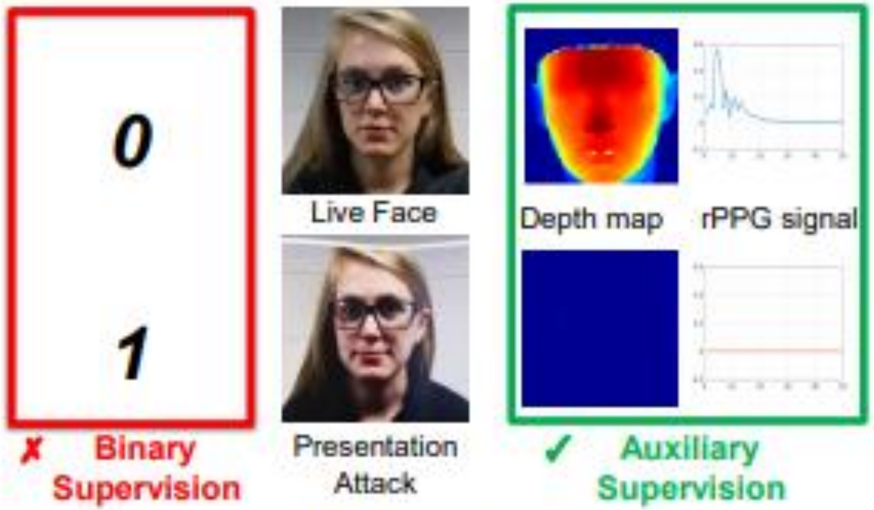


Figure 1. Conventional CNN-based face anti-spoof approaches utilize the binary supervision, which may lead to overfitting given the enormous solution space of CNN. This work designs a novel network architecture to leverage two auxiliary information as supervision: the depth map and rPPG signal, with the goals of improved generalization and explainable decisions during inference.

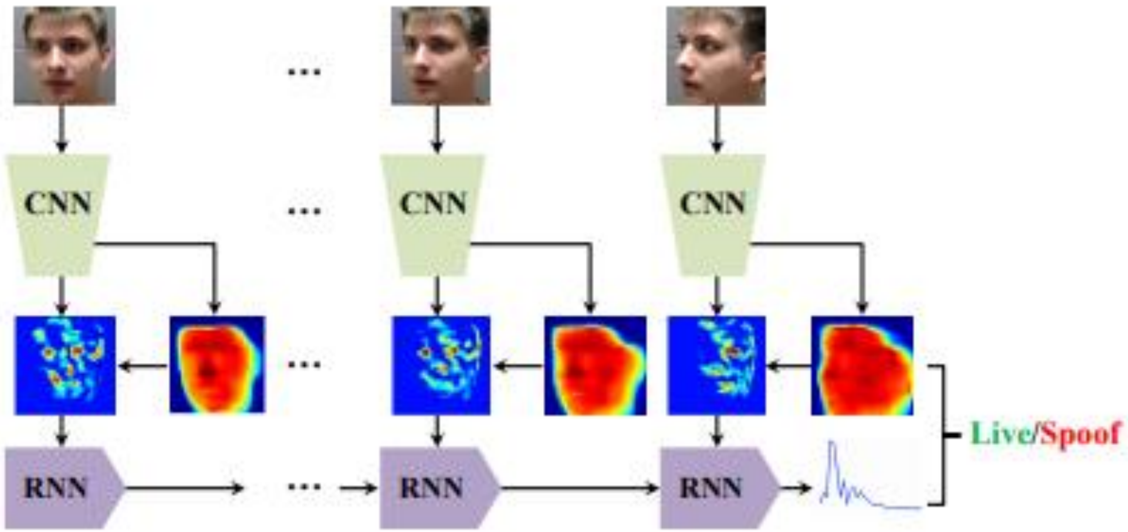


Figure 2. The overview of the proposed method.

201810957 염지현

[2] Learning Deep Models for Face Anti-Spoofing: Binary or Auxillary Supervision

Yaojie Liu, Amin Jourabloo, Xiaoming Liu

Problem
: Face Anti-Spoofing

Strategy
- rPPG SIGNAL
- Depth estimation

Background technique
: CNN

Dataset
: NVAA, CASIA-MFSO, Replay-Attrack, MSU-MFSD, Oulu-NPU, Siw

[3] A Prior-Less Method for Multi-Face Tracking in Unconstrained Videos

Chung-Ching Lin, Ying Hung

Figure 1: An example of multi-face tracking in unconstrained videos. The Bruno Mars music video shows the task is challenging due to partial occlusion and significant variations of lighting condition, camera angle, expression, and head pose across shots.

201810957 염지현

[3] A Prior-Less Method for Multi-Face Tracking in Unconstrained Videos

Chung-Ching Lin, Ying Hung

Problem

: Multi-Face Tracking

Strategy

- New multi-face tracking
- re-identification algorithm

Background technique

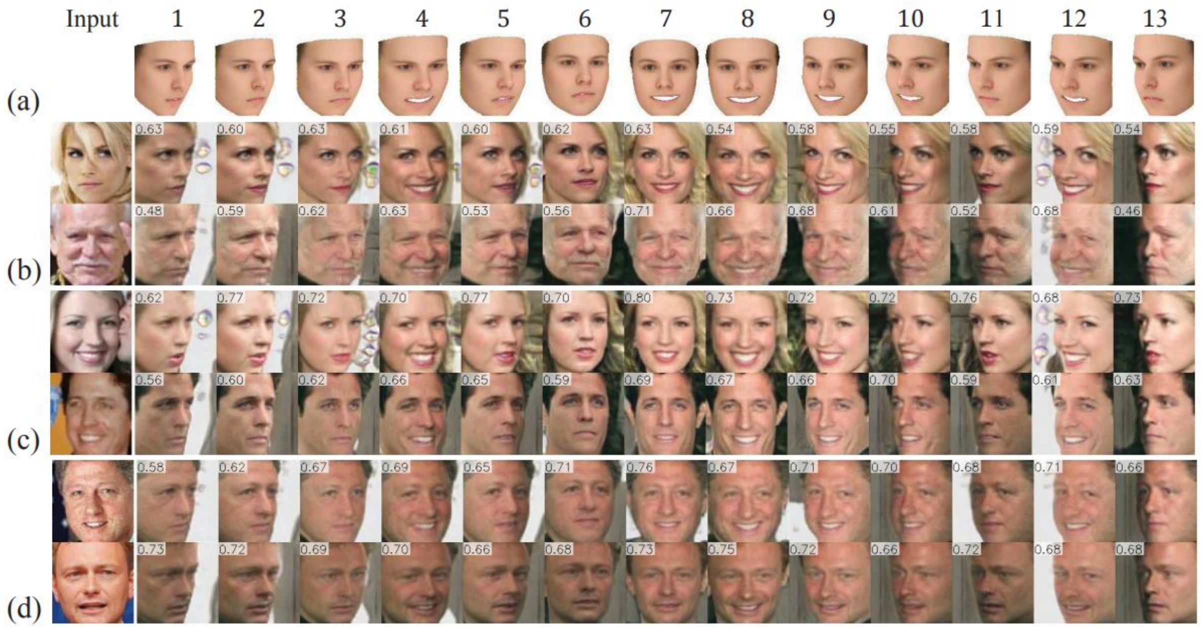
: CNN

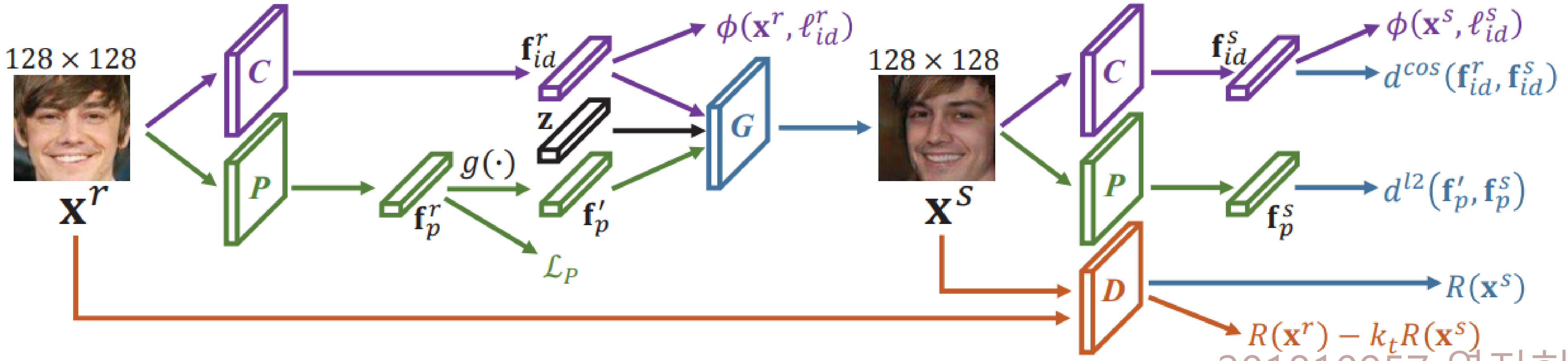
Dataset

: EHMVD

[4] FaceID-GAN: Learning a symmetry Three-Player GAN for identity-Preserving Face Synthesis

Yujun Shen, Ping Luo, Junjie Yan, Xiaogang Wang, Xiaoou Tang





201810957 염지현

[4] FaceID-GAN: Learning a symmetry Three-Player GAN for identity-Preserving Face Synthesis

Yujun Shen, Ping Luo, Junjie Yan, Xiaogang Wang, Xiaoou Tang

Problem
: Face Synthesis

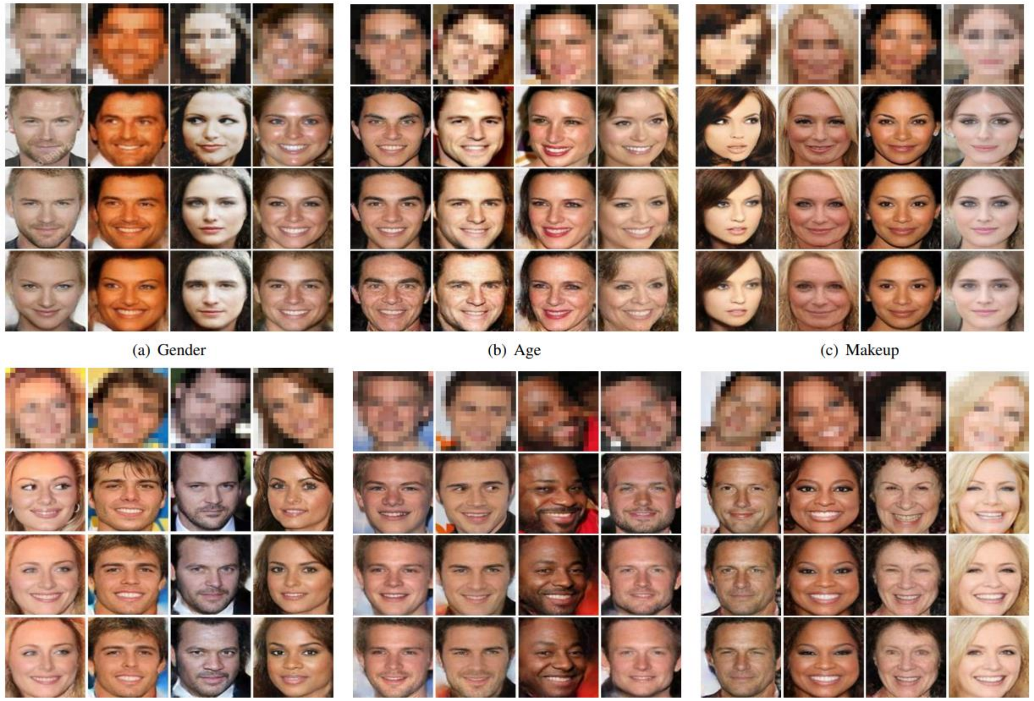
Strategy
- FaceID-GAN
  → Three players game, identity classifier : 입력 및 출력 이미지의 신원 특징을 추출하는데 사용
- Identity classifier + Discriminator 협력 → 얼굴 정체성 및 품질 향상
- 얼굴 정체성을 유지하면서 여러 각도의 얼굴을 생성할 수 있는 게 핵심

Background technique
: GAN

Dataset
: IJB-A

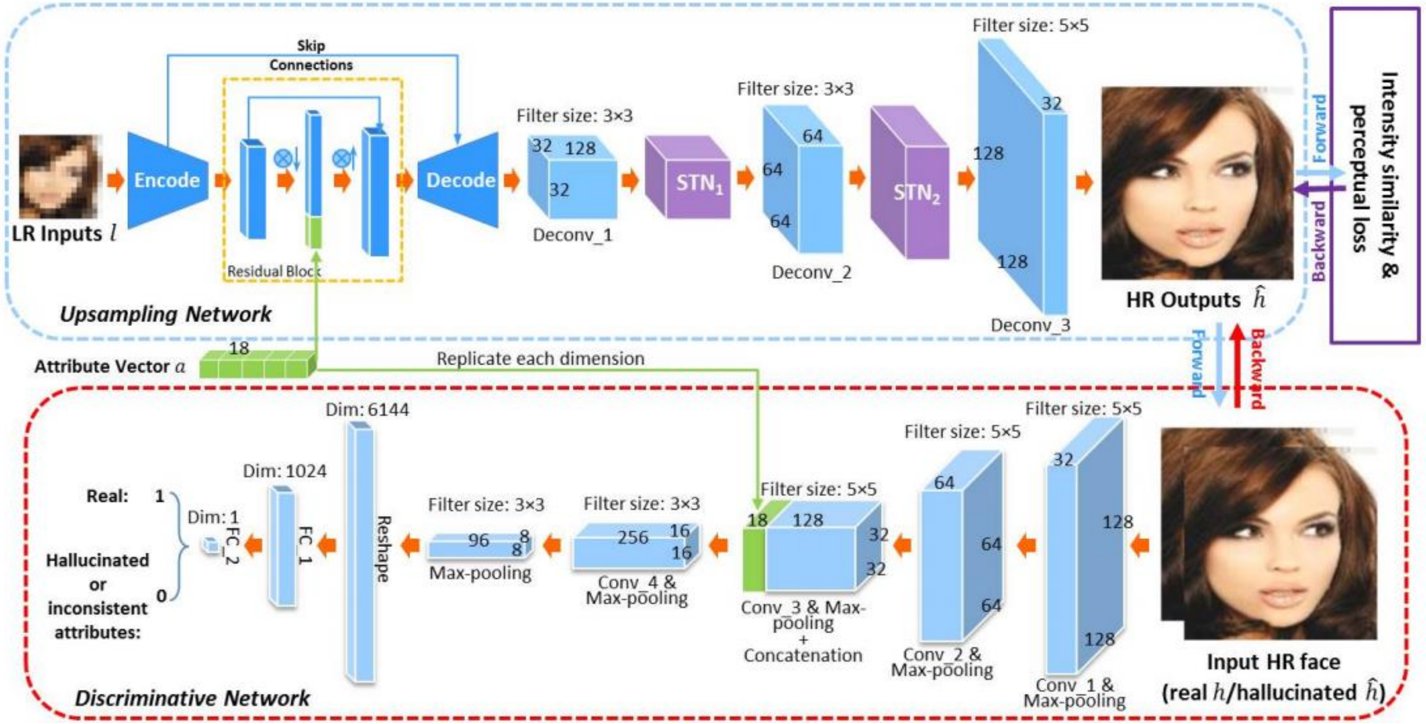
[5] Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes

Xin Yu, Basura Femando, Richard Hartley, Faith Porikli



(a) Gender (b) Age (c) Makeup

(d) Nose (e) Beard (f) Eyes



201810957 염지현

[5] Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes

Xin Yu, Basura Femando, Richard Hartley, Faith Porikli

Problem
: Face Hallucination

Strategy
- Low-Resolution face image를 업 샘플링 하는 대신 먼저 얼굴 속성을 사용하여 Lower-Resolution
  이미지 인코딩 후 feature map을 super-resolving image로 변환
- attribute information과 feature map을 skip connection으로 연결하기 위한 오토 인코더 사용
  → 시각적 정보 + 의미 정보

+) 특정 속성을 추가하거나 제거하기 위해 얼굴 속성을 조정하여 업 샘플링 된 HR 얼굴을 추가로 수정 가능

Background technique
: CNN, GAN

Dataset
: CelebA

[6] Learning From Millions of 3D Scans for Large-Scale 3D Face Recognition

Syed Zulqarnain Gilani, Ajmal Mian

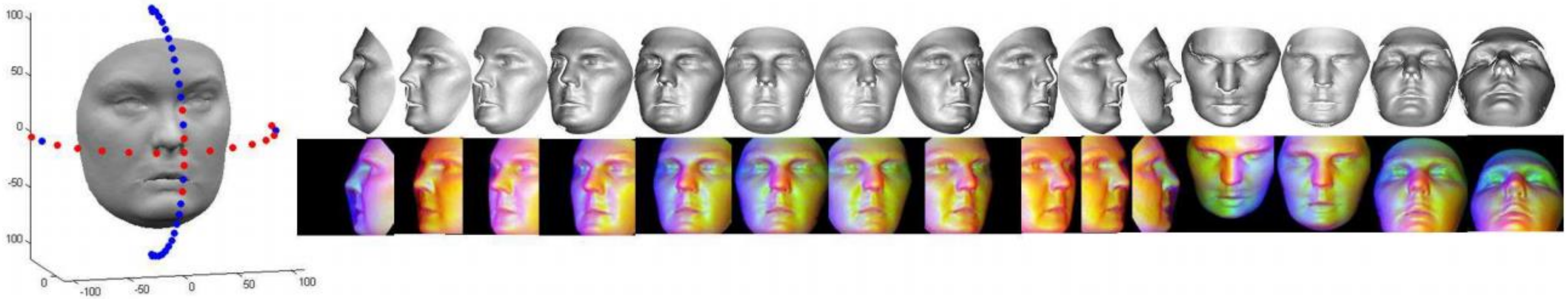
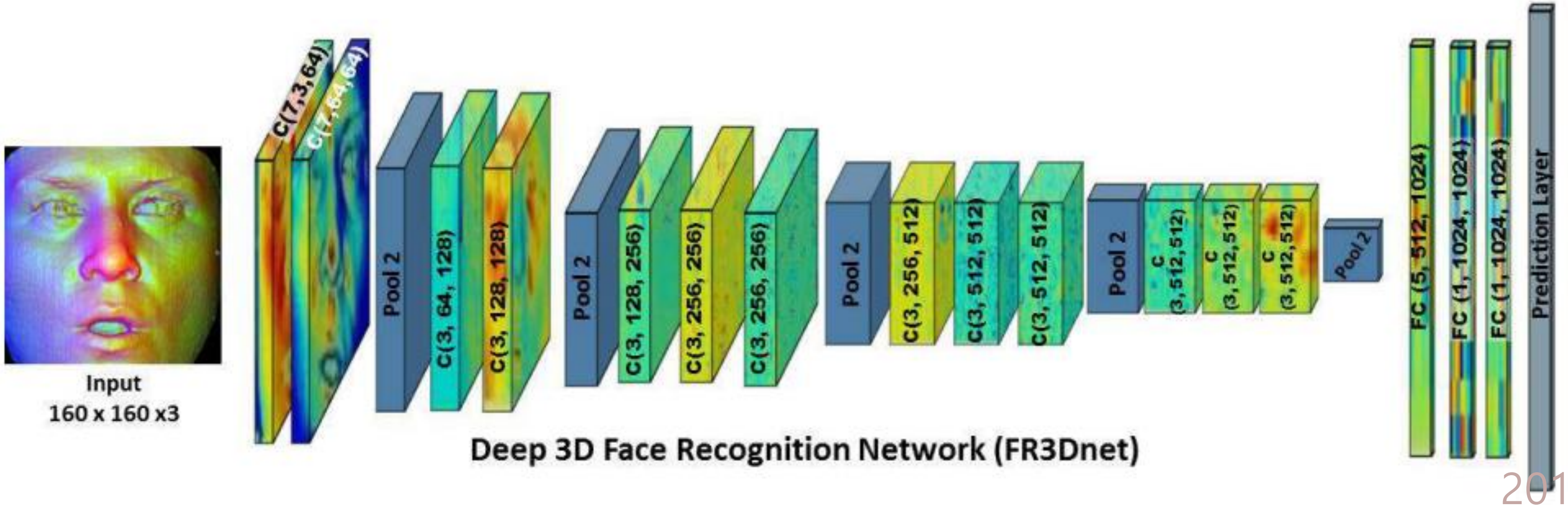


Figure 3. Position of cameras on a hemisphere surrounding the 3D face and the 15 poses generated as a result.



201810957 염지현

[6] Learning From Millions of 3D Scans for Large-Scale 3D Face Recognition

Syed Zulqarnain Gilani, Ajmal Mian

Problem
: 3D Face Recognition

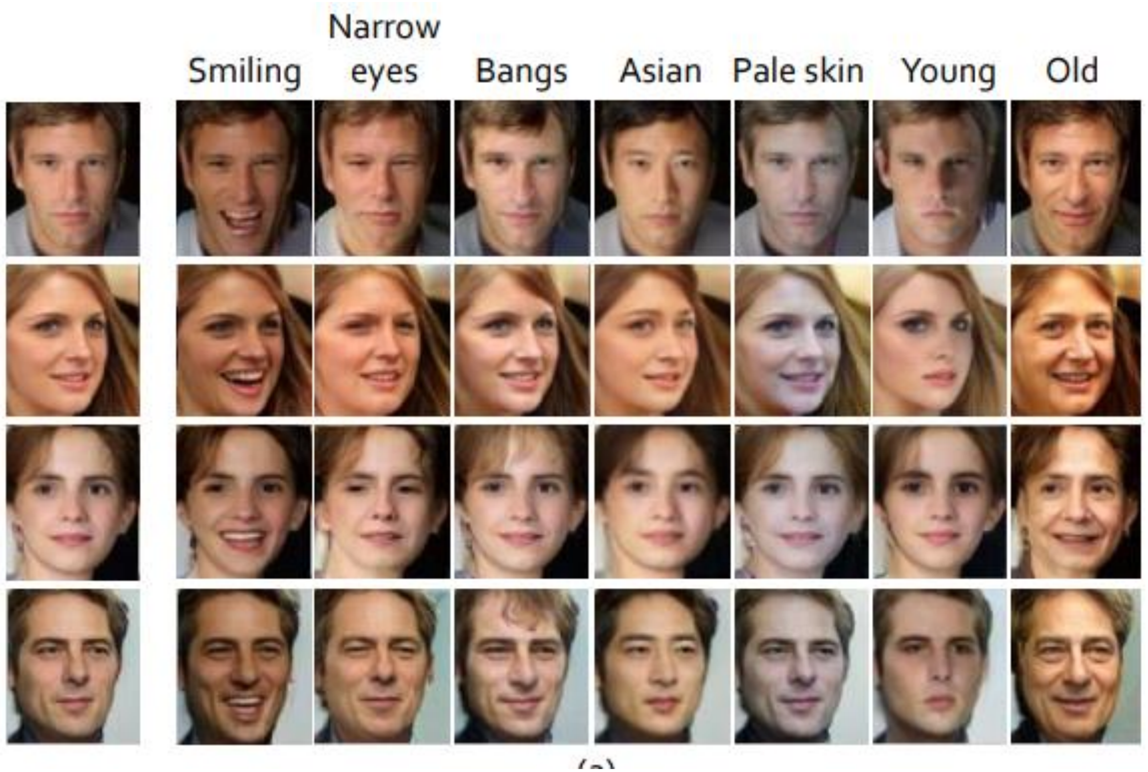
Strategy
- 100K identities에 대한 3.1 Million face datasets merge

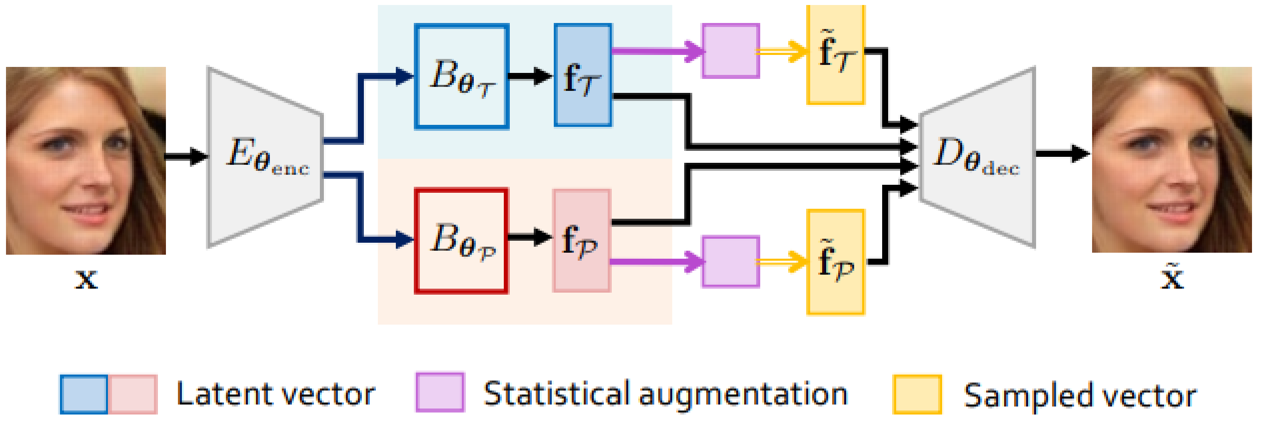
Background technique
: CNN

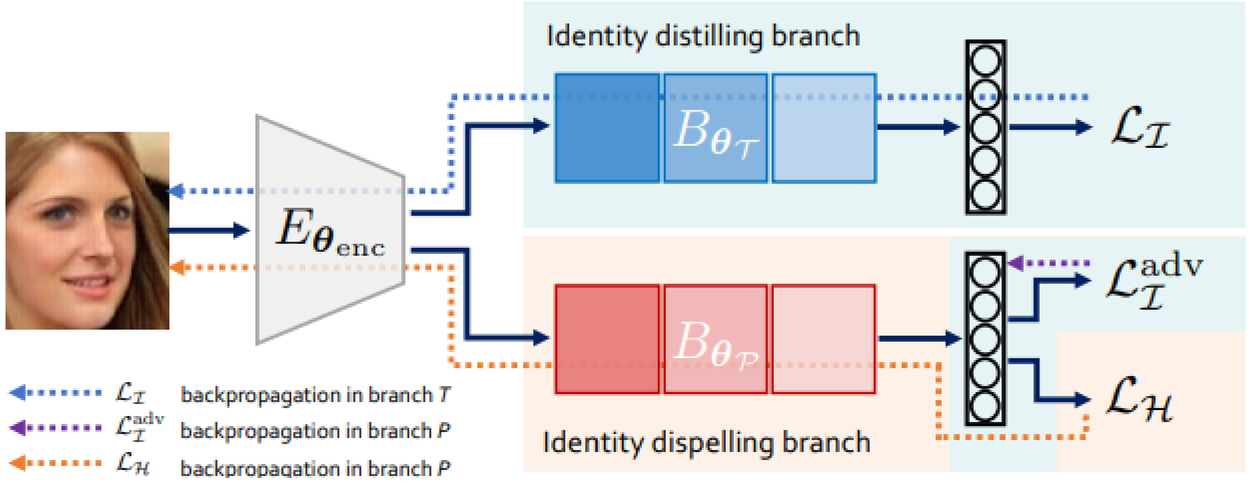
Dataset
: RGCv2, BU3DFE, BU4DFE, Bosphorus, CASIA, GavabDB, TexasFRD, 3D-TEC, UMBDB, ND-2006

[7] Exploring Disentangled Feature Representation Beyond Face Identification

Yu Liu, Fangyin Wei, Jing Shao, Lu Sheng, Junjie Yan, Xiaogang Wang







201810957 염지현

[7] Exploring Disentangled Feature Representation Beyond Face Identification

Yu Liu, Fangyin Wei, Jing Shao, Lu Sheng, Junjie Yan, Xiaogang Wang

Problem

: Feature Representation(identity preserving)

Strategy

- D^2A2 → 신원 추출 기능(신원 확인을 위한) + 신원 삭제 기능(시스템을 속이기 위한)
- 인코더를 통해 데이터에서 특징을 추출 후 다시 입력 공간으로 매핑하는 디코더 제안

Background technique

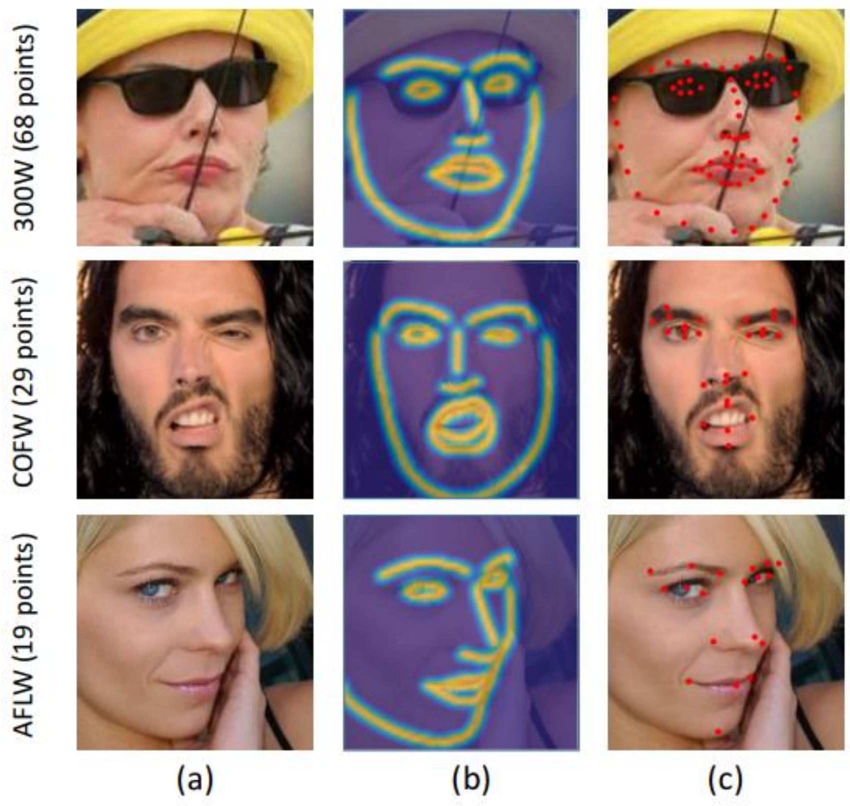
: CNN, GAN

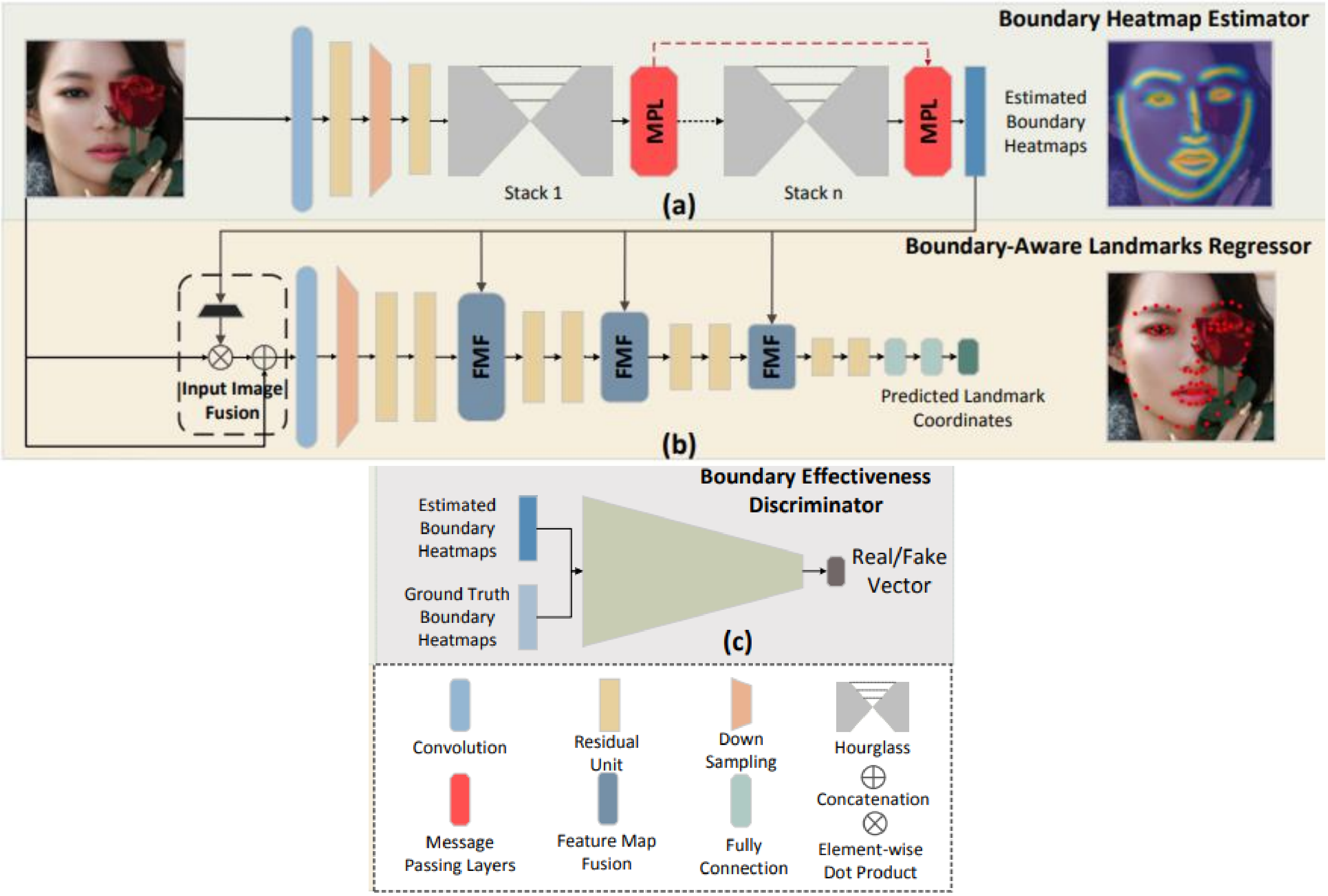
Dataset

: MSCeleb-1M

[8] Look at Boundary: A Boundary-Aware Face Alignment Algorithm

Wayne Wu, Chen Qian, Shuo Yang, Quan Wang, Yici Cai, Qiang Zhou





201810957 염지현

[8] Look at Boundary: A Boundary-Aware Face Alignment Algorithm

Wayne Wu, Chen Qian, Shuo Yang, Quan Wang, Yici Cai, Qiang Zhou

Problem
: Face Landmark detection

Strategy
1. face boundary heatmaps 추정
2. 추정한 heatmaps으로부터 landmark regression
   → face boundary와 landmark 사이의 관계를 탐색하기 위한 discriminator

Background technique
: CNN, GAN

Dataset
: WFLW, AFLW, COFW, 300W

[9] Towards Pose Invariant Face Recognition in the Wild

Jian Zhao, Yu Cheng, Yan Xu, Lin Xiong, Jianshu Li, Fang Zhao, Karlekar Jayashree, Sugiri Pranata, Shengmei Shen, Junliang Xing, Shuicheng Yan, Jiashi Feng

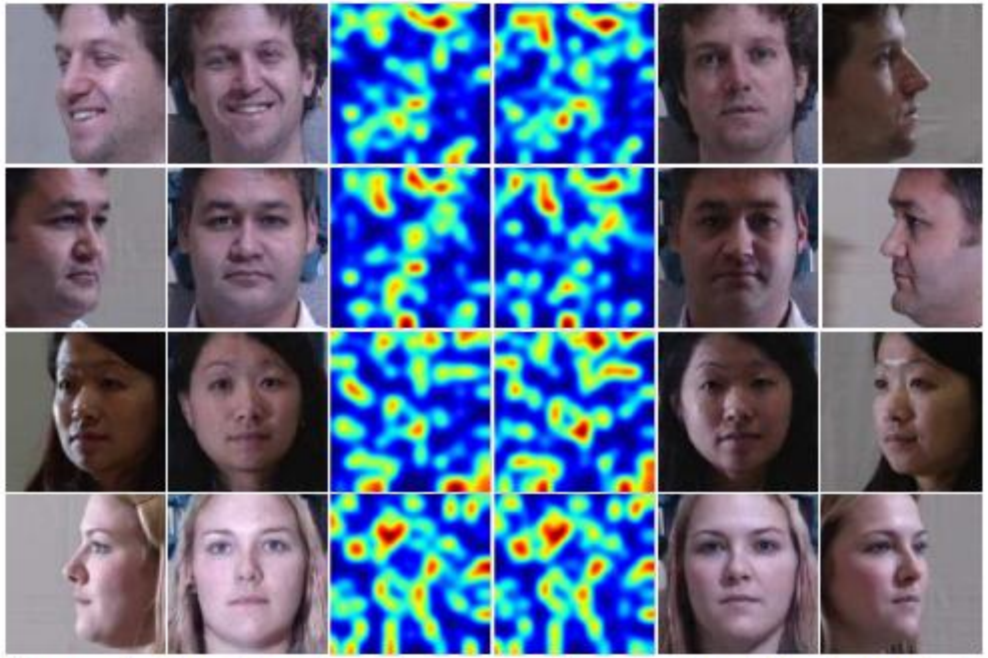


Figure 1: Pose invariant face recognition in the wild. *Col.* 1 & 6: distinct identities under different poses with other unconstrained factors (different expressions and lighting conditions); *Col.* 2 & 5: recovered frontal faces with our proposed PIM model; *Col.* 3 & 4: learned facial representations with our proposed PIM model. PIM can learn pose-invariant representations and recover photorealistic frontal faces effectively. The representations are extracted from the penultimate layer of PIM.

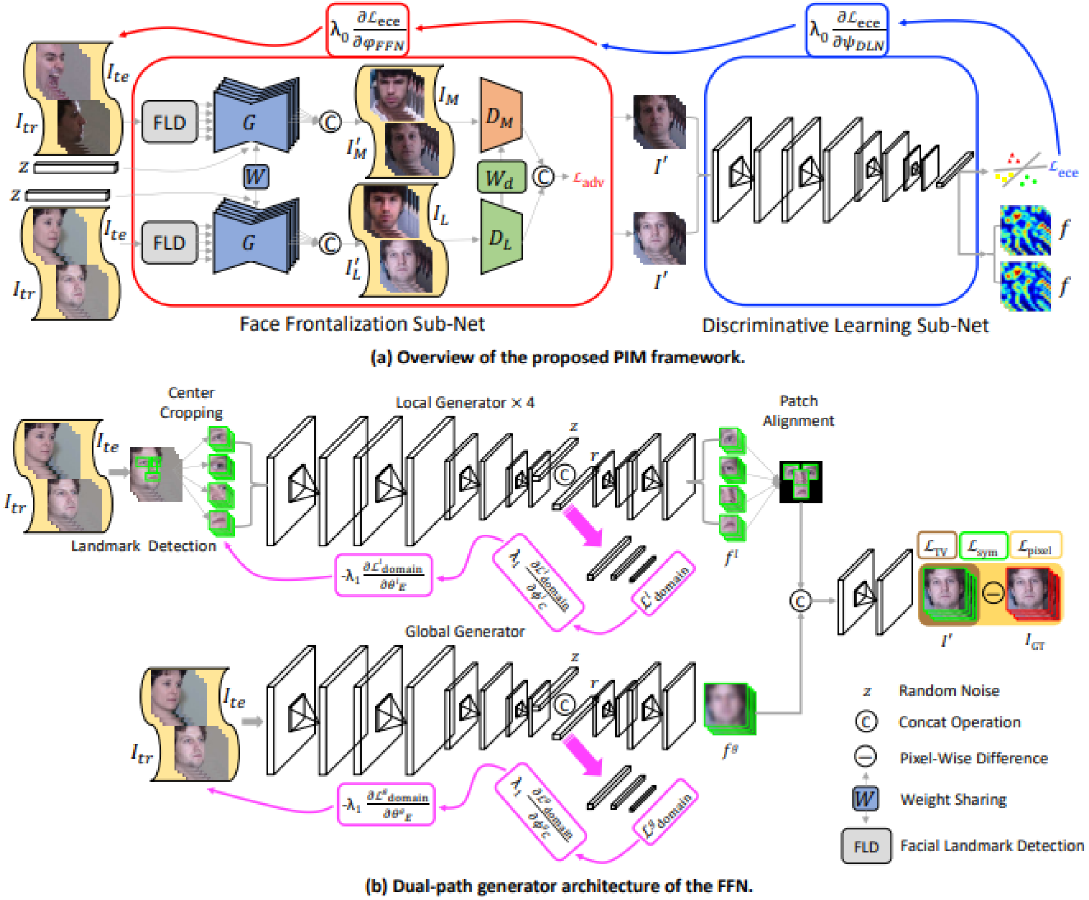


Figure 2: Pose Invariant Model (PIM) for face recognition in the wild. The PIM contains an Face Frontalization sub-Net (FFN) and a Discriminative Learning sub-Net (DLN) that jointly learn end-to-end. FFN is a dual-path (*i.e.*, simultaneously perceiving global structures and local details) GAN augmented by (1) unsupervised cross-domain (*i.e.*, I_{tr} and I_{te}) adversarial training and (2) a siamese discriminator with a "learning to learn" strategy — convolutional parameters (*i.e.*, W_d) dynamically predicted by the "learner" D_L of the discriminator and transferred to D_M. DLN is a generic Convolutional Neural Network (CNN) for face recognition optimized by the proposed enforced cross-entropy optimization. It takes in the frontalized face images from FFN and outputs learned pose invariant facial representations.

[9] Towards Pose Invariant Face Recognition in the Wild

Jian Zhao, Yu Cheng, Yan Xu, Lin Xiong, Jianshu Li, Fang Zhao, Karlekar Jayashree, Sugiri Pranata, Shengmei Shen, Junliang Xing, Shuicheng Yan, Jiashi Feng

Problem

: Face Recognition (Pose Variation)

Strategy

- PIM(Pose Invariant Model)
  - FFN(Face Frontalization Sub-Net) → 정면 image 생성
  - DNL(Discriminative LEARNING Sub-Net) → 포즈 불변 얼굴 표현 출력

Background technique

: GAN, CNN

Dataset

: Multi-PIE, LFW, CFP

[10] Real-Time Rotation-Invariant Face Detection With Progressive Calibration Networks

Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen

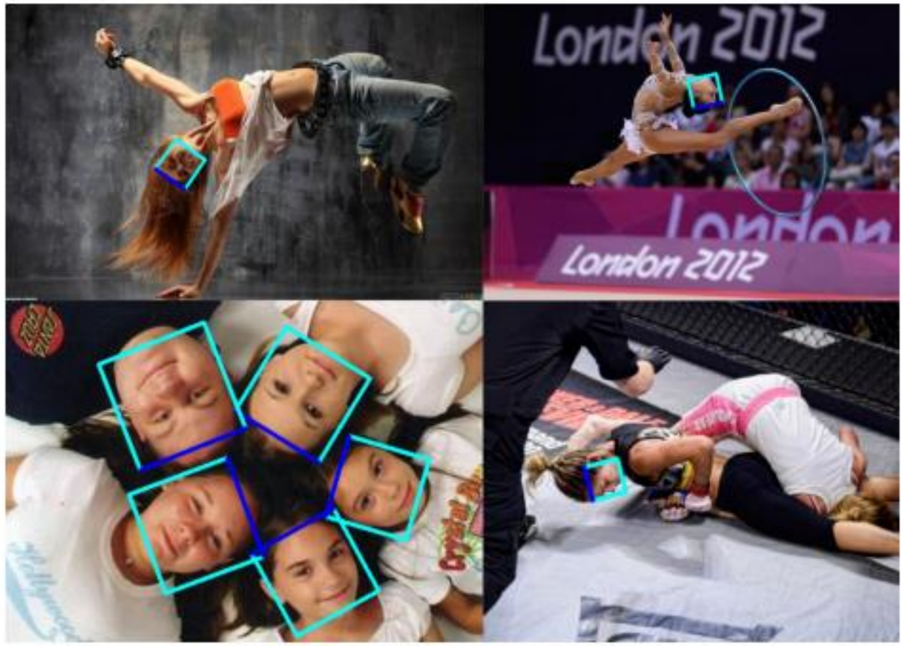


Figure 1. Many complex situations need rotation-invariant face detectors. The face boxes are the outputs of our detector, and the blue line indicates the orientation of faces.

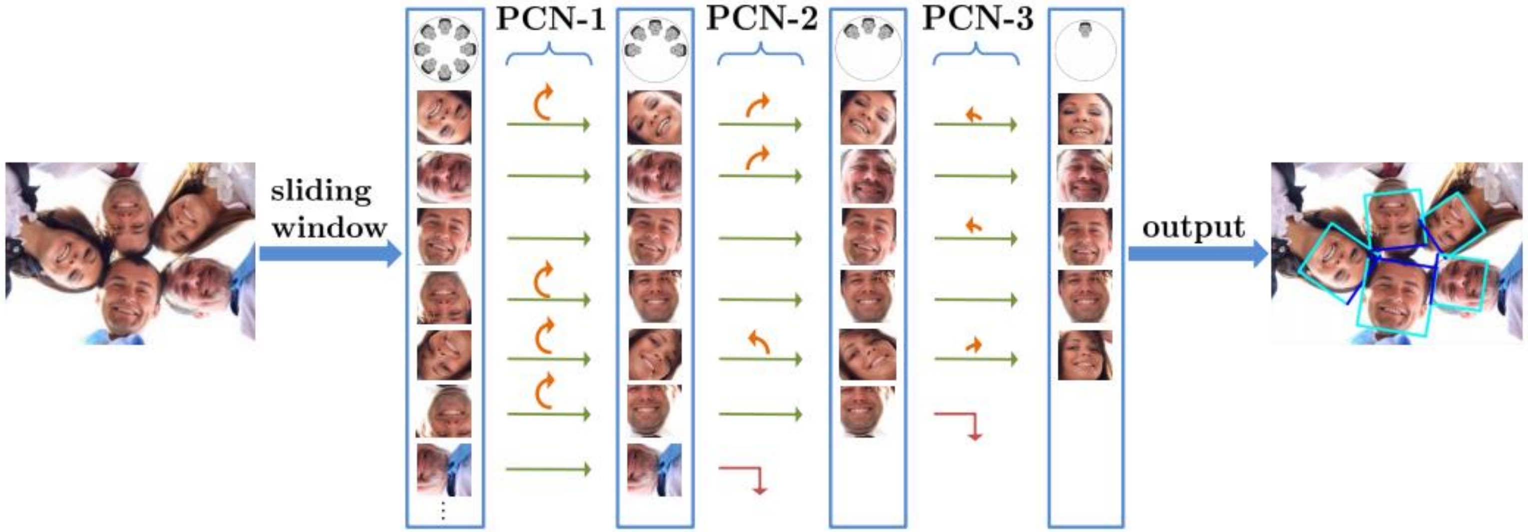


Figure 3. An overview of our proposed progressive calibration networks (PCN) for rotation-invariant face detection. Our PCN progressively calibrates the RIP orientation of each face candidate to upright for better distinguishing faces from non-faces. Specifically, PCN-1 first identifies face candidates and calibrates those facing down to facing up, halving the range of RIP angles from $[-180°, 180°]$ to $[-90°, 90°]$. Then the rotated face candidates are further distinguished and calibrated to an upright range of $[-45°, 45°]$ in PCN-2, shrinking the RIP ranges by half again. Finally, PCN-3 makes the accurate final decision for each face candidate to determine whether it is a face and predict the precise RIP angle.

[10] Real-Time Rotation-Invariant Face Detection With Progressive Calibration Networks

Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen

Problem
: Detecting faces with arbitrary rotation-in-plane

Strategy
- PCN(Progressive Calibration Networks)
    - 얼굴과 얼굴이 아닌 부분을 구분
    - 각 얼굴 후부의 RIP 방향 조정

Background technique
: CNN

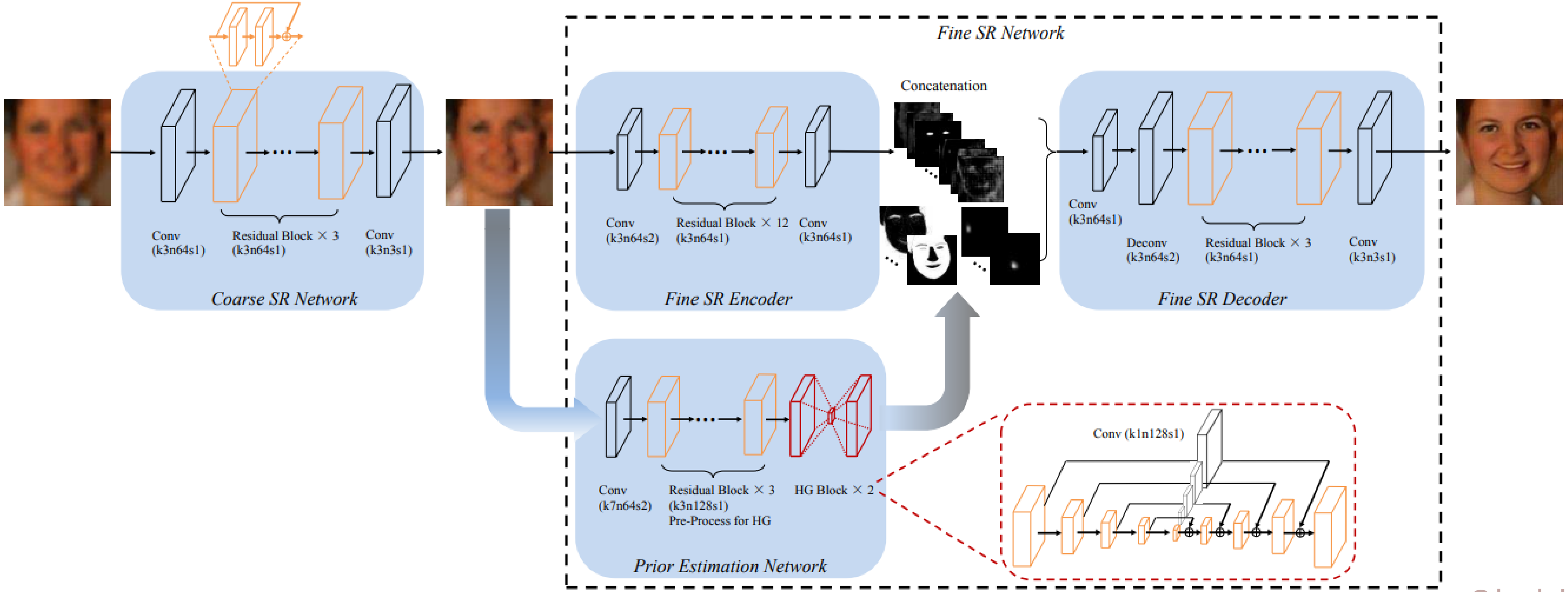
Dataset
: FDDB, WiderFace

201810957 염지현

[11] FSRNet: End-to-End Learning Face Super-Resolution With Facial Priors

Yu Chen, Ying Tai, Xiaoming Liu, Chunhua Shen, Jian Yang

Figure 1: **Visual results of different super-resolution methods.**



201810957 염지현

[11] FSRNet: End-to-End Learning Face Super-Resolution With Facial Priors

Yu Chen, Ying Tai, Xiaoming Liu, Chunhua Shen, Jian Yang

Problem
: Face Super-Resolution

Strategy
- FSRNet
  : Coarse SR Network의 output → fine SR Network & Prior Estimation Network 의 입력을 통해
  이미지의 특징을 추출한 후 Fine SR Decoder를 통해 고화질 이미지 생성
- FSRGAN
  ; FSRNet + adversarial loss

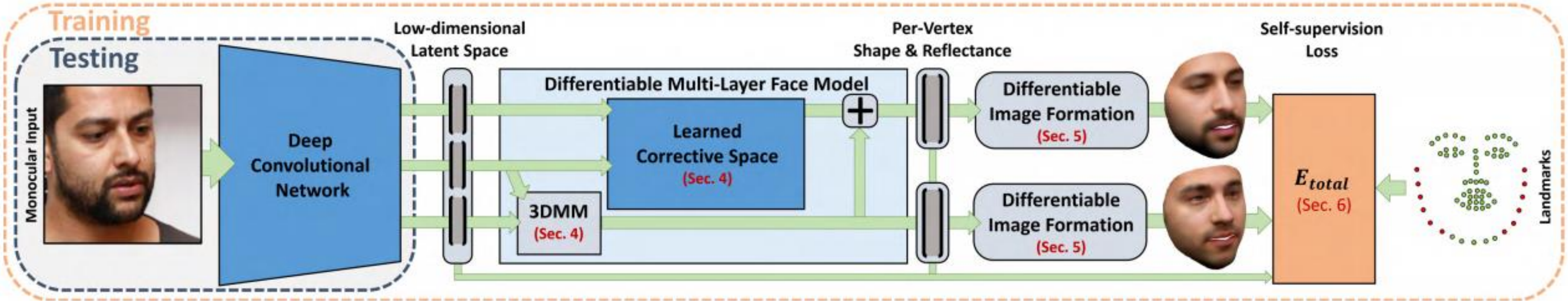
Background technique
: CNN, GAN

Dataset
: Helen, CelebA

[12] Self-Supervised Multi-Level Face Model Learning for Monocular Reconstruction at Over 250 Hz

Ayush Tewari, Michael Zollhöfer, Pablo Garrido, Florian Bernard, Hyeongwoo Kim, Patrick Pérez, Christian Theobalt

Our novel monocular reconstruction approach estimates high-quality facial geometry, skin reflectance (including facial hair) and incident illumination at over 250 Hz. A trainable multi-level face representation is learned jointly with the feed forward inverse rendering network. End-to-end training is based on a self-supervised loss that requires no dense ground truth.



201810957 염지현

[12] Self-Supervised Multi-Level Face Model Learning for Monocular Reconstruction at Over 250 Hz

Ayush Tewari, Michael Zollhöfer, Pablo Garrido, Florian Bernard, Hyeongwoo Kim, Patrick Pérez, Christian Theobalt

Problem
: 3D Face Reconstruction

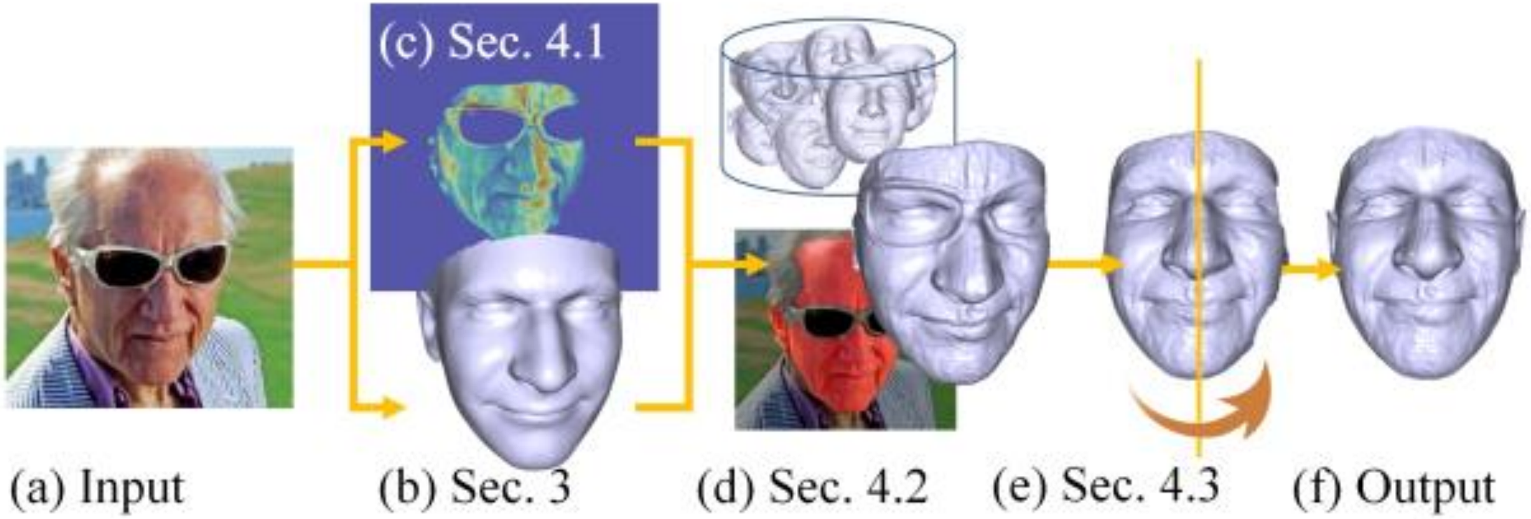
Strategy
- 기존의 연구들과 달리 얼굴의 기하학적인 부분, 피부 반사, 조명 등을 고려
- differentiable expert-designed renderer + self-supervised training loss & end-to-end learning

Background technique
: CNN

Dataset
: FaceWarehouse

[13] Extreme 3D Face Reconstruction: Seeing Through Occlusions

Anh Tuấn Trần, Tal Hassner, Iacopo Masi, Eran Paz, Yuval Nirkin, Gérard Medioni



201810957 염지현

[13] Extreme 3D Face Reconstruction: Seeing Through Occlusions

Anh Tuấn Trần, Tal Hassner, Iacopo Masi, Eran Paz, Yuval Nirkin, Gérard Medioni

Problem
: 3D Face Reconstruction

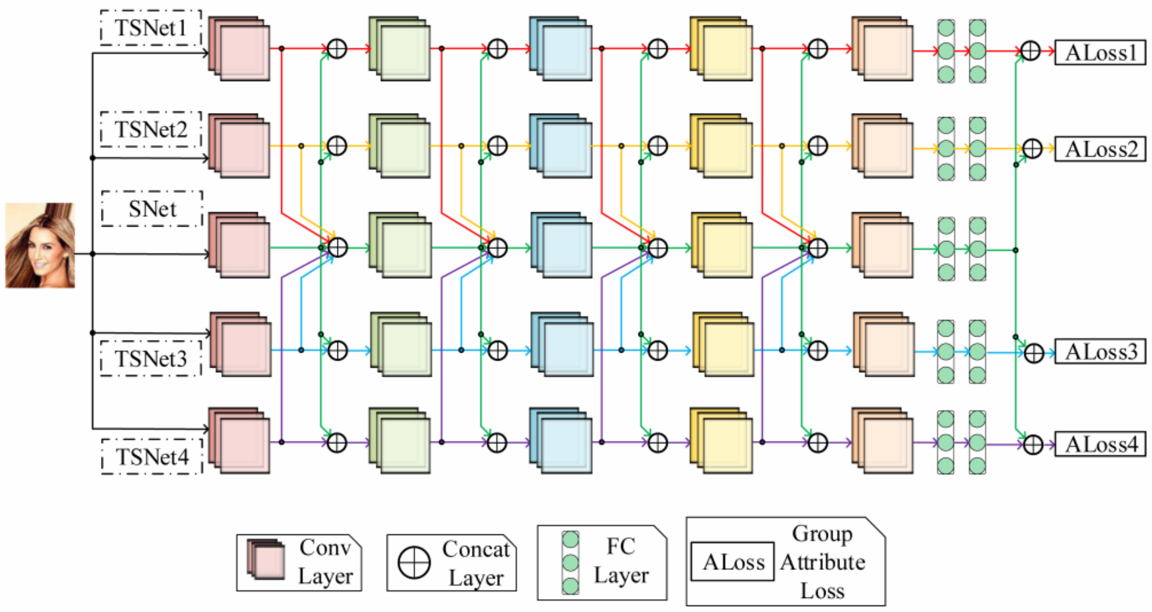
Strategy
1. Course 3D face shape 추정
2. bump map으로 표시된 detail을 layer 별로 layering
   +) bump map : Deep Convolution encoder-decoder 사용
   +) 가려진 얼굴에서 3D Face Reconstruction → 가려진 얼굴 추정

Background technique
: CNN

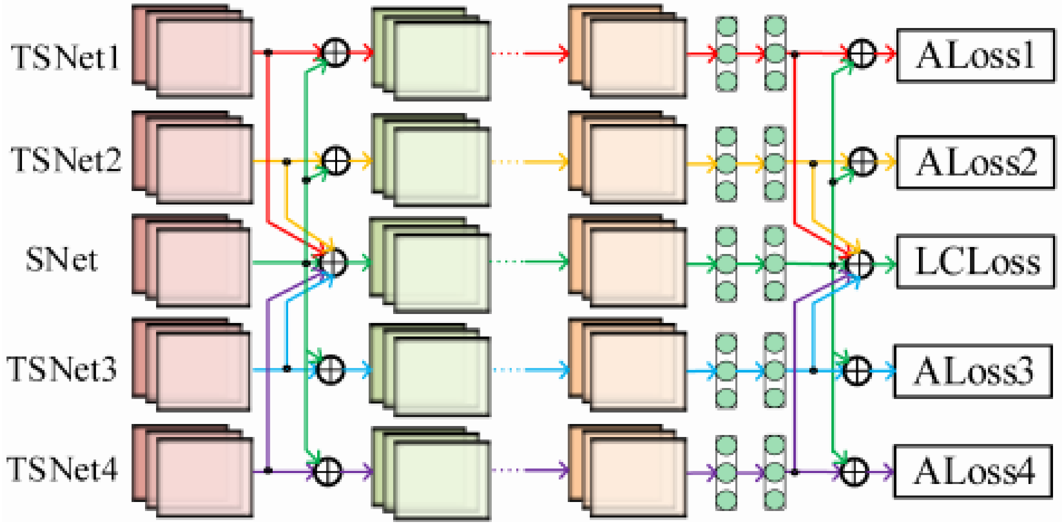
Dataset
: https://github.com/anhttran/extreme_3d_faces

[14] Partially Shared Multi-Task Convolutional Neural Network With Local Constraint for Face Attribute Learning

Jiajiong Cao, Yingming Li, Zhongfei Zhang



PS-
MCNN



PS-
MCNNLC

[14] Partially Shared Multi-Task Convolutional Neural Network With Local Constraint for Face Attribute Learning

Jiajiong Cao, Yingming Li, Zhongfei Zhang

Problem

: Face Attribute learning problem

Strategy

1. PS-MCNN(Partially Shared Multi-task Convolutional Neural Network)
   4개의 TSNets(Task Specific Networks)
   1개의 SNet(Shared Network)
   - PS(Partially Shared) 구조로 이루어져 있음
2. PS-MCNNLC(Partially Shared Multi-task Convolutional Neural Network with Local Constraint)
   - identity information을 활용하기 위해 각 샘플의 특징과 동일한 identity를 가진 neighbor와의
   차이를 최소화 하는 LCLoss 도입

Background technique

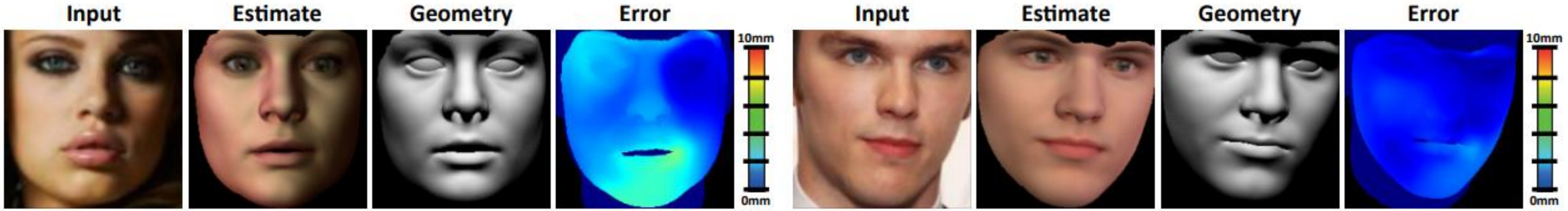
: CNN

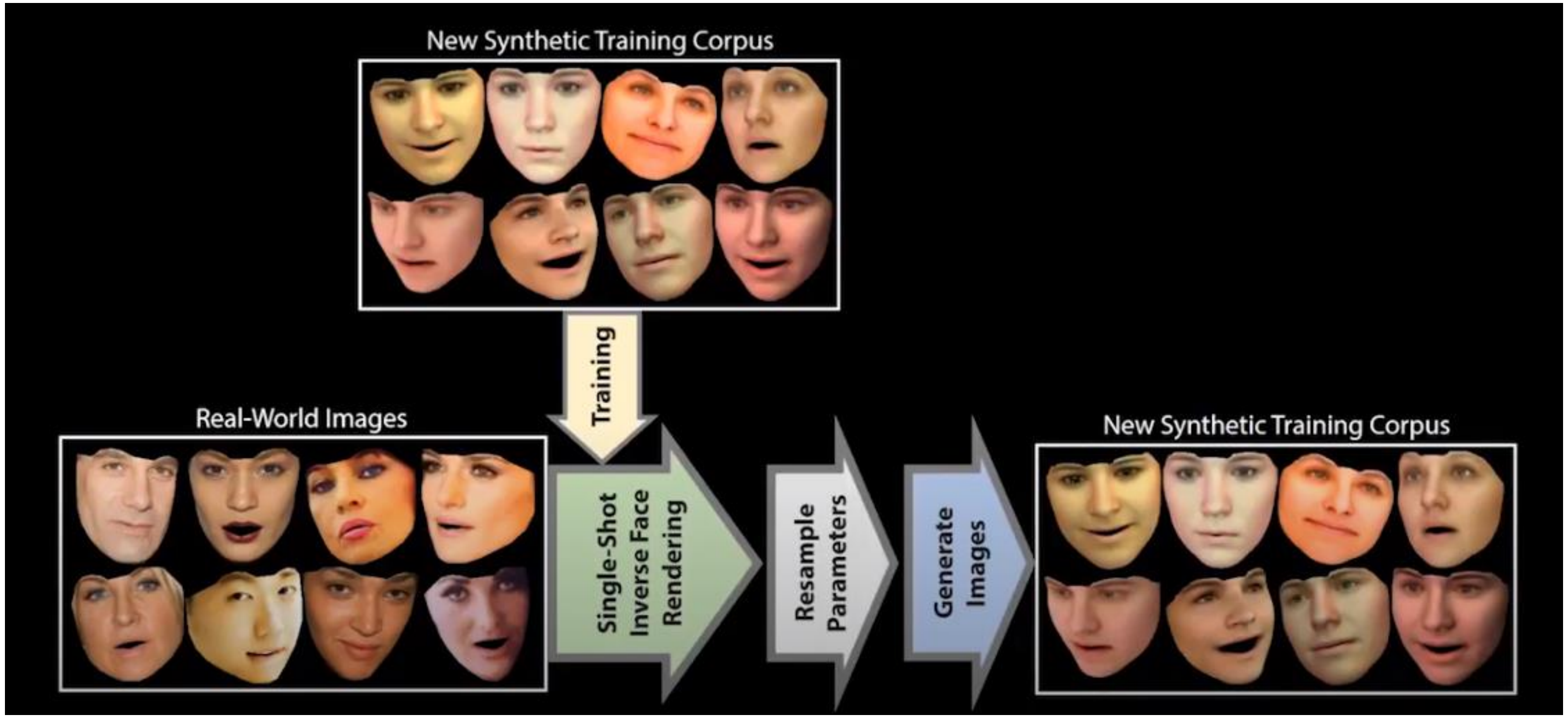
Dataset

: CelebA, LFWA

[15] InverseFaceNet: Deep Monocular Inverse Face Rendering

Hyeongwoo Kim, Michael Zollhöfer, Ayush Tewari, Justus Thies, Christian Richardt, Christian Theobalt





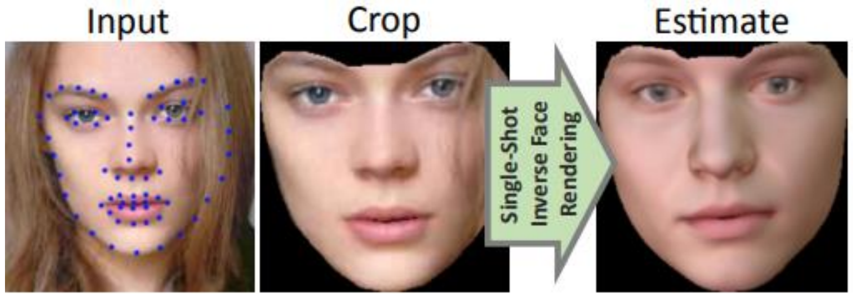
201810957 염지현

[15] InverseFaceNet: Deep Monocular Inverse Face Rendering

Hyeongwoo Kim, Michael Zollhöfer, Ayush Tewari, Justus Thies, Christian Richardt, Christian Theobalt

Problem
: Inverse Face Rendering

Strategy
1. Input → Detections → Crop → (Single-Shot Inverse Face Rendering) → Estimate

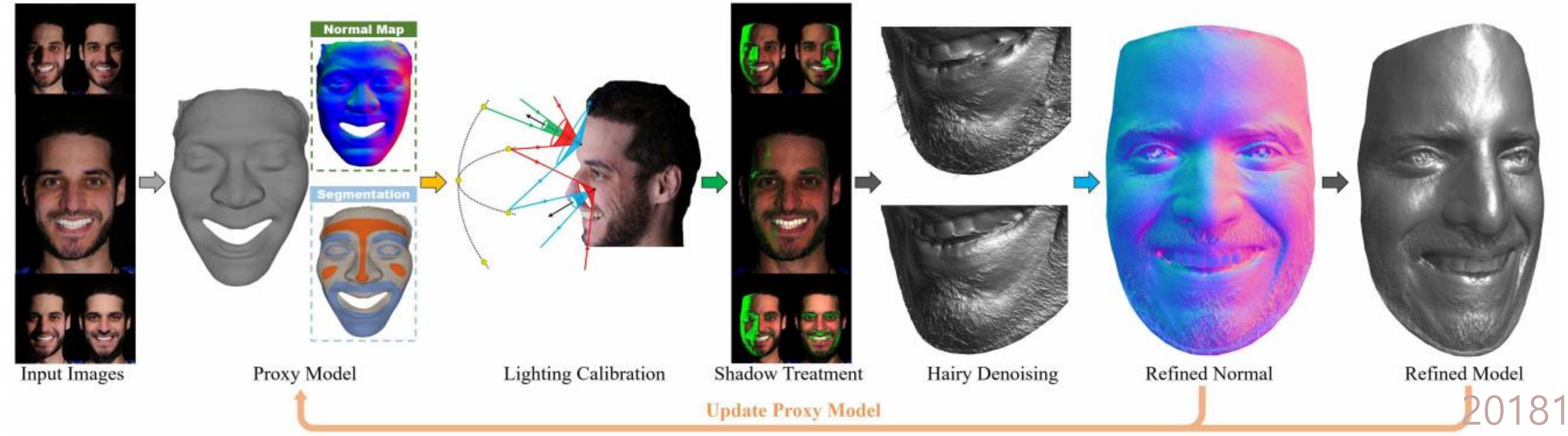


Background technique
: CNN

Dataset
: FaceWarehouse

[16] Sparse Photometric 3D Face Reconstruction Guided by Morphable Models

Xuan Cao, Zhang Chen, Anpei Chen, Xin Chen, Shiying Li, Jingyi Yu



201810957 염지현

[16] Sparse Photometric 3D Face Reconstruction Guided by Morphable Models

Xuan Cao, Zhang Chen, Anpei Chen, Xin Chen, Shiying Li, Jingyi Yu

Problem

: 3D Face Reconstruction

Strategy

1. input : 다양한 각도의 조명을 비추고 촬영한 이미지 + 입력 이미지와 일치하는 포즈와 표정을 가진 3D 모핑 모델 → Proxy model
2. Proxy Model에서 모든 조명의 위치와 조명을 추정
   → 조명의 위치는 피사체 얼굴에 상대적이므로 각 얼굴 부분에 대한 입사 조명 방향 쉽게 계산
3. 그림자 감지
4. 신뢰할 수 있는 3개의 입사광 선택한 후 고해상도 normal map 계산
5. bidirectional extremum filter 추가로 개발 → 노이즈 제거
6. depth gradient maps을 통합 → face geometry를 얻음

Background technique

: CNN

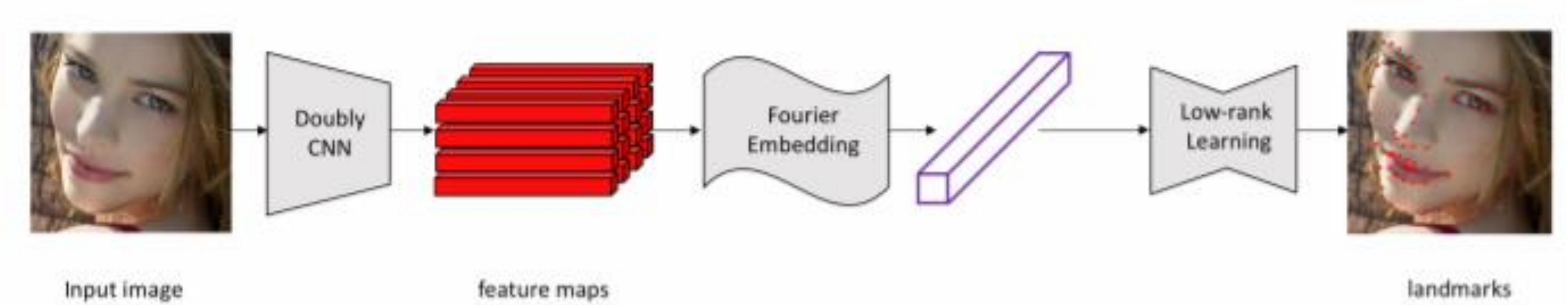
Dataset

: 언급 x

201810957 염지현

[17] Direct Shape Regression Networks for End-to-End Face Alignment

Xin Miao, Xiantong Zhen, Xianglong Liu, Cheng Deng, Vassilis Athitsos, Heng Huang



201810957 염지현

[17] Direct Shape Regression Networks for End-to-End Face Alignment

Xin Miao, Xiantong Zhen, Xianglong Liu, Cheng Deng, Vassilis Athitsos, Heng Huang

Problem
: Face Landmark Detection

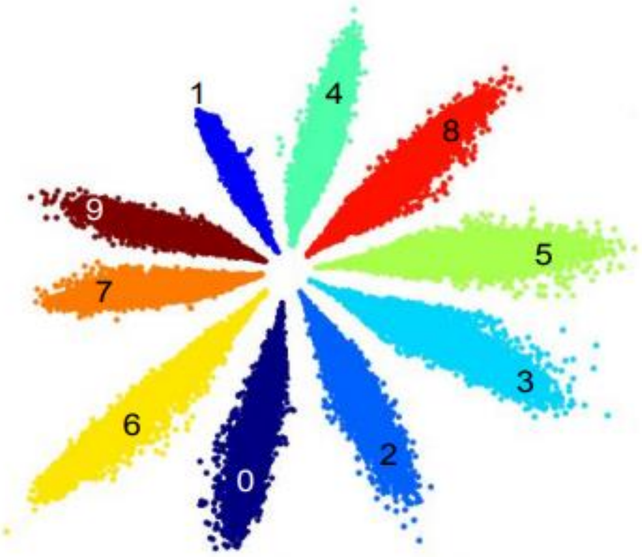
Strategy
- Doubly convolutional layer
- Fourier feature pooling layer

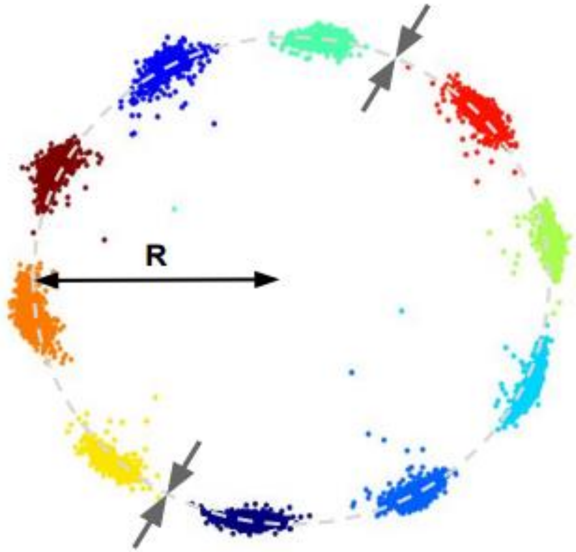
Background technique
: CNN

Dataset
: AFLW, 300W, CelebA, MAFL, 300VW

[18] Ring Loss: Convex Feature Normalization for Face Recognition

Yutong Zheng, Dipan K. Pal, Marios Savvides





$$L_R = \frac{\lambda}{2m} \sum_{i=1}^{m} (\|\mathcal{F}(\mathbf{x}_i)\|_2 - R)^2$$

201810957 염지현

[18] Ring Loss: Convex Feature Normalization for Face Recognition

Xin Miao, Xiantong Zhen, Xianglong Liu, Cheng Deng, Vassilis Athitsos, Heng Huang

Problem
: softmax를 강화하기 위한 Ring loss(정규화 접근 방식) 제안

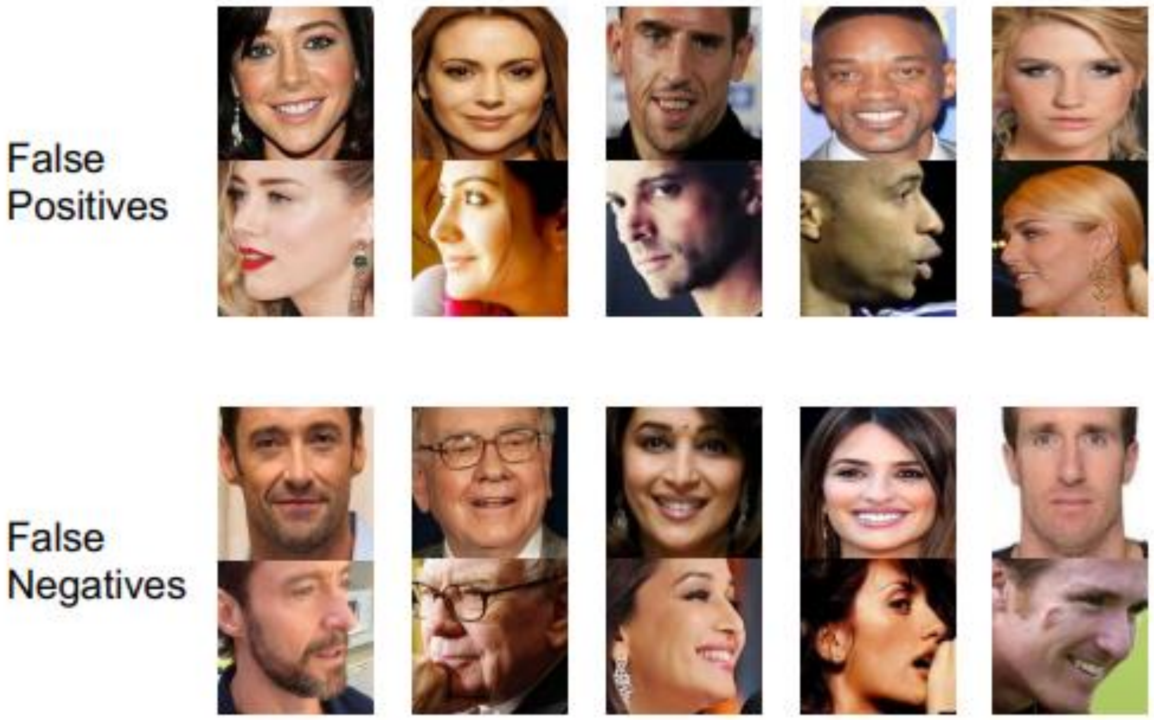
Strategy
- Softmax normalization 적용 및 convexity를 유치하면서 단위원으로 제한하여 강력한 기능 제공

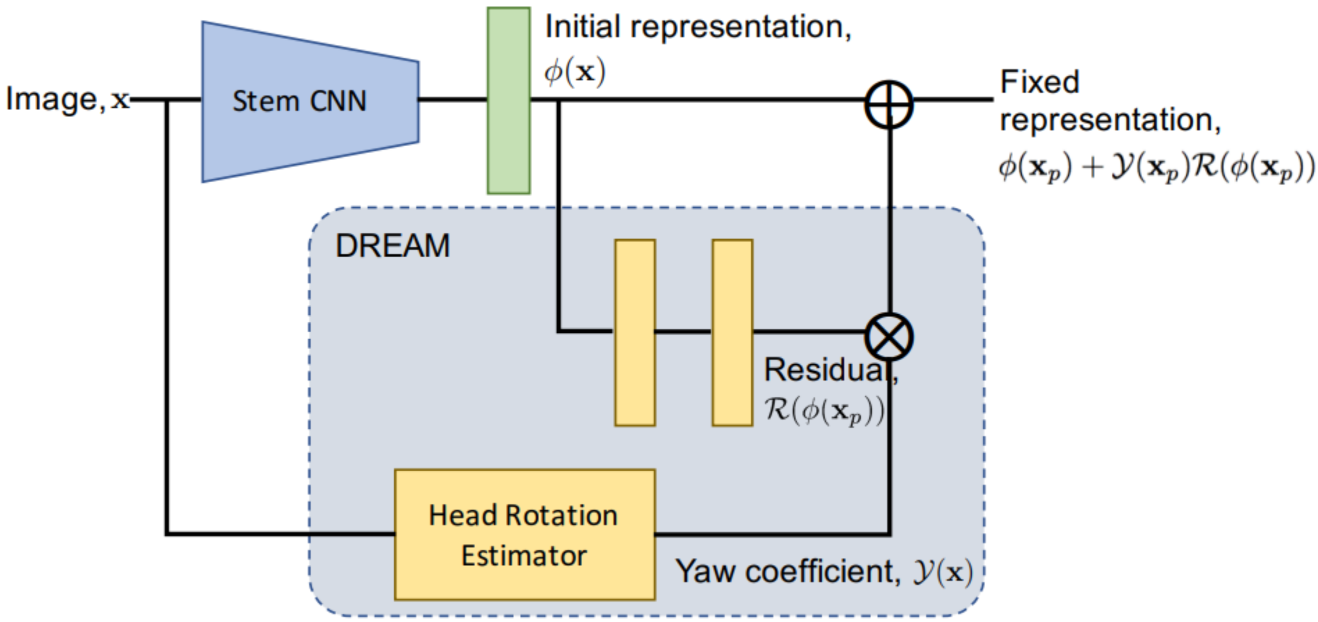
Background technique
: softmax

Dataset
: WebFace, CelebA 1M

[19] Pose-Robust Face Recognition via Deep Residual Equivariant Mapping

Kaidi Cao, Yu Rong, Cheng Li, Xiaoou Tang, Chen Change Loy





[19] Pose-Robust Face Recognition via Deep Residual Equivariant Mapping

Kaidi Cao, Yu Rong, Cheng Li, Xiaoou Tang, Chen Change Loy

Problem

: Pose-Robust Face Recognition

Strategy

- 다양한 pose에서 변하지 않는 geometry feature를 찾자!
  → 프로필 얼굴 pose을 표준(정면) 포즈로 변환
- DREAM Block(Deep Residual EquivAriant Mapping)
  - Input representation에 동적으로 residuals 추가
  - Profile과 frontal의 불일치 연결

Background technique

: ResNet, CNN

Dataset

: CFP, IJB-A, MS-Celeb-1M

[20] Disentangling Features in 3D Face Shapes for Joint Face Reconstruction and Recognition

Feng Liu, Ronghang Zhu, Dan Zeng, Qijun Zhao, Xiaoming Liu

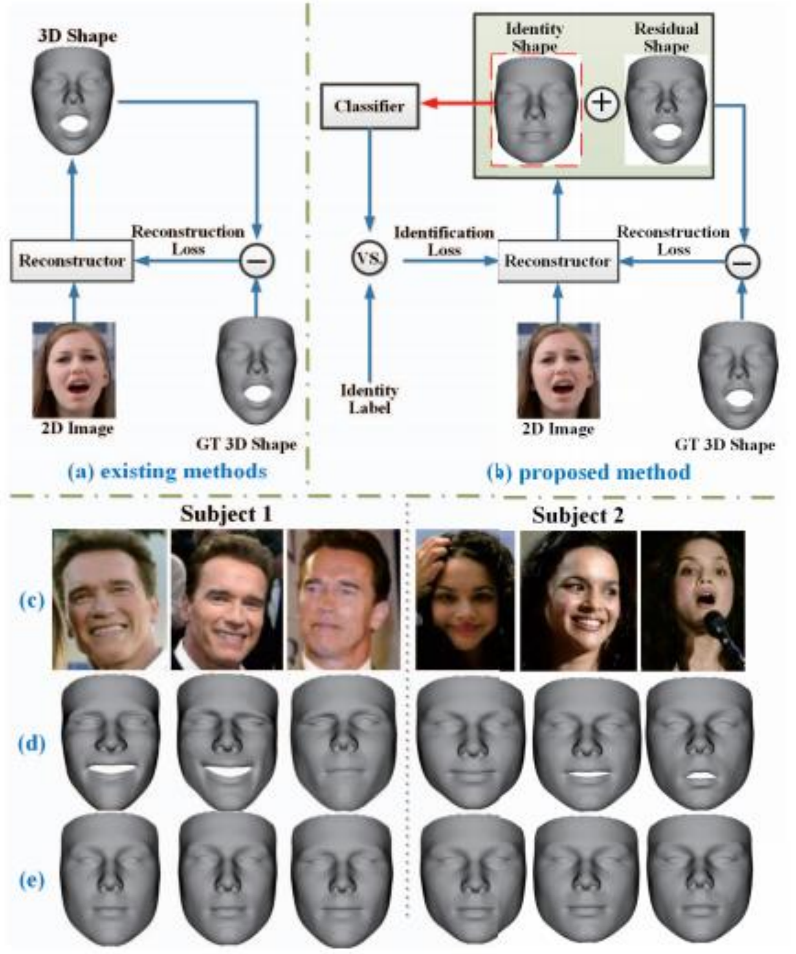
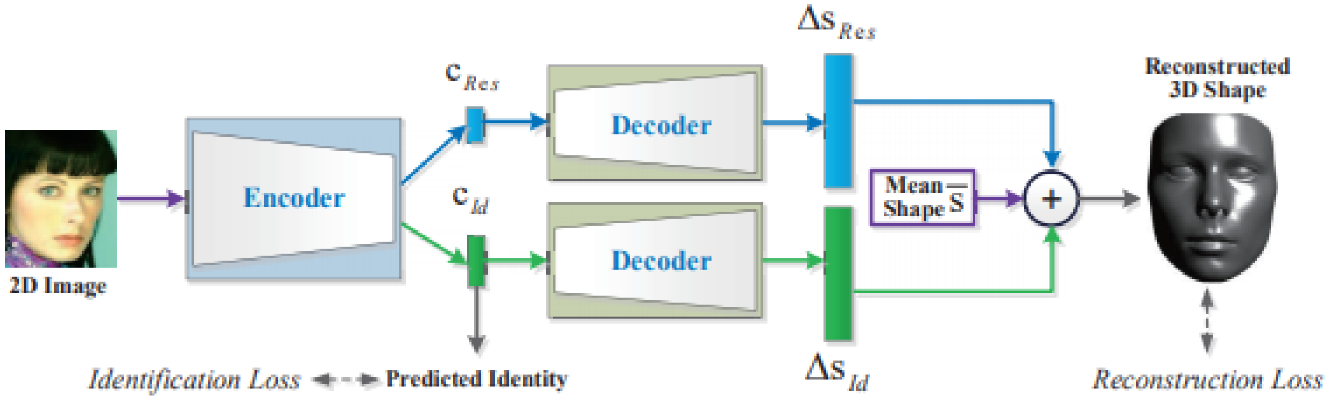


Figure 1. Comparison between the learning process of (a) existing methods and (b) our proposed method. GT denotes Ground Truth. (d) and (e) are 3D face shapes and disentangled identity shapes reconstructed by our method for the images in (c) from LFW [15].



201810957 염지현

[20] Disentangling Features in 3D Face Shapes for Joint Face Reconstruction and Recognition

Feng Liu, Ronghang Zhu, Dan Zeng, Qijun Zhao, Xiaoming Liu

Problem
: 3D Face Reconstruction and Recognition

Strategy
- Encoder-decoder network to disentangle shape features during 3D face reconsturciton
  - 2D image에서 조밀한 3D face shape regression
  - 복합 3D face shape을 기반으로 3D face shapes identity and residual을 명시적, 개별적으로 다룸

Background technique
: CNN

Dataset
: MICC, BU3DFE, LFW, YTF

[21] CosFace: Large Margin Cosine Loss for Deep Face Recognition

Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, Wei Liu

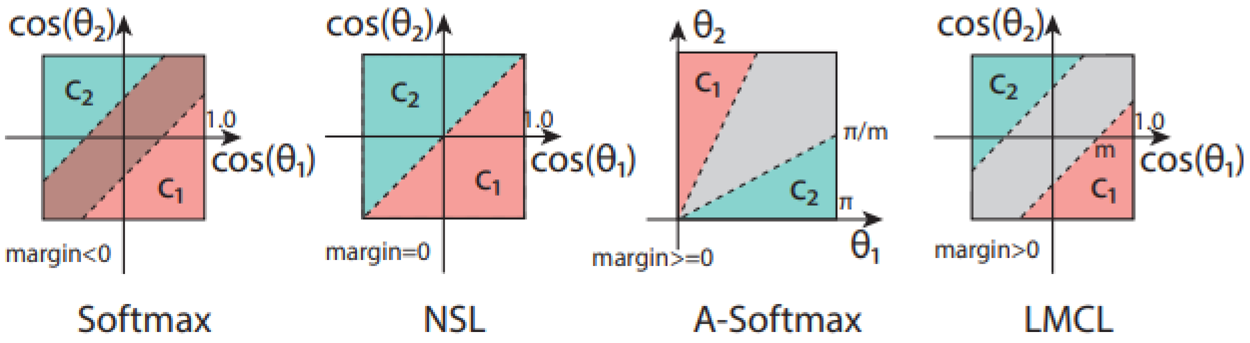
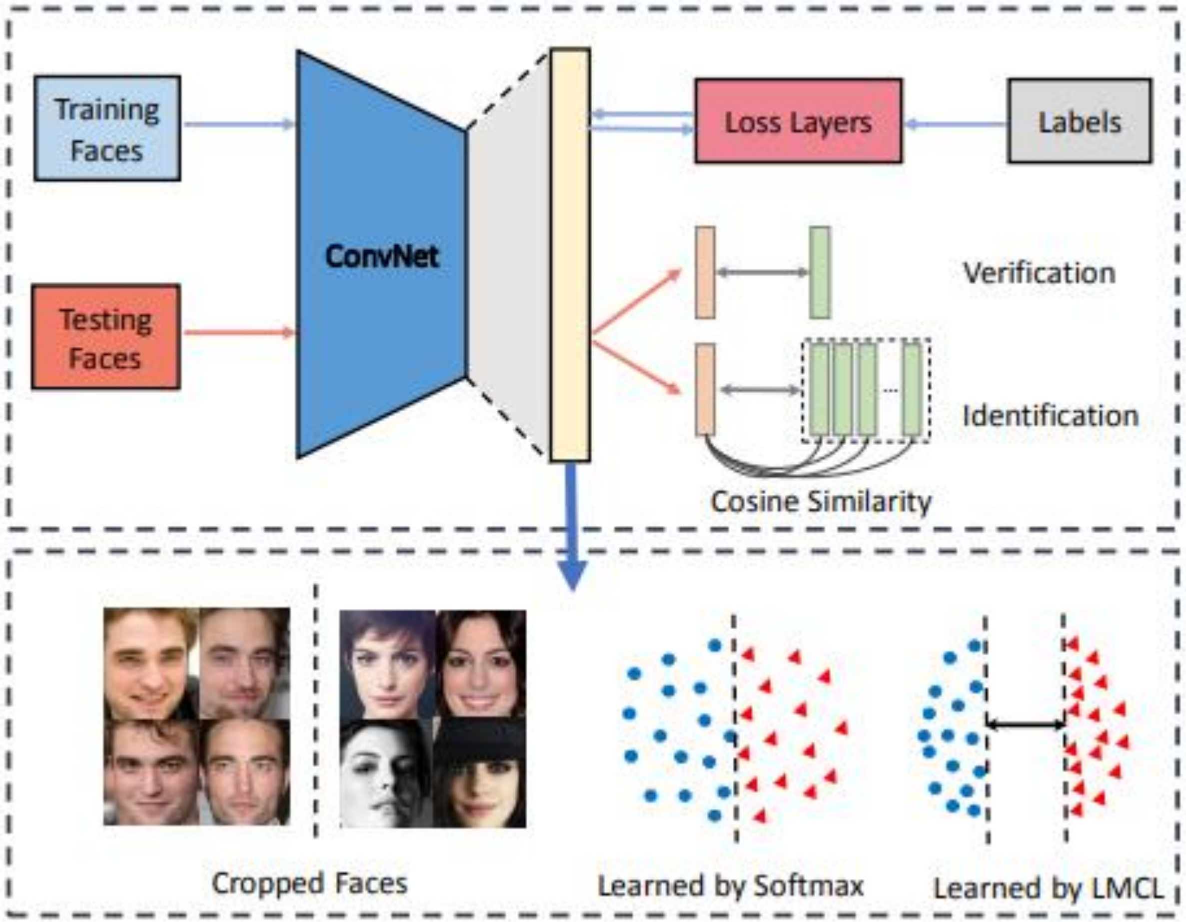


Figure 2. The comparison of decision margins for different loss functions the binary-classes scenarios. Dashed line represents decision boundary, and gray areas are decision margins.



201810957 염지현

[21] CosFace: Large Margin Cosine Loss for Deep Face Recognition

Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, Wei Liu

Problem

: softmax loss → Face Recognition 시 구별 X

Strategy

- CosFace framework – LMCL(Large Margin Cosine Loss)

Background technique

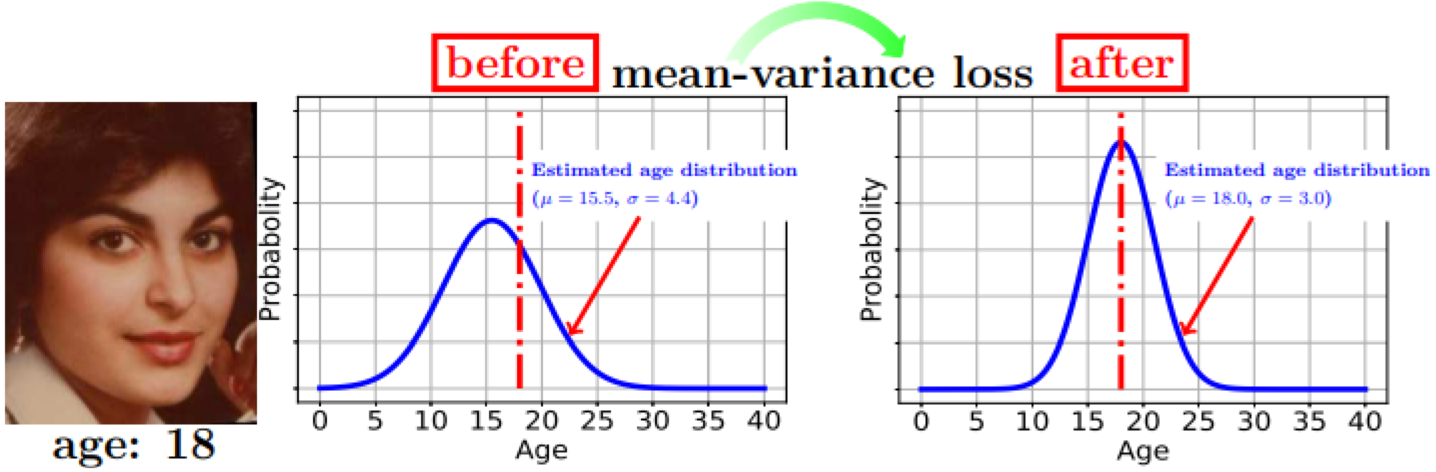
: CNN

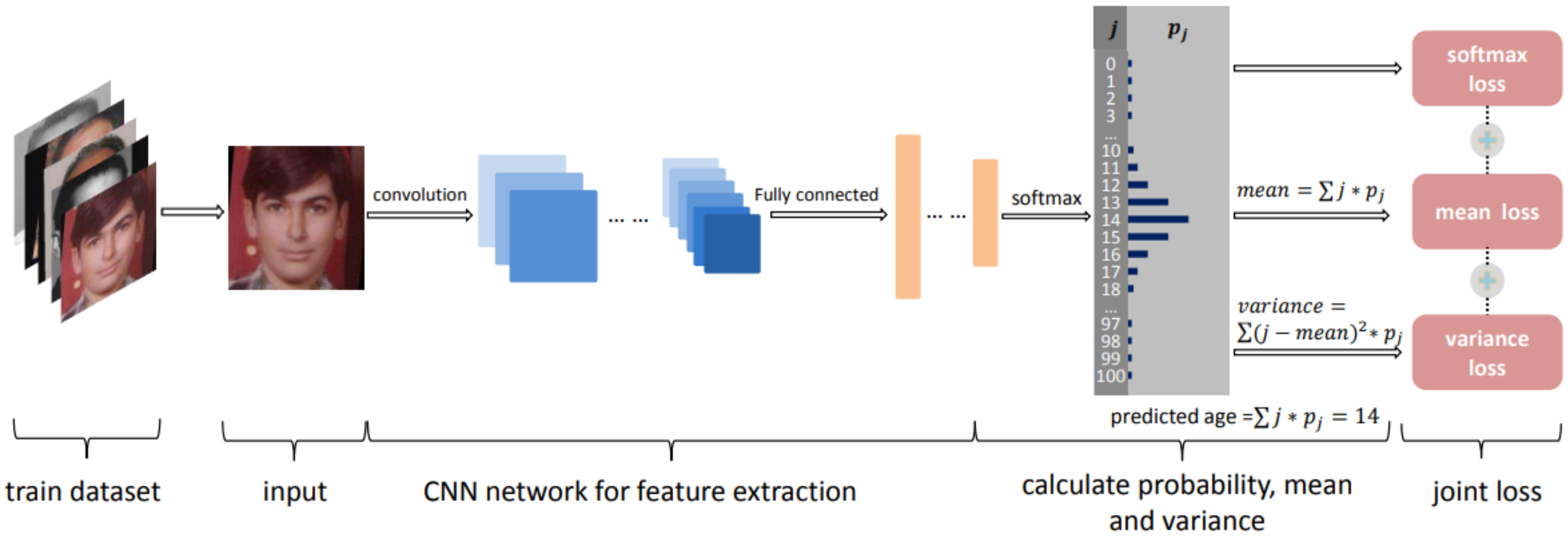
Dataset

: LFW, YTF

[22] Mean-Variance Loss for Deep Age Estimation From a Face

Hongyu Pan, Hu Han, Shiguang Shan, Xilin Chen





201810957 염지현

[22] Mean-Variance Loss for Deep Age Estimation From a Face

Hongyu Pan, Hu Han, Shiguang Shan, Xilin Chen

Problem
: Mean-Variance Loss for Deep Age Estimation from a Face

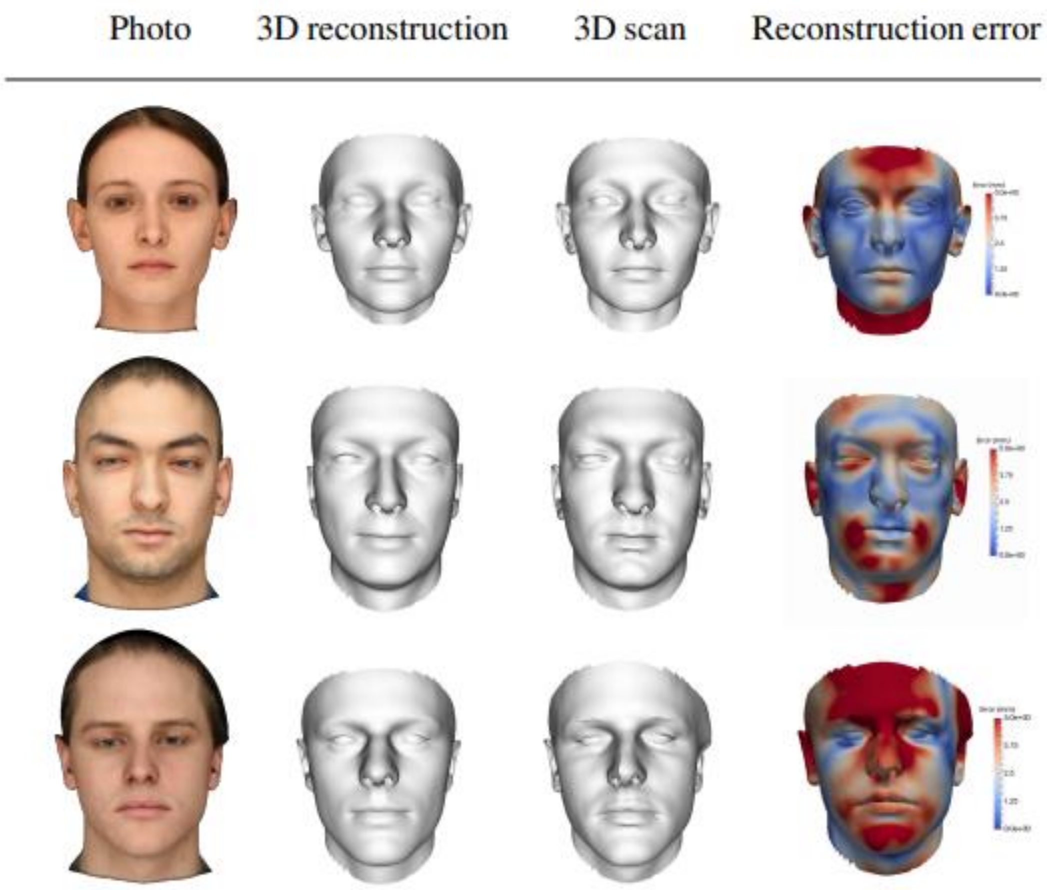
Strategy
- Mean-Variance Loss(distribution learning을 통해 robust age estimation 가능)

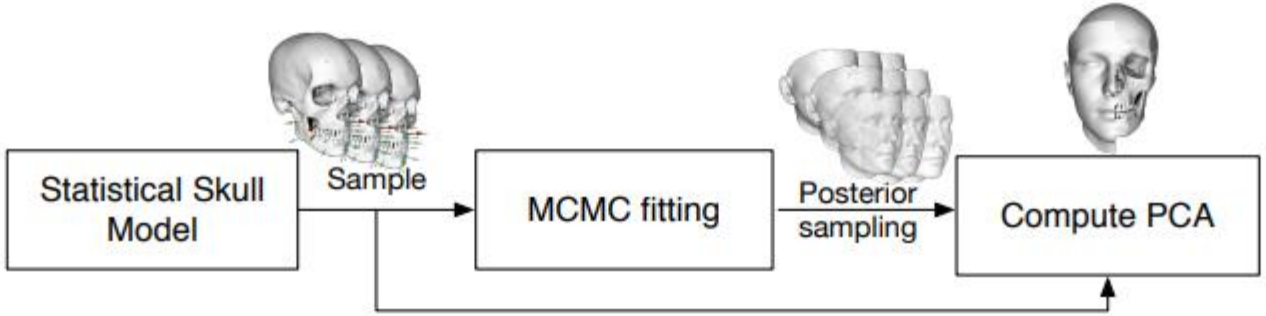
Background technique
: CNN

Dataset
: FG-NET, MORPH Album II, CLAP2016, AADB

[23] Probabilistic Joint Face-Skull Modelling for Facial Reconstruction

Dennis Madsen, Marcel Lüthi, Andreas Schneider, Thomas Vetter





201810957 염지현

[23] Probabilistic Joint Face-Skull Modelling for Facial Reconstruction

Dennis Madsen, Marcel Lüthi, Andreas Schneider, Thomas Vetter

Problem

: Face-Skull Modeling for Facial Reconstruct

Strategy

- Face-shape Model, skull shape Model, tissue depth maker information
- Hastings Algorithm과 함께 tissue depth distribution과 일치하는 posterior distribution of face 계산

Background technique

: MCMC(Markov Chain Monte Carlo)

Dataset

: 자체적인 dataset 생성

[24] Deep Face Detector Adaptation without Negative Transfer or Catastrophic Forgetting

Muhammad Abdullah Jamal, Haoxiang Li, Boqing Gong

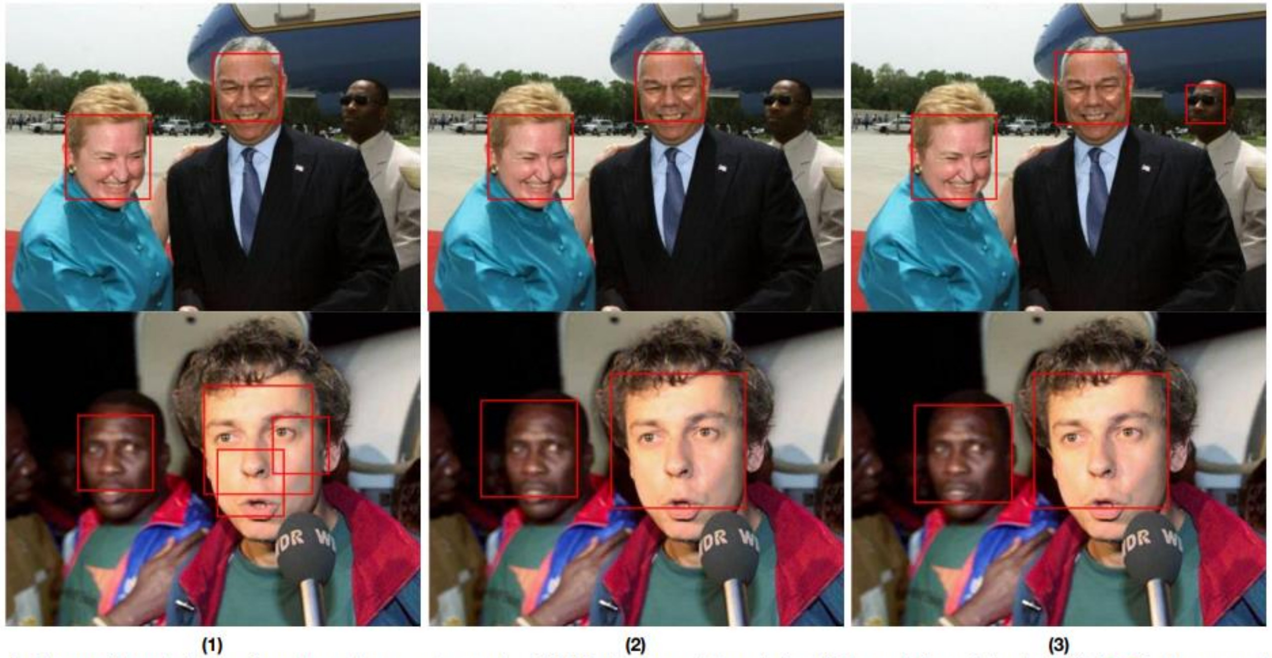


Figure 1. From left to right are face detection results on the FDDB dataset with **a state-of-the-art face detector (1)** [1, 2], the same detector but adapted by our method to the target domain (FDDB) **with no data annotation (2)**, and **with some data annotations (3)**.

[24] Deep Face Detector Adaptation without Negative Transfer or Catastrophic Forgetting

Muhammad Abdullah Jamal, Haoxiang Li, Boqing Gong

Problem
: Face Detection

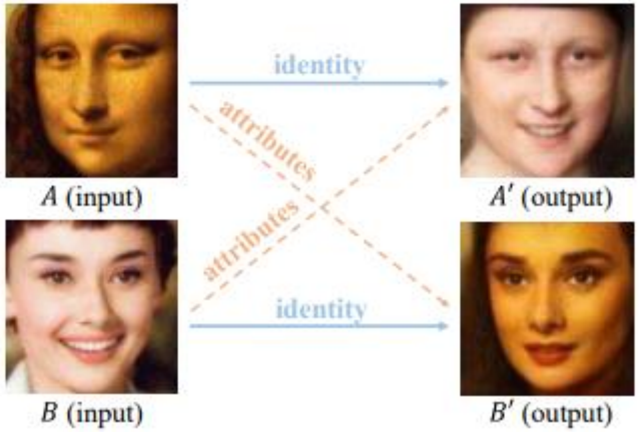
Strategy
- Residual loss : to avoid negative transfer
- Residual classifier : to alleviate catastrophic forgetting

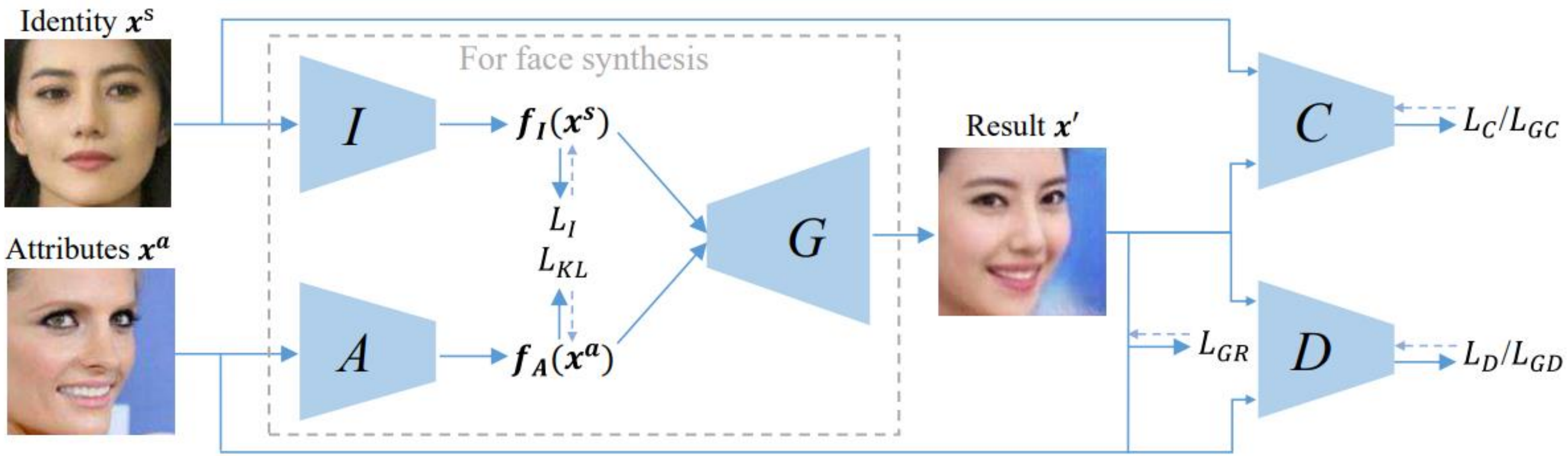
Background technique
: face detector

Dataset
: FDDB

[25] Towards Open-Set Identity Preserving Face Synthesis

Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, Gang Hua





201810957 염지현

[25] Towards Open-Set Identity Preserving Face Synthesis

Jianmin Bao, Dong Chen, Fang Wen, Houqiang Li, Gang Hua

Problem
: Face Synthesis(preserving identity)

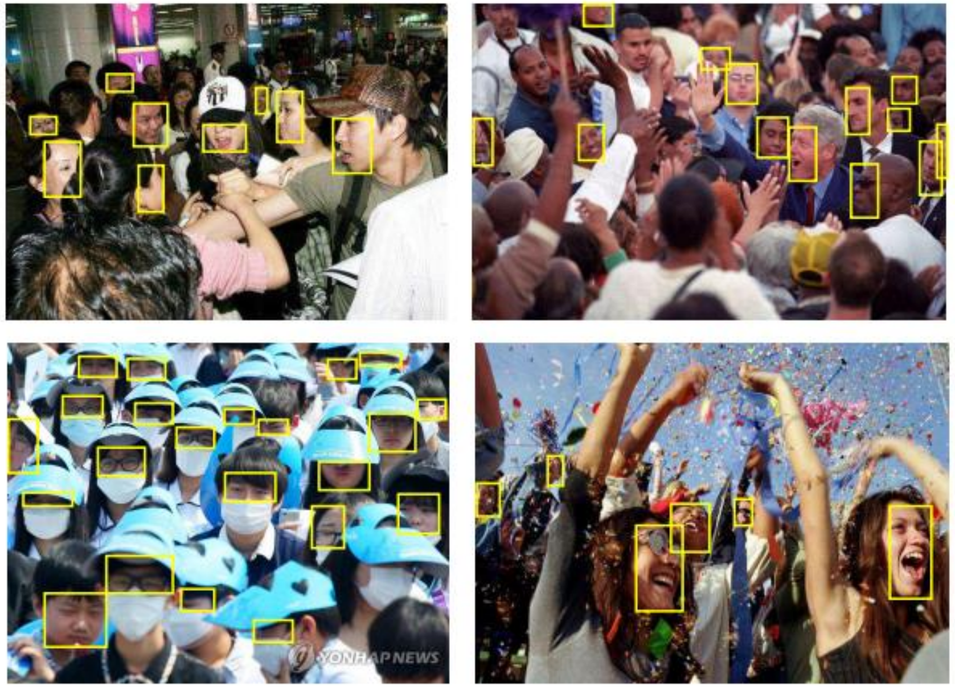
Strategy
- Open-set Identity Preserving GAN : Identity vector와 attribute vector recombine
  (attribute : pose, emotion, illumination, background)

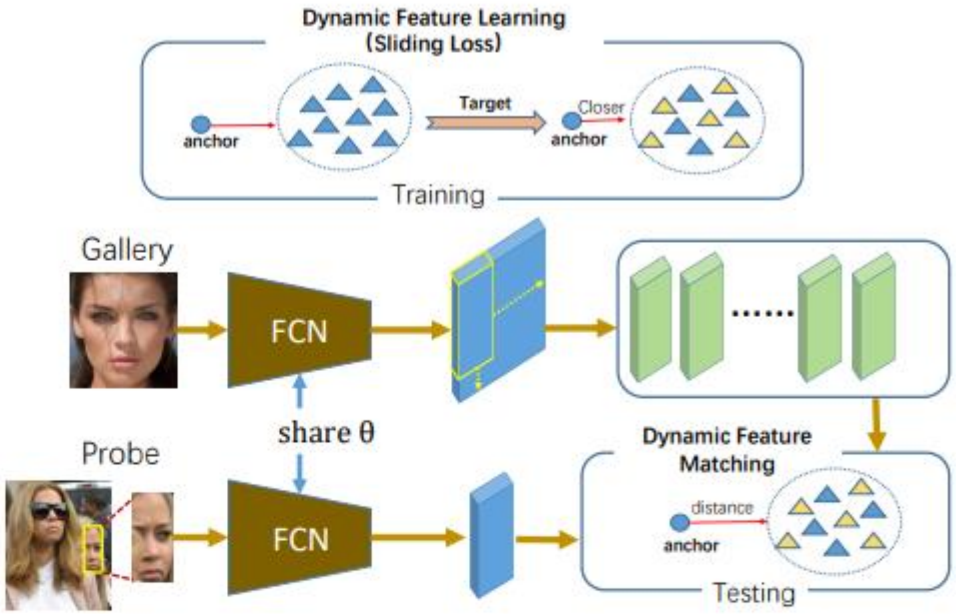
Background technique
: GAN

Dataset
: FaceScrub, CASIA-WebFace, MS-Celeb-1M

[26] Dynamic Feature Learning for Partial Face Recognition

Lingxiao He, Haiqing Li, Qi Zhang, Zhenan Sun





201810957 염지현

[26] Dynamic Feature Learning for Partial Face Recognition

Lingxiao He, Haiqing Li, Qi Zhang, Zhenan Sun

Problem
: Partial Face Recognition

Strategy
- DFM(FCN + SRC(Sparse Representation Classifier))

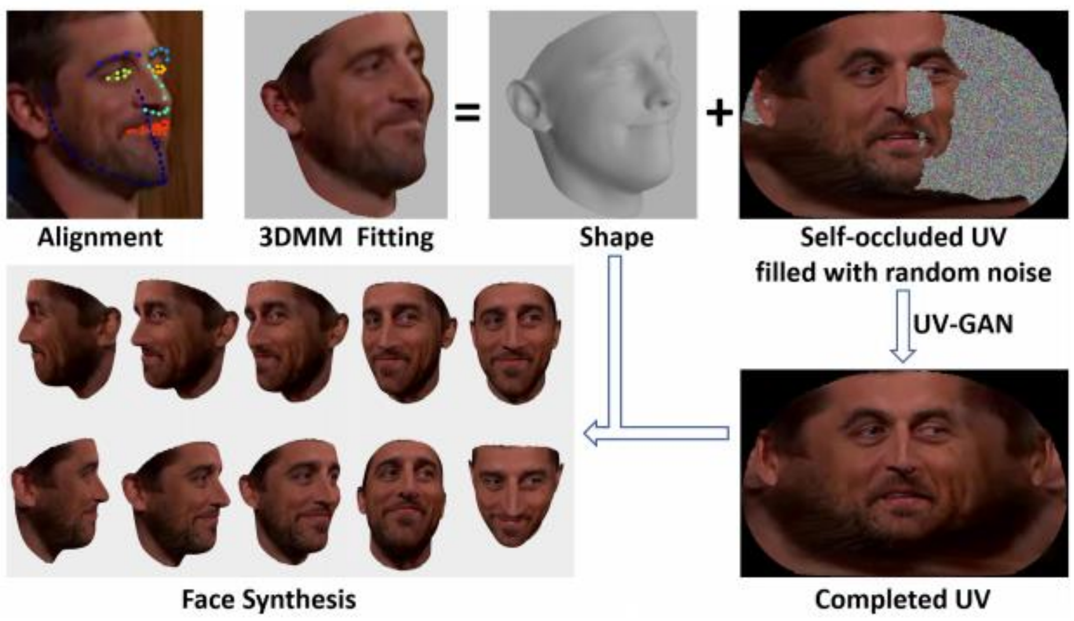
Background technique
: CNN

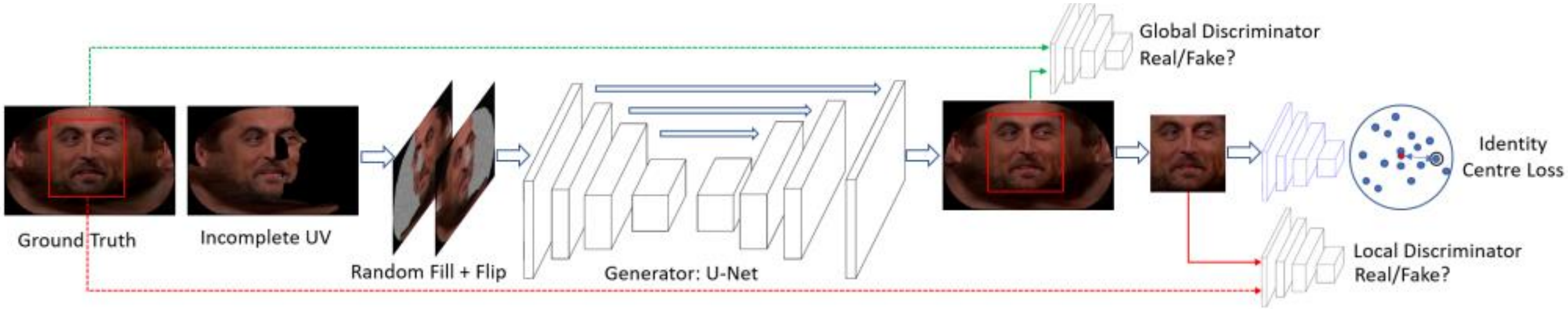
Dataset
: LFW, YTF, CASIA-NIR-Distance

201810957 염지현

[27] UV-GAN: Adversarial Facial UV Map Completion for Pose-Invariant Face Recognition

Jiankang Deng, Shiyang Cheng, Niannan Xue, Yuxiang Zhou, Stefanos Zafeiriou





201810957 염지현

[27] UV-GAN: Adversarial Facial UV Map Completion for Pose-Invariant Face Recognition

Jiankang Deng, Shiyang Cheng, Niannan Xue, Yuxiang Zhou, Stefanos Zafeiriou

Problem

: Facial UV Map 생성(for Pose-invariant Face Recognition)

Strategy

- UV-GAN
  - 3DMM으로 피팅 → 미완성 UV-Map 수집 → UV-GAN을 통해 완전한 UV-Map 생성
    → 3D shape에 face synthesis

Background technique

: GAN

Dataset

: CASIA

[28] A Face-to-Face Neural Conversation Model

Hang Chu, Daiqing Li, Sanja Fidler

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.		and then? and then i went home alone.	we drank a bottle of champagne.	and then i went to bed.
she doesn't know where he is.		i don't know where he is.	i'm sorry.	i don't know where she is.
and he sleeps only one hour a night.		he's a great man.	he sleeps in the same bed.	he's a good man.
a night that marked the opening of a new chapter.		in world history.	for the future.	in the history of the world.
i hope you're not a hothead like sonny.		he's a good kid.	he's got a lot of something.	he's a good kid.

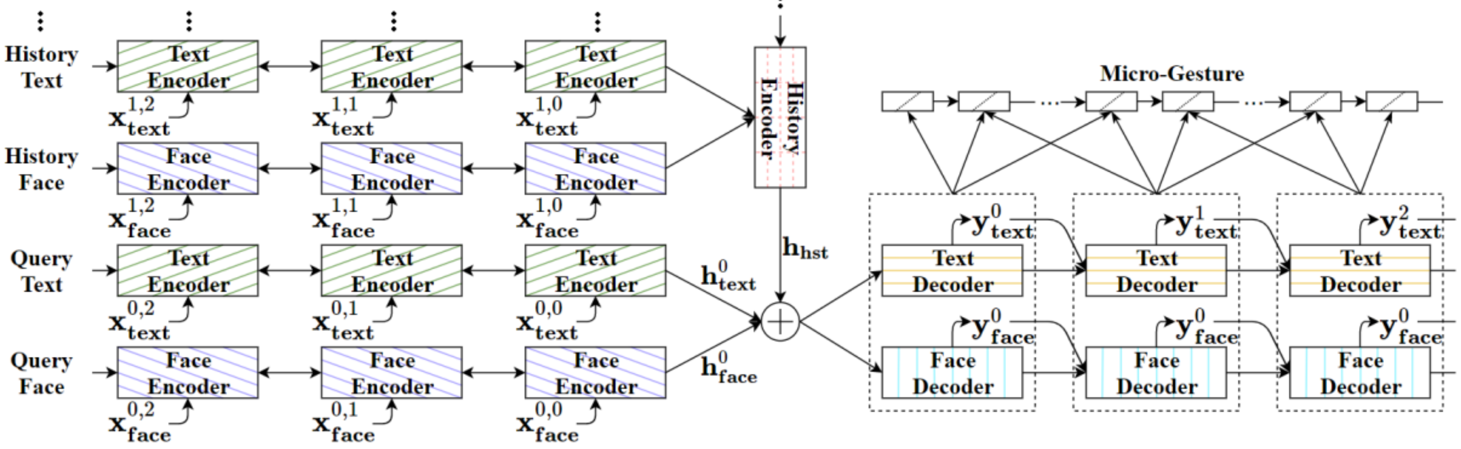


Figure 6: Our face-to-face conversation model. Our model consists of 6 RNNs shown in different colors. First, the text-face sequences of query and conversation history are encoded by text and face encoders (only one history sentence is depicted). History sentence encodings are further encoded by the history encoder. Next, encodings are added to form the context vector, which the text and face decoders are conditioned on. Finally, we generate frame-level, micro-gesture animation controls based on the word decodings.

201810957 염지현

[28] A Face-to-Face Neural Conversation Model

Hang Chu, Daiqing Li, Sanja Fidler

<u>Problem</u>

: Facial gesture를 통해 conversation(text) 생성

<u>Strategy</u>

- A Face=to-Face Neural Conversation Model → Decoder
    - 하위 layer : Generating the verbal response & Coarse facial expression
    - 상위 layer : fills in the subtle gestures

<u>Background technique</u>

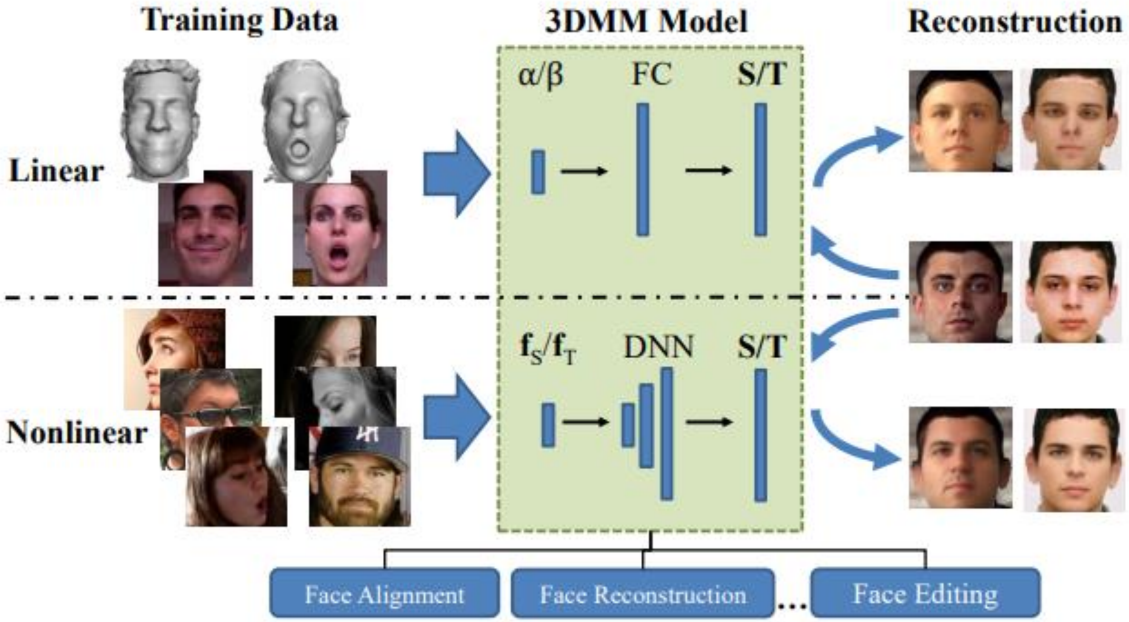
: RNN, GAN

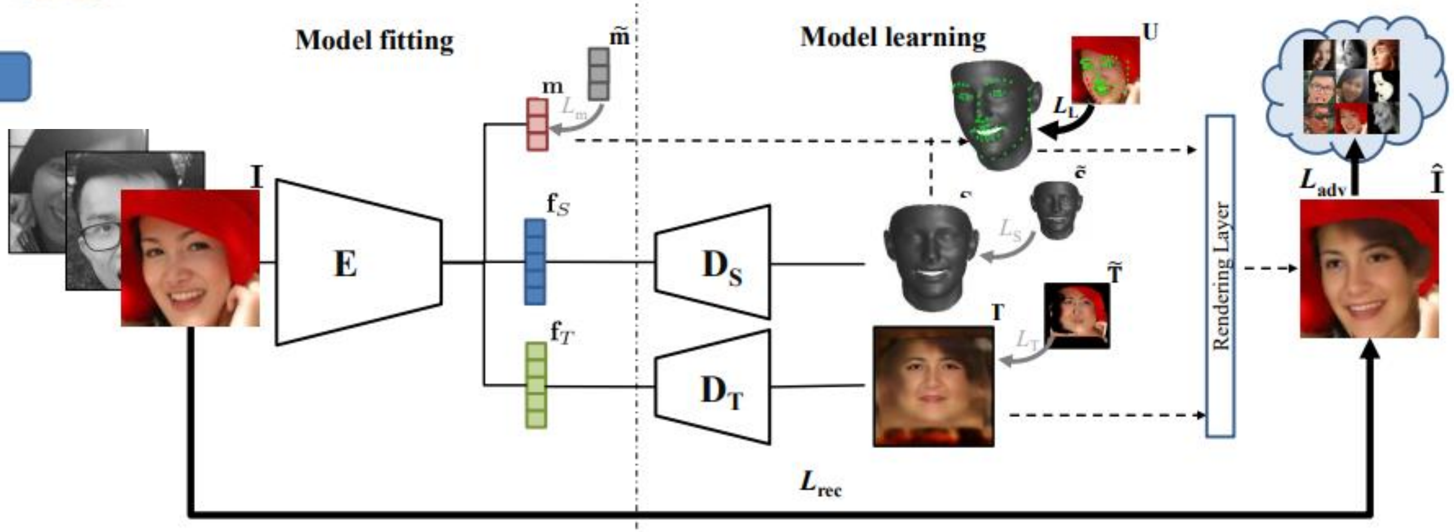
<u>Dataset</u>

: Movie Chat

[29] Nonlinear 3D Face Morphable Model

Luan Tran, Xiaoming Liu





201810957 염지현

[29] Nonlinear 3D Face Morphable Model

Luan Tran, Xiaoming Liu

Problem
: 잘 정제된 2D image 및 3D model 없이 wild한 image로부터 3D Model 만들기

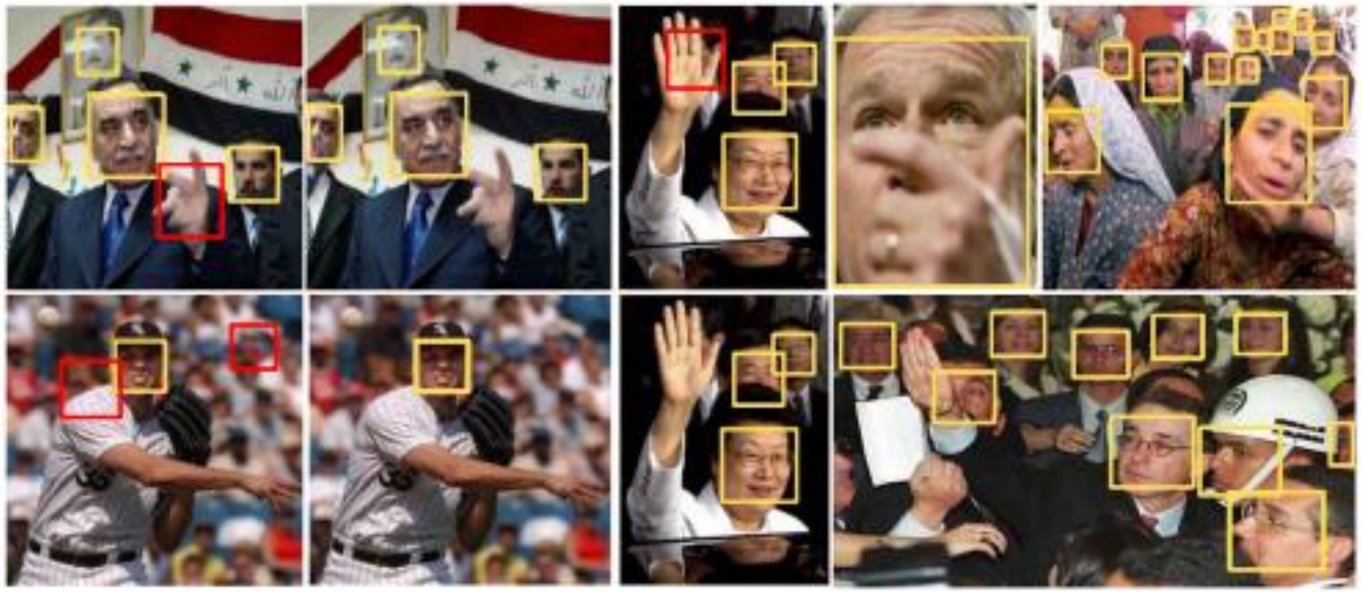
Strategy
- 인코더 : projection, shape, texture parameters 추정
- 디코더(2개) : 각각 shape & texture parameters과 3D shape & texture mapping

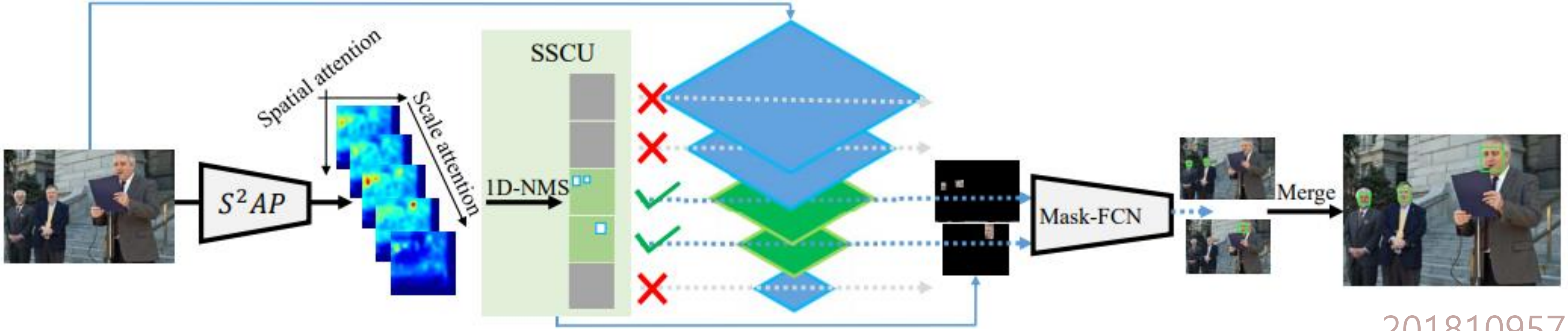
Background technique
: CNN

Dataset
: AFLW2000

[30] Beyond Trade-Off: Accelerate FCN-Based Face Detector With Higher Accuracy

Guanglu Song, Yu Liu, Ming Jiang, Yujie Wang, Junjie Yan, Biao Leng





[30] Beyond Trade-Off: Accelerate FCN-Based Face Detector With Higher Accuracy

Guanglu Song, Yu Liu, Ming Jiang, Yujie Wang, Junjie Yan, Biao Leng

Problem

: Facial Detection

Strategy

- S^2AP-FCN을 기반으로 scale과 space orthogonal direction으로 분리
  (specific scale + valid location in image pyramid)

Background technique

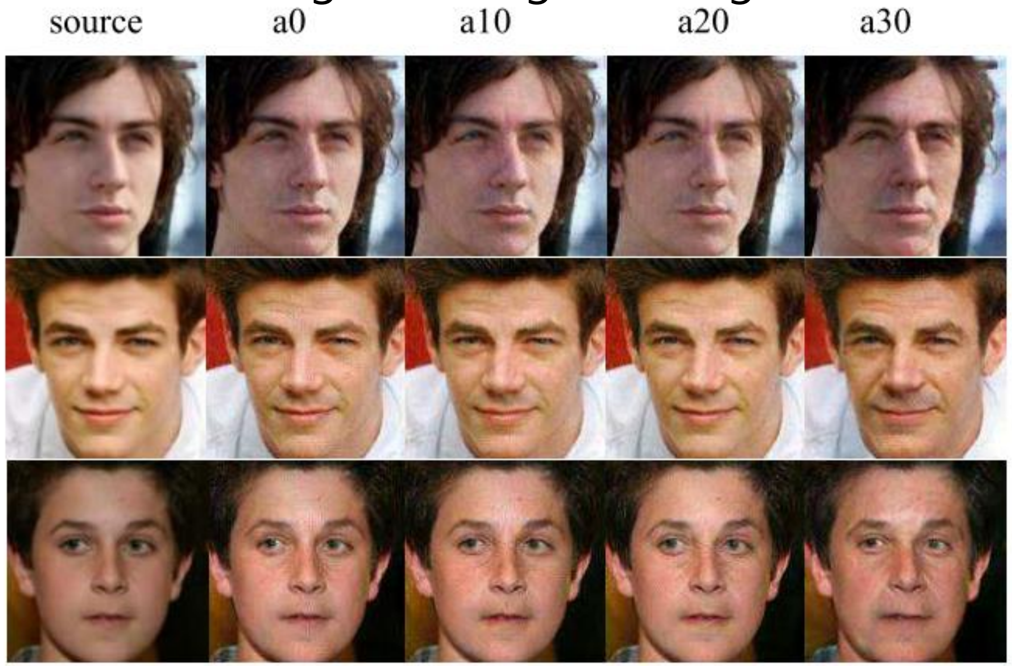
: CNN

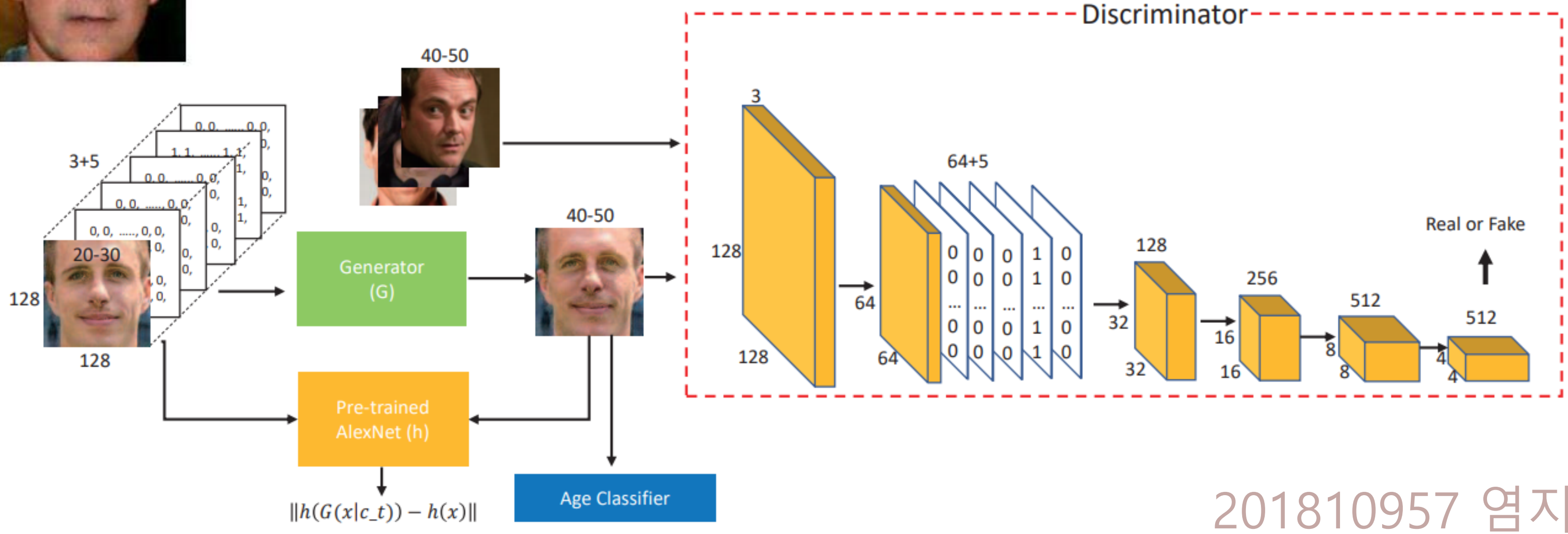
Dataset

: FDDB, AFW, MALF

[31] Face Aging With Identity-Preserved Conditional Generative Adversarial Networks

Zongwei Wang, Xu Tang, Weixin Luo, Shenghua Gao





201810957 염지현

[31] Face Aging With Identity-Preserved Conditional Generative Adversarial Networks

Zongwei Wang, Xu Tang, Weixin Luo, Shenghua Gao

Problem
: Face Aging with Identity-Preserved

Strategy
- IPCGANs(Identity-Preserve Conditional GANs)
  - Target age group 내에 노화 패턴을 합성
  - Identity-Preserve module로 Identity 보전(특징을 유사하게 만들기)
  - Discriminator : aged face와 target age의 일치성 확인

Background technique
: GAN

Dataset
: CACD

[32] Deep Semantic Face Deblurring

Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, Ming-Hsuan Yang

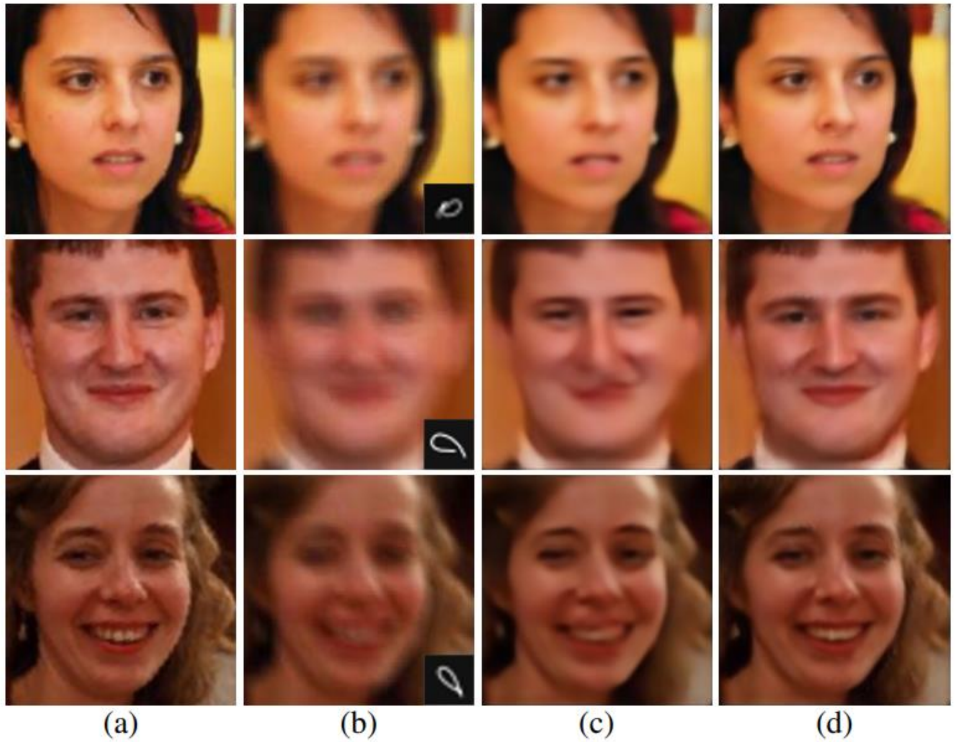


Figure 1. **Face deblurring results.** We exploit the semantic information of face within an end-to-end deep CNN for face image deblurring. (a) Ground truth images (b) Blurred images (c) Ours w/o semantics (d) Ours w/ semantics.

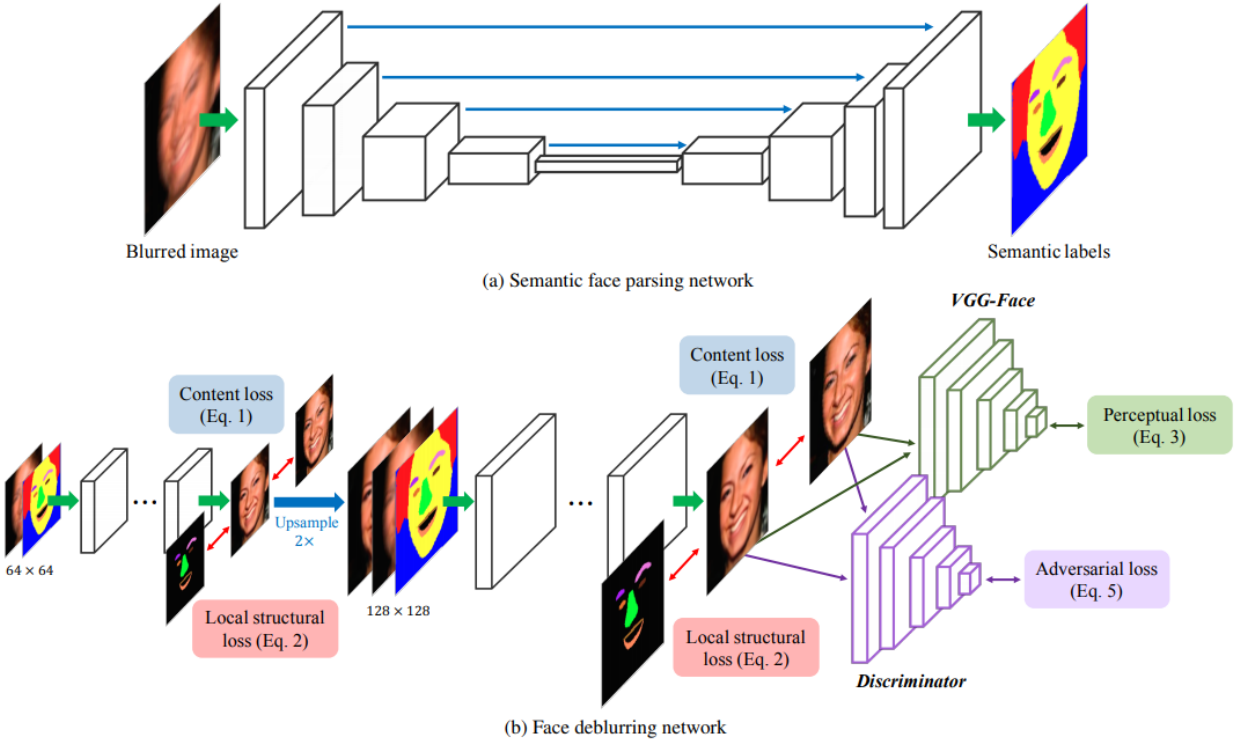


Figure 2. **Overview of the proposed semantic face deblurring network.** The proposed network consists of two sub-networks: a semantic face parsing network and a multi-scale deblurring network. The face parsing network generates the semantic labels of the input blurred images. The multi-scale deblurring network has two scales. We concatenate the blurred image and semantic labels as the input to the first scale. At the second scale, the input is the concatenation of the upsampled deblurred image from the first scale, the blurred image and the corresponding semantic labels. Each scale of the deblurring network receives the supervision from the pixel-wise content loss and local structural losses. We impose the perceptual and adversarial losses at the output of the second scale.

[32] Deep Semantic Face Deblurring

Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, Ming-Hsuan Yang

Problem
: Face Deblurring

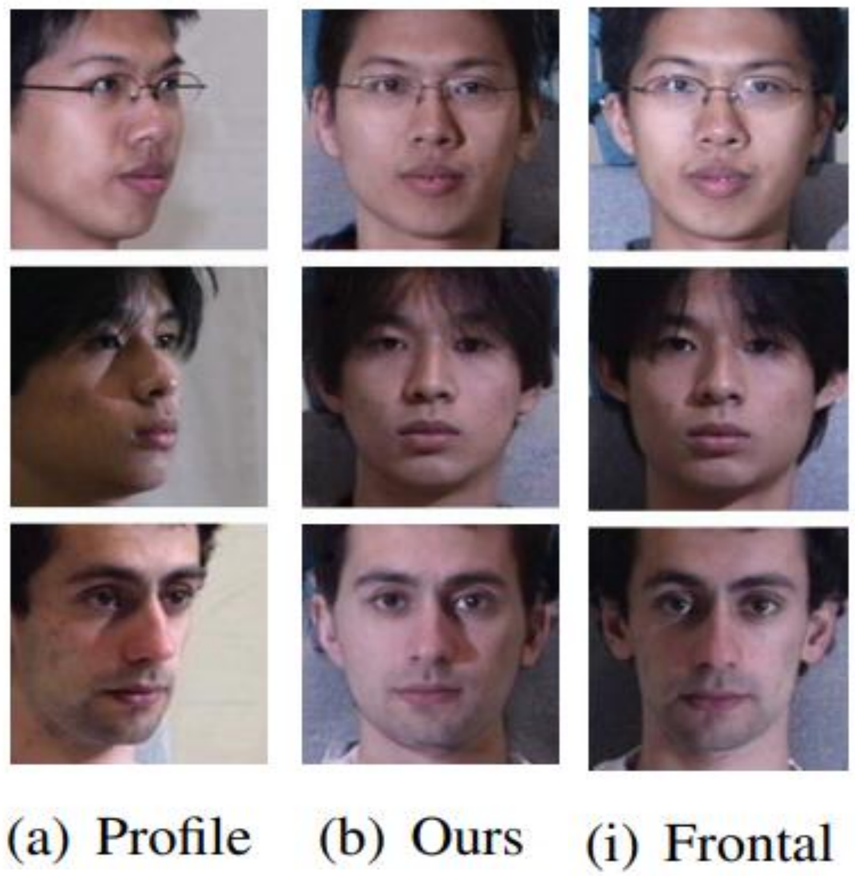
Strategy
- Semantic Face Deblurring Network: several key semantic components를 이용
  - Face Parsing network: semantic labels 생성
  - Sub-Network
    - Blurring image + semantic label → Deblurring image
    - Discriminator → Perceptual loss, Adversarial loss

Background technique
: GAN, CNN

Dataset
: Helen, CMU PIE, CelebA

[33] Pose-Guided Photorealistic Face Rotation

ibo Hu, Xiang Wu, Bing Yu, Ran He, Zhenan Sun



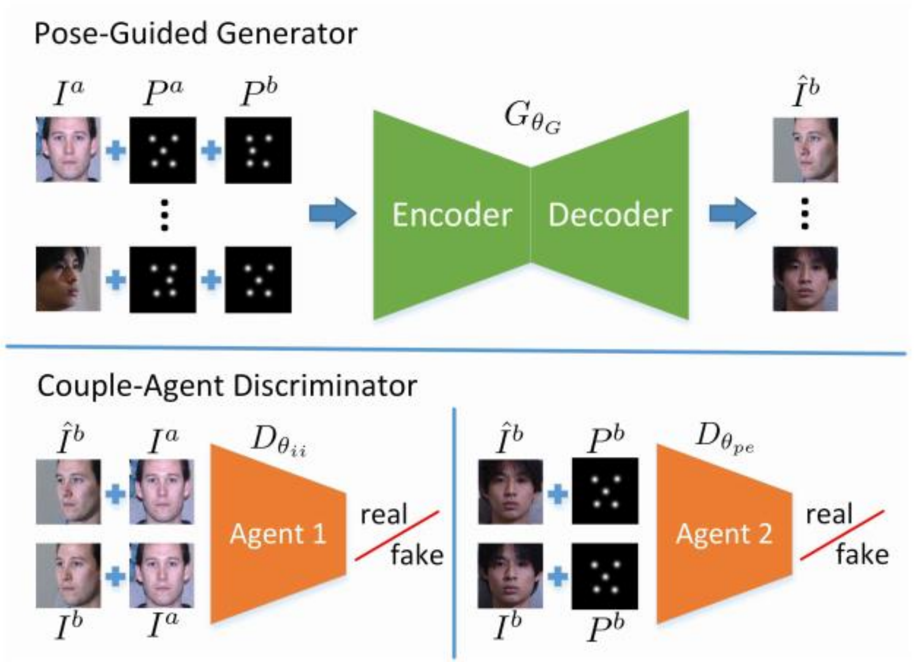


Figure 1. The framework of CAPG-GAN. The generator incorporates pose information by landmark heatmaps in the learning process. The couple-agent discriminator distinguishes generated pairs from ground-truth pairs for photorealistic synthesization.

201810957 염지현

[33] Pose-Guided Photorealistic Face Rotation

ibo Hu, Xiang Wu, Bing Yu, Ran He, Zhenan Sun

Problem

: Face Rotation

Strategy

- CAPG-GAN(Couple-Agent Pose-Guided GAN)
  - Head pose information → Facial Landmark Heatmaps를 통해 인코딩 됨
  - Pose-Guided Generator → mask image 생성
  - Couple-Agent Discriminator
    - Agent1: source image와 generated image 판별
    - Agent2: Pose

Background technique

: GAN

Dataset

: Multi-PIE, LFW

[34] Super-FAN: Integrated Facial Landmark Localization and Super-Resolution of Real-World Low Resolution Faces in Arbitrary Poses With GANs

Adrian Bulat, Georgios Tzimiropoulos

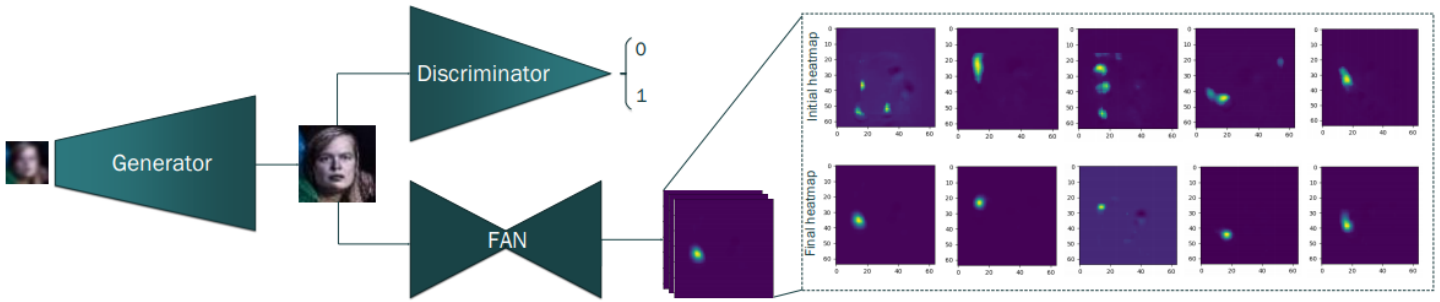


Figure 2: The proposed Super-FAN architecture comprises three connected networks: the first network is a newly proposed Super-resolution network (see sub-section 4.1). The second network is a WGAN-based discriminator used to distinguish between the super-resolved and the original HR image (see sub-section 4.2). The third network is FAN, a face alignment network for localizing the facial landmarks on the super-resolved facial image and improving super-resolution through a newly-introduced heatmap loss (see sub-section 4.3).

201810957 염지현

[34] Super-FAN: Integrated Facial Landmark Localization and Super-Resolution of Real-World Low Resolution Faces in Arbitrary Poses With GANs

Adrian Bulat, Georgios Tzimiropoulos

Problem
: Improve facial resolution and detect facial landmark

Strategy
- Super-FAN
    - First network : Super-resolution network(Generator)
    - Second network : Distinguish between Super-resolved and the original image(Discriminator)
    - Third network : Detect facial landmark on Super-resolved image

Background technique
: GAN

Dataset
: WiderFace

[35] Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors

Adrian Bulat, Georgios Tzimiropoulos

Figure 1. **Annotations are imprecise.** We show annotations of nine annotators on two images of the mouth. Each color indicates a different landmark. Note the inconsistencies of annotations even on the more discriminative landmarks such as the corner of the mouth. This could be harmful to both the training and evaluation of detectors, thus motivating the use of supervisory signals which does not rely on human annotations.

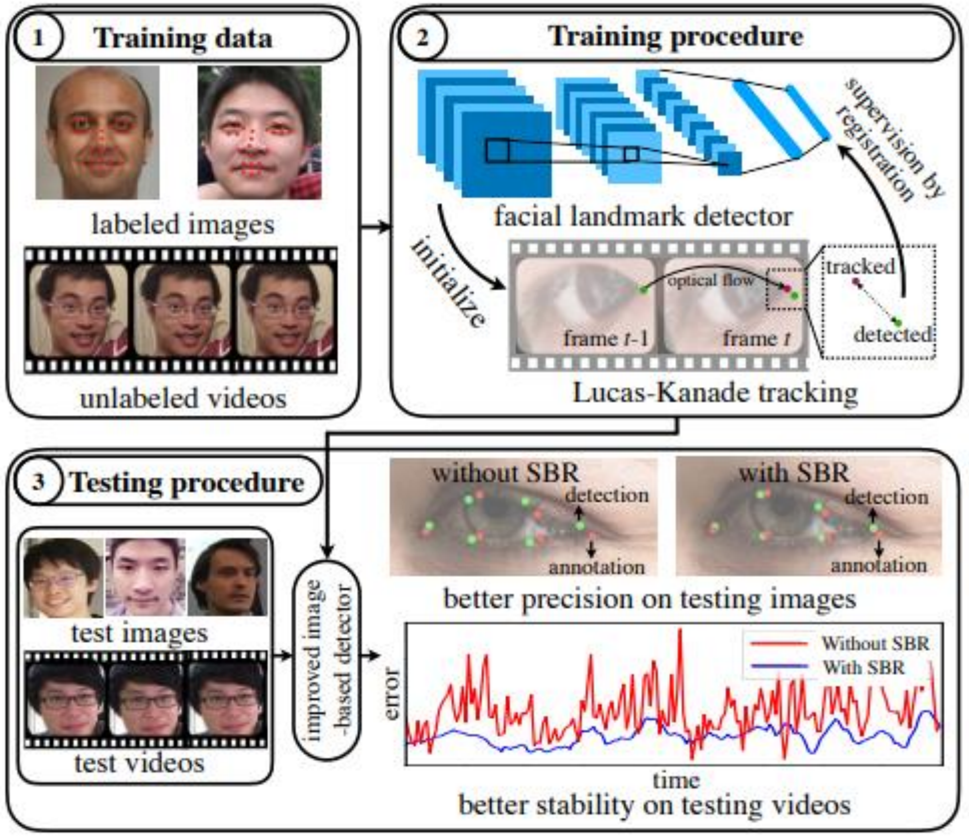


Figure 2. The **supervision-by-registration (SBR) framework** takes labeled images and unlabeled video as input to train an image-based facial landmark detector which is more precise on images/video and also more stable on video.

[35] Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors

Adrian Bulat, Georgios Tzimiropoulos

Problem
: adjacent frame에서 동일한 landmar를 detect

Strategy
- SBR(Supervision-by-Registration)
  - Labeling된 image와 unlabeling된 video 입력
    → 이미지/비디오에서 정확 & 안정되게 landmark detection

Background technique
: CNN

Dataset
: image – 300W, ALFW / video – 300VW, Youtube-Celebrities

[36] Style Aggregated Network for Facial Landmark Detection

Xuanyi Dong, Yan Yan, Wanli Ouyang, Yi Yang

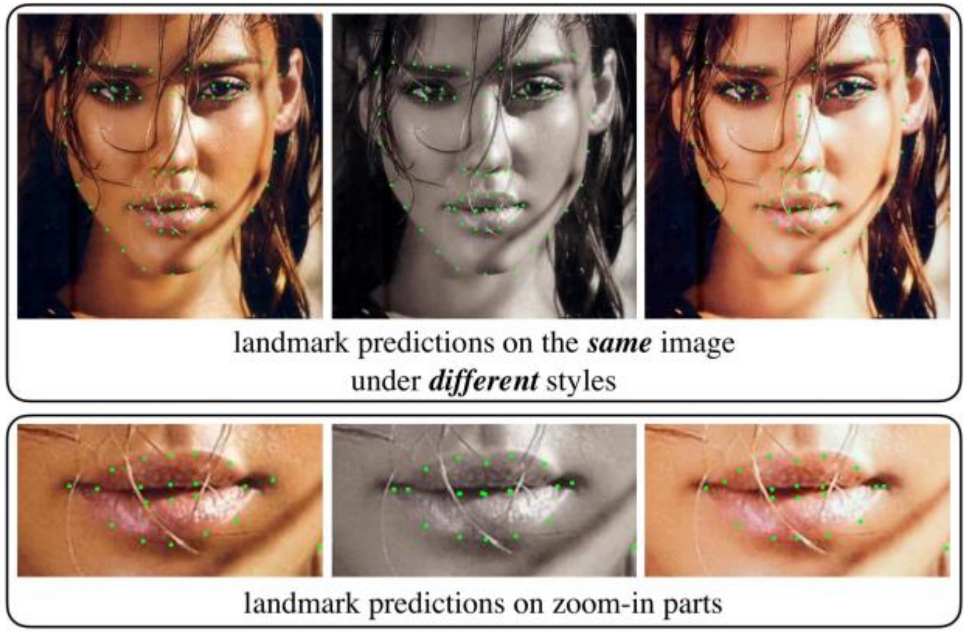
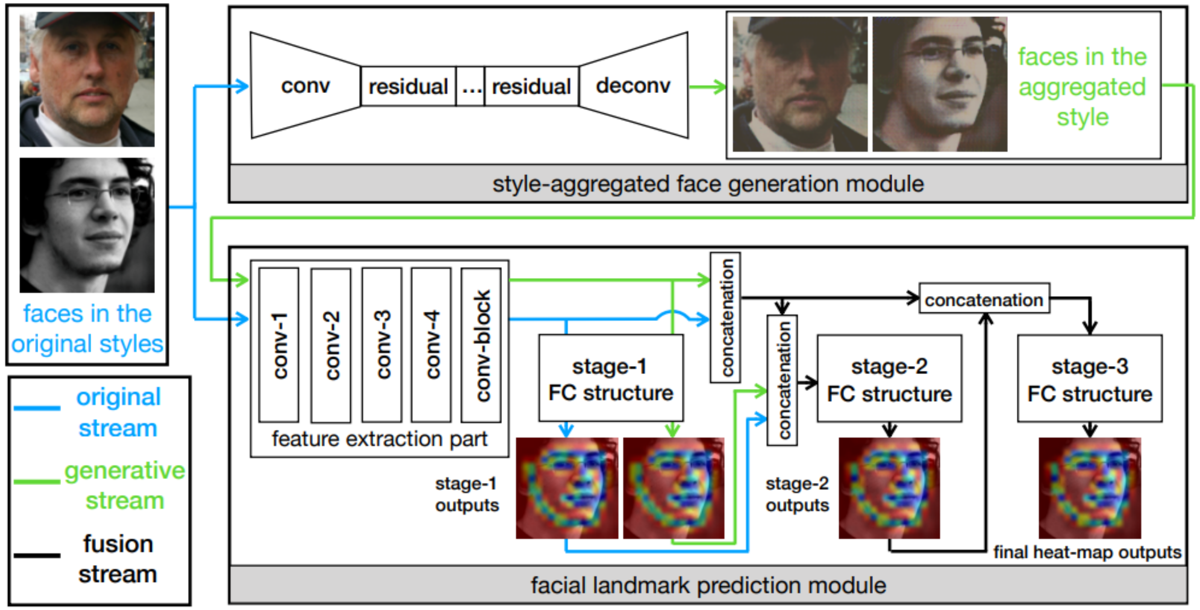


Figure 1. A face image in *three* different styles and the locations of the facial landmarks predicted by a facial landmark detector on them. The image styles, e.g., grayscale vs. color images, light vs. dark, intense vs. dull, can be quite distinct owing to various collection sources. The contents of the above three images are identical. The only difference is the image style. We apply a well-trained facial landmark detector to localize the facial landmarks. The zoom-in parts show the deviation among the predicted locations of the same facial landmarks on different styled images.



[36] Style Aggregated Network for Facial Landmark Detection

Xuanyi Dong, Yan Yan, Wanli Ouyang, Yi Yang

Problem
: 이미지 자체의 style의 변화와 상관없이 landmark detection 하기

Strategy
- SAN(Style Aggregated Network)
  - Style aggregated face generation module : 입력 이미지 다른 style 변화
  - Facial landmark prediction module : original image와 style-aggregated image에서 landmark detection
  - Style-aggregated image를 사용하여 환경적인 변화에 robust한 face image 유지

Background technique
: GAN

Dataset
: ALFW, 300-W

[37] Robust Facial Landmark Detection via a Fully-Convolutional Local-Global Context Network

Daniel Merget, Matthias Rock, Gerhard Rigoll

Figure 5. Best viewed in the digital version. Qualitative results of our approach on the 300-W benchmark. The images are sorted according to their error (top left is best, bottom right is worst). All but the first two images shown are worse than the average case. More precisely, the mean and median errors of the images in rows 1 through 4 are in the 76.5% and 83.6% quantiles of the test set, respectively. The 5th row displays the 10 worst results. Note that the results are displayed in color, but our network only uses grayscale information.

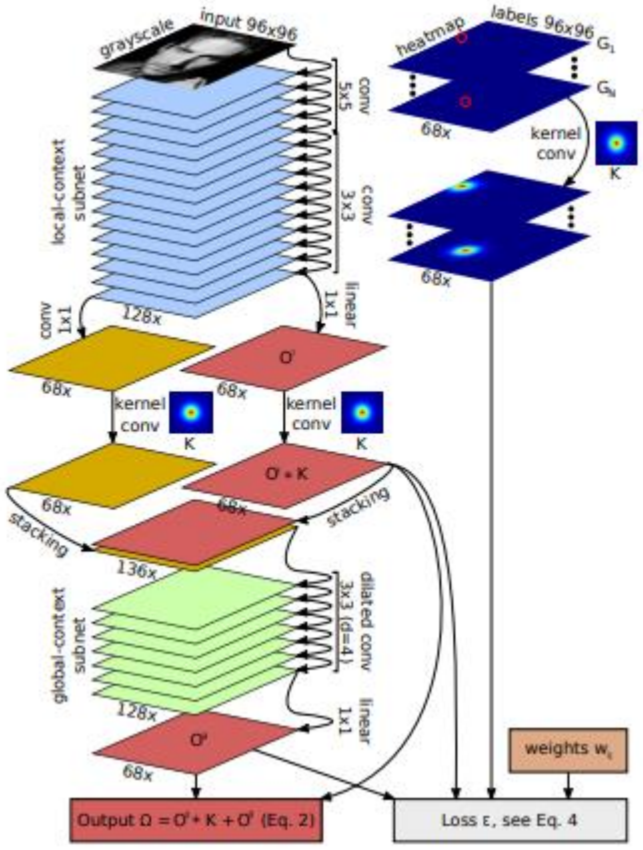


Figure 1. The network architecture used throughout this work.

201810957 염지현

[37] Robust Facial Landmark Detection via a Fully-Convolutional Local-Global Context Network

Daniel Merget, Matthias Rock, Gerhard Rigoll

Problem

: The lack of Global Context(constrained receptive field), Robust Facial Landmark Detection

Strategy

- Fully-convolutional neural network에 global context 직접 추가하기
  - Global context subNet
  - Local context subnet

    → 이 둘을 implicit kernel convolution

Background technique

: CNN

Dataset

: IJB-B, 300-W, AFLW

[38] Learning Facial Action Units From Web Images With Scalable Weakly Supervised Clustering

Kaili Zhao, Wen-Sheng Chu, Aleix M. Martinez

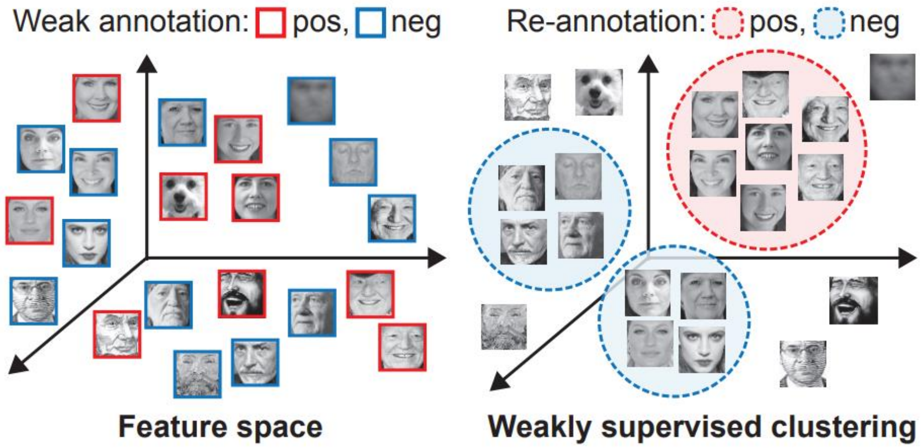


Figure 1. An illustration of weakly-supervised clustering: Unlike the original feature space (**left**) that encloses different semantics and noisy annotations in neighboring images, weakly supervised clustering (**right**) finds a new embedding space where image clusters possess visual-semantic coherence. The proposed approach scales up to a large number of images, and offers outlier/noise pruning by design. Each cluster is re-annotated as the same class by majority voting, and will be included for training AU detectors.

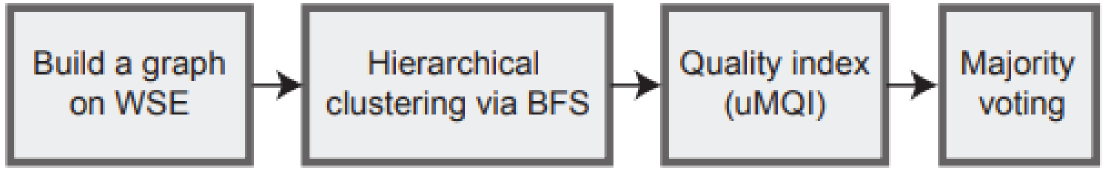


Figure 4. Pipeline for WSC re-annotation

201810957 염지현

[38] Learning Facial Action Units From Web Images With Scalable Weakly Supervised Clustering

Kaili Zhao, Wen-Sheng Chu, Aleix M. Martinez

Problem
: Clustering

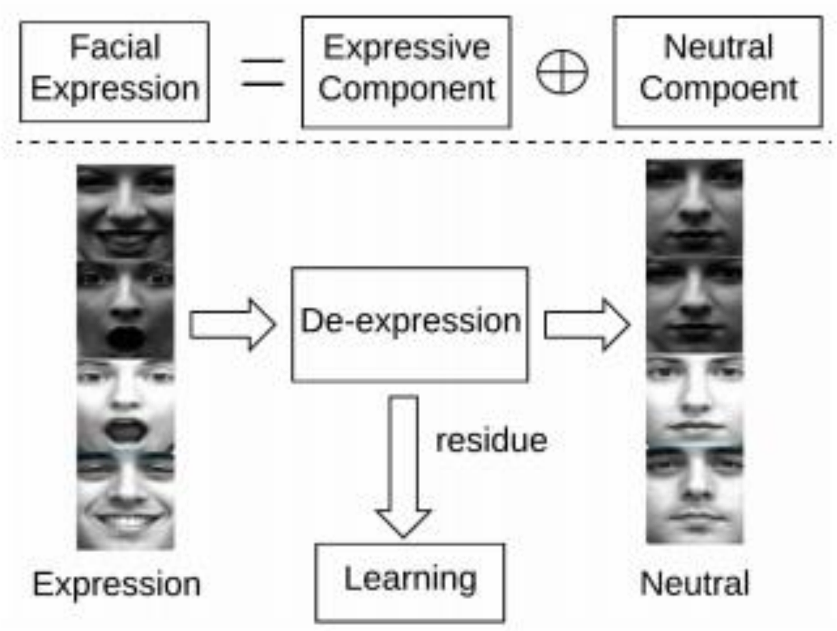
Strategy
- 학습된 임베딩 공간을 사용 → 시각적 & 의미적으로 유사한 이미지 그룹 식별 & AU classifier 훈련

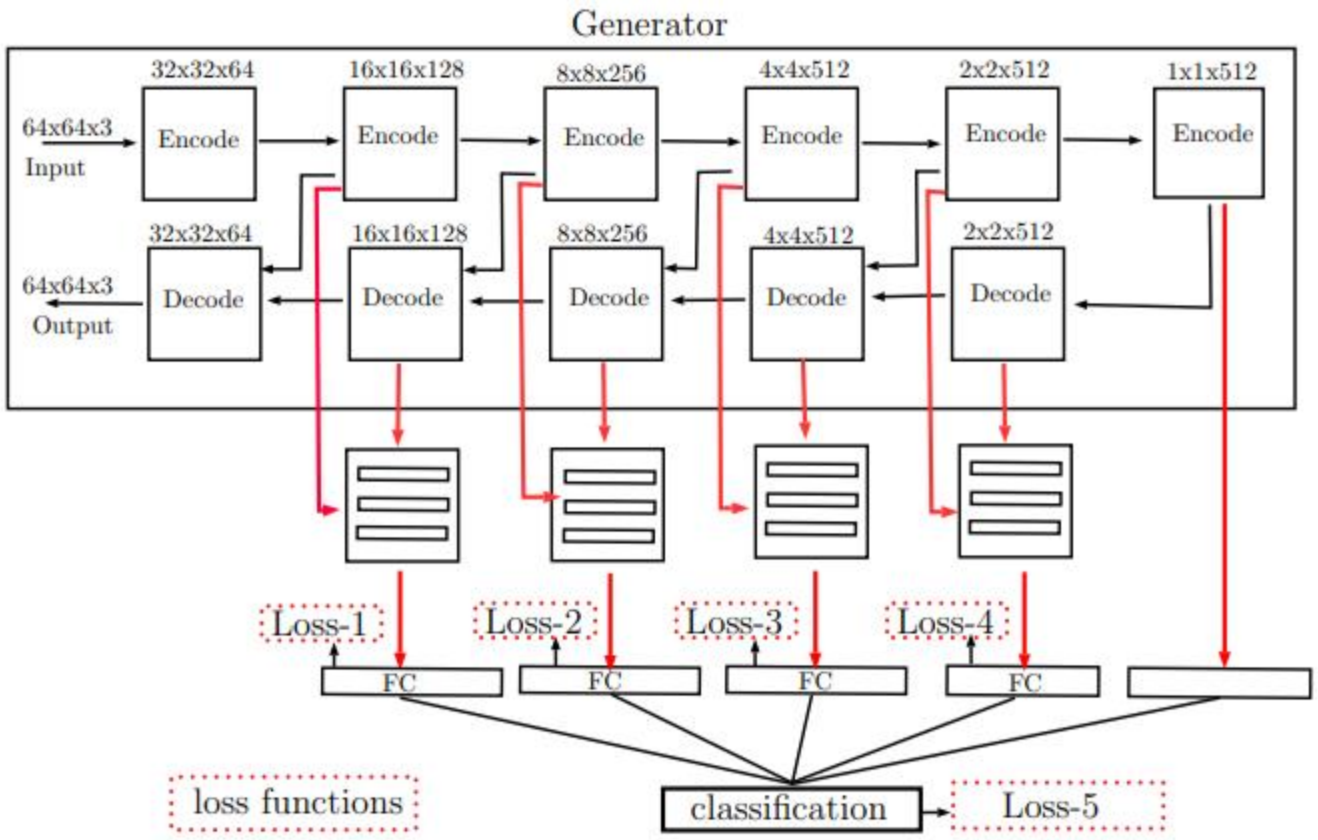
Background technique
: SVM

Dataset
: EmotioNet

[39] Facial Expression Recognition by De-Expression Residue Learning

Huiyuan Yang, Umur Ciftci, Lijun Yin





[39] Facial Expression Recognition by De-Expression Residue Learning

Huiyuan Yang, Umur Ciftci, Lijun Yin

Problem
: Facial Expression Recognition

Strategy
- DeRL(De-expression Residue Learning)을 통해 구성 요소의 정보를 추출하여 facial expression 인식
- Cgan을 통해 input image를 neutral image로 generate
- Expression 정보 → 중간 layer에 저장하고 이 부분을 학습하여 expression을 분류

Background technique
: GAN

Dataset
: BU-4DFE, BP4Dspontaneous

[40] Weakly Supervised Facial Action Unit Recognition Through Adversarial Training

Guozhu Peng, Shangfei Wang

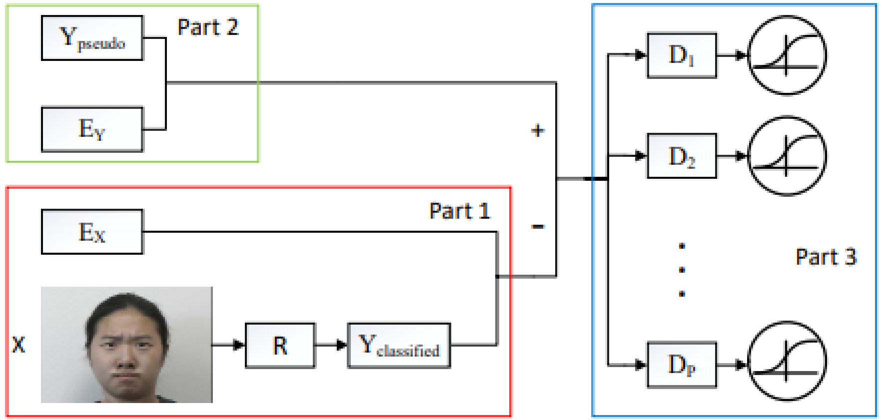


Figure 1: The framework of RAN. In Part 1, the facial feature X is inputted into recognizer R and get the "fake" AU vector, the "real" AU data generated in section 2.2 are in Part 2. In part 3, P discriminators are trained, "real" or "fake" AU data are inputted to corresponding discriminator with the same expression label. See text for details.

[40] Weakly Supervised Facial Action Unit Recognition Through Adversarial Training

Huiyuan Yang, Umur Ciftci, Lijun Yin

Problem
: Facial Action Unit Recognition

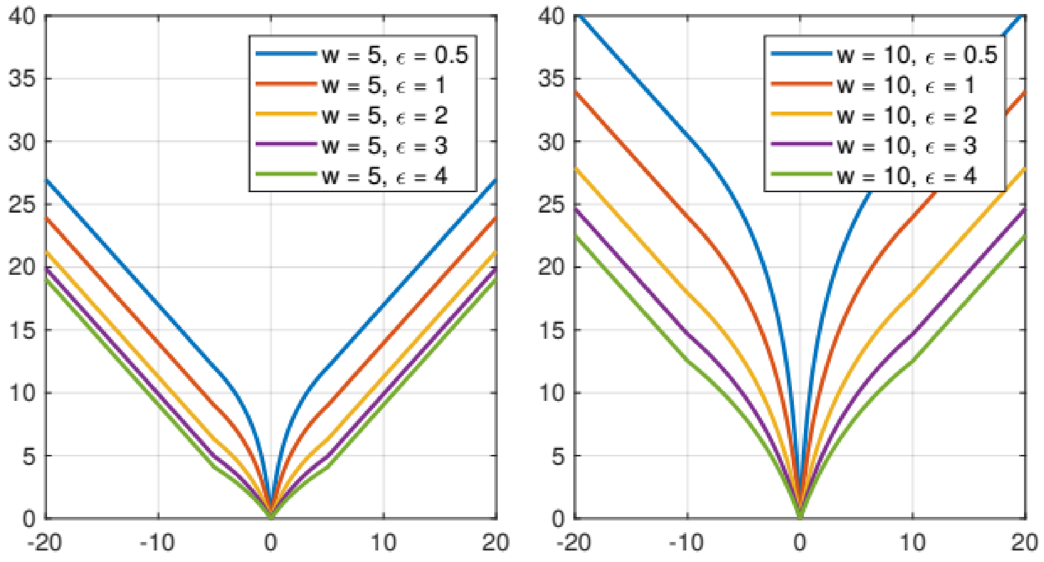
Strategy
- AU recognition method(Recognition model R, AU classifier, Discrimination model D)
- AU classifier는 R을 통해 학습하고 D를 통해 추정한다.

Background technique
: GAN

Dataset
: cohn-kande

[41] Wing Loss for Robust Facial Landmark Localisation With Convolutional Neural Networks

Zhen-Hua Feng, Josef Kittler, Muhammad Awais, Patrik Huber, Xiao-Jun Wu



(a) $w = 5$ (b) $w = 10$

Figure 1. Our Wing loss function (Eq. 5) plotted with different parameter settings, where w limits the range of the non-linear part and ϵ controls the curvature. By design, we amplify the impact of the samples with small and medium range errors to the network training.

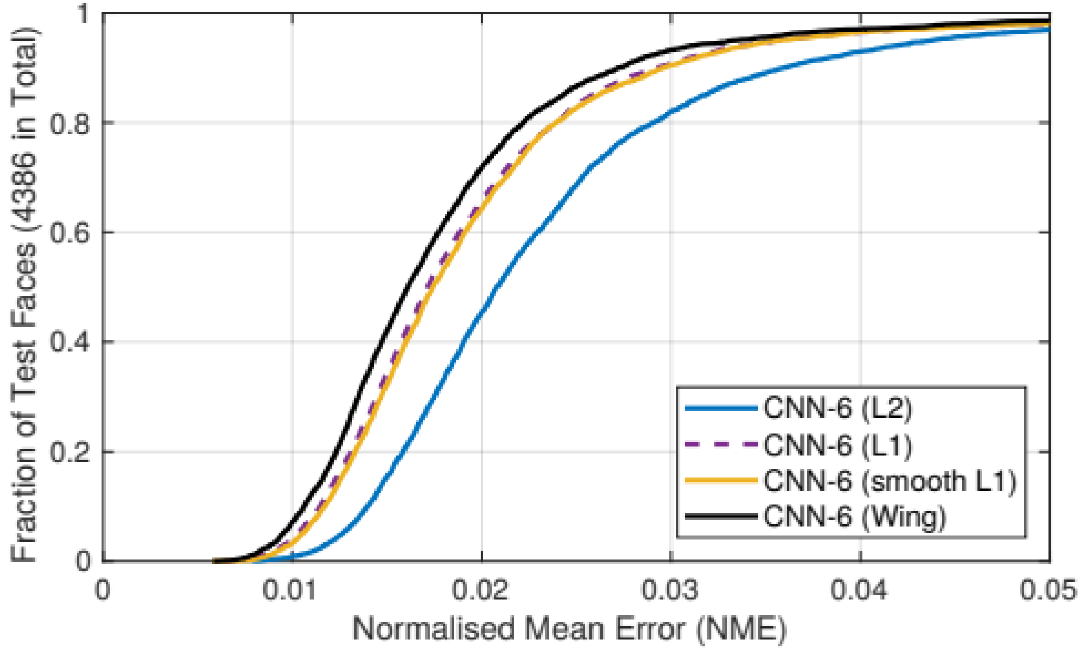


Figure 4. CED curves comparing different loss functions on the AFLW dataset, using the AFLW-Full protocol.

$$wing(x) = \begin{cases} w \ln(1 + |x|/\epsilon) & \text{if } |x| < w \\ |x| - C & \text{otherwise} \end{cases},$$

201810957 염지현

[41] Wing Loss for Robust Facial Landmark Localisation With Convolutional Neural Networks

Zhen-Hua Feng, Josef Kittler, Muhammad Awais, Patrik Huber, Xiao-Jun Wu

Problem

: Wing Loss for Robust Facial Landmark Localisation

Strategy

- Wing Loss's key idea
  - Increase contribution of the samples with small and medium size errors to the training of the regression network

Background technique

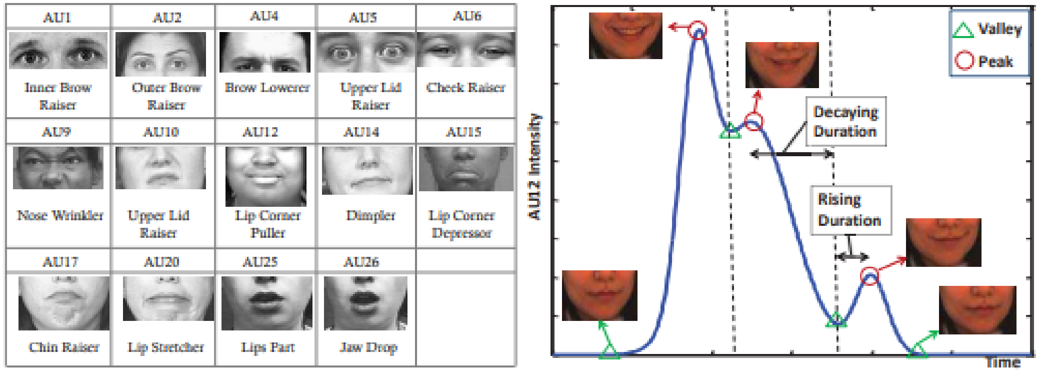
: CNN

Dataset

: AFLW, 300W

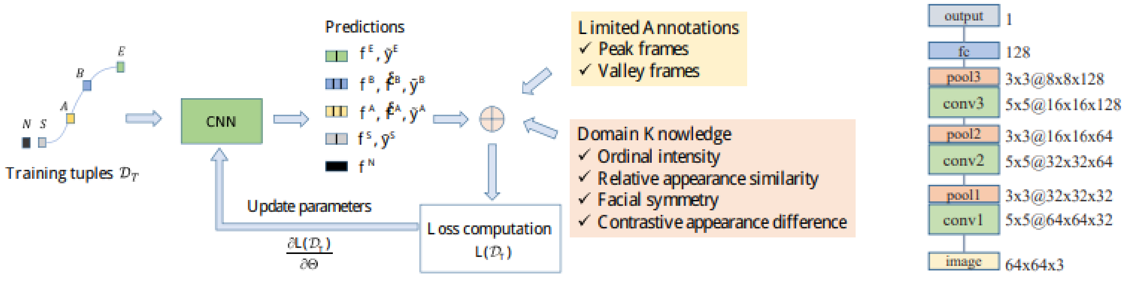
[42] Weakly-Supervised Deep Convolutional Neural Network Learning for Facial Action Unit Intensity Estimation

Yong Zhang, Weiming Dong, Bao-Gang Hu, Qiang Ji



(a) (b)

Figure 1. (a) Facial appearance of facial action units. (b) The intensity curve of AU12 in a sequence from [20].



(a) (b)

Figure 2. (a) The pipeline of the proposed method. Each frame in a tuple and also flipped face images of A and B go through the same CNN. The predictions are collected, including the representation from the fully collected layer and the predicted AU intensity. The loss is computed with using the predictions and the supervisory information, i.e., limited annotations and domain knowledge. The gradient of the loss is used to update the parameters of CNN. (b) The structure of CNN

201810957 염지현

[42] Weakly-Supervised Deep Convolutional Neural Network Learning for Facial Action Unit Intensity Estimation

Yong Zhang, Weiming Dong, Bao-Gang Hu, Qiang Ji

Problem

: Facial Action Unit Intensity Estimation

Strategy

- Knowledge-base-semi-supervised Deep CNN
    - Loss : Limited annotation & Domain knowledge를 통해 구함

Background technique

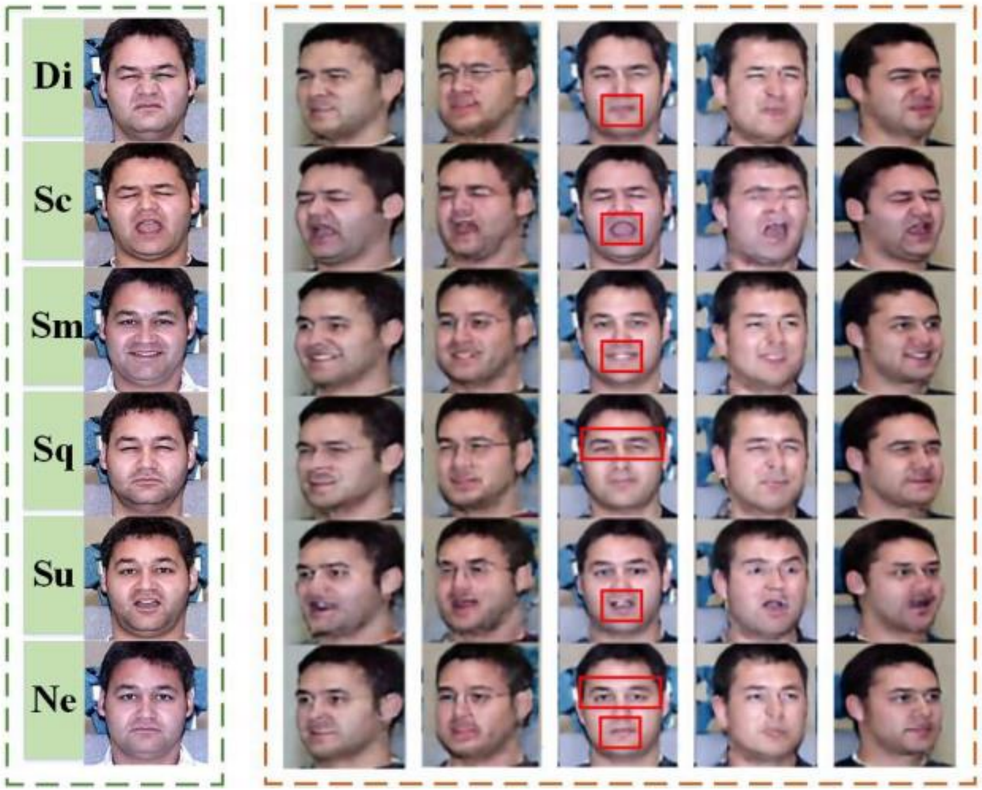
: CNN

Dataset

: FERA2015, DISFA

[43] Joint Pose and Expression Modeling for Facial Expression Recognition

Feifei Zhang, Tianzhu Zhang, Qirong Mao, Changsheng Xu



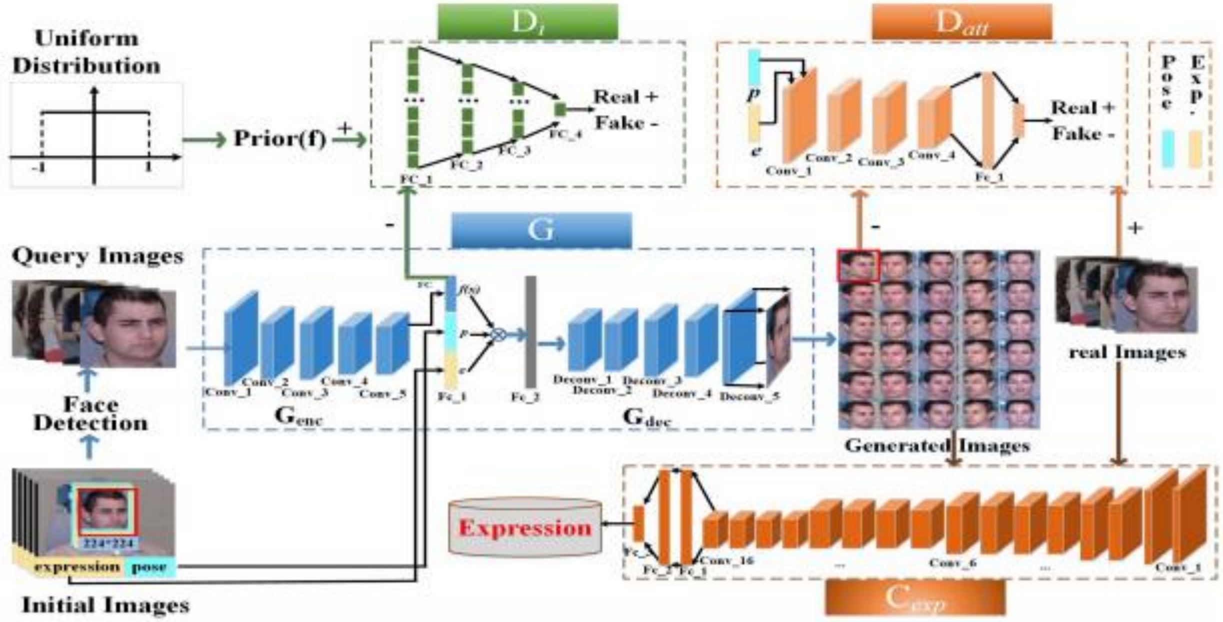


Figure 2. The overall architecture of the proposed model, which incorporates a generator G, two discriminators D_{att} and D_i, and a classifier C_{exp}. Conditioned by the expression and pose codes e and p, the proposed model can generate facial images with different expressions under arbitrary poses to enlarge and enrich the training set for the FER task.

201810957 염지현

[43] Joint Pose and Expression Modeling for Facial Expression Recognition

Feifei Zhang, Tianzhu Zhang, Qirong Mao, Changsheng Xu

Problem
: Facial Expression Recognition

Strategy
- 서로 다른 pose와 expression을 공동으로 활용하는 end-to-end Deep learning model

Background technique
: GAN

Dataset
: SFEW, Multi-PIE, BU-3DFE

[44] Modeling Facial Geometry Using Compositional VAEs

Timur Bagautdinov, Chenglei Wu, Jason Saragih, Pascal Fua, Yaser Sheikh

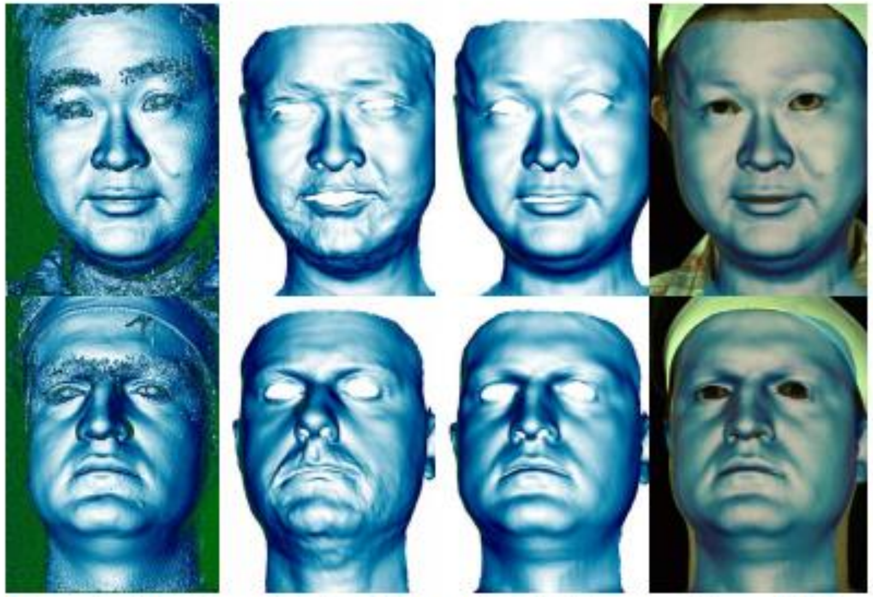
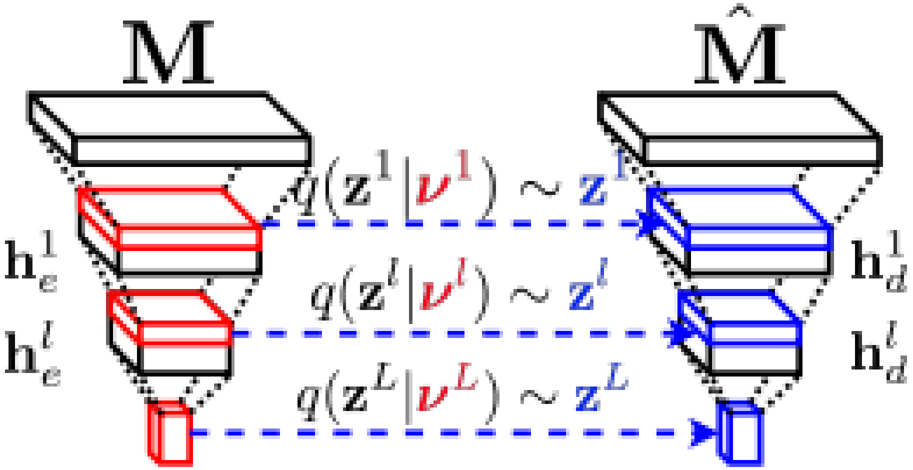


Figure 5. Visual results for fitting noisy depth maps. From left-to-right: input depth map, rendered mesh (LINEAR), rendered mesh (OURS), rendered mesh (OURS) overlaid with the image.



(d) Compositional VAE

[44] Modeling Facial Geometry Using Compositional VAEs

Timur Bagautdinov, Chenglei Wu, Jason Saragih, Pascal Fua, Yaser Sheikh

Problem

: Modeling Facial Geometry

Strategy

- Lower layers - capture global geometry
- Higher layers – 더 많은 local deformations(변형) 인코딩
  → 인간의 얼굴 구조를 의미있는 세부 수준으로 자연스럽게 분해 가능

Background technique

: CNN

Dataset

: 300-W

[45] Optimizing Filter Size in Convolutional Neural Networks for Facial Action Unit Recognition

Shizhong Han, Zibo Meng, Zhiyuan Li, James O'Reilly, Jie Cai, Xiaofeng Wang, Yan Tong

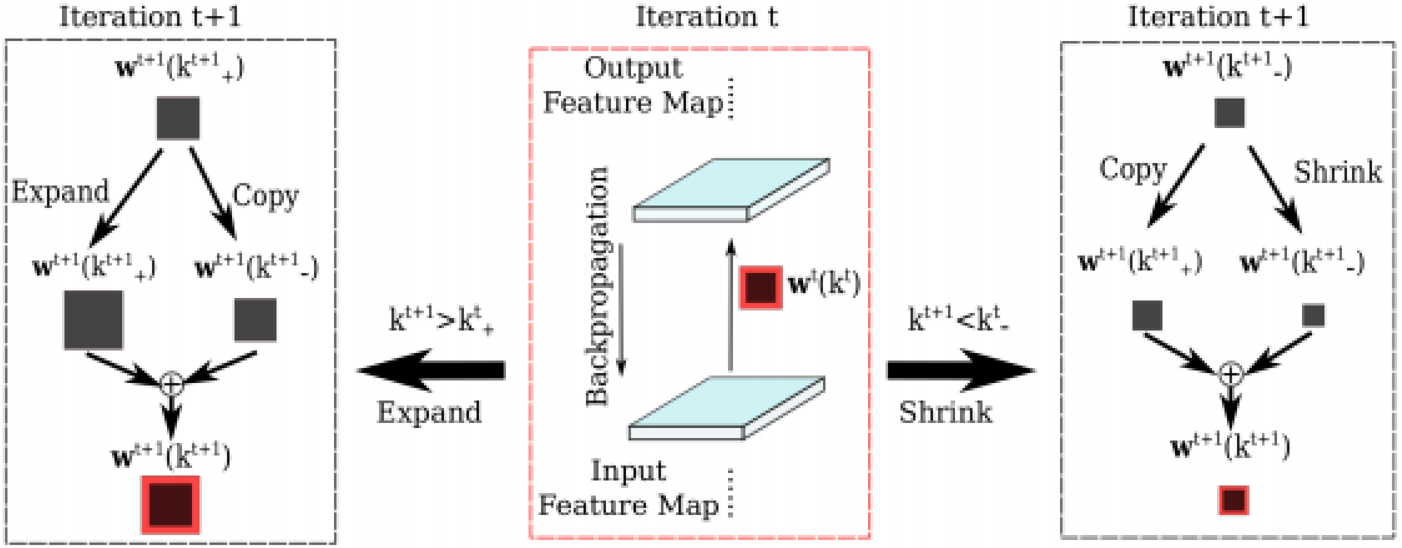


Figure 3. When the filter size k is updated during backpropagation, it may be out of the interval $[k^t_-, k^t_+)$. In this case, transformation operations are needed to update the sizes of the upper-bound and lower-bound filters after updating their coefficients. Specifically, an expanding operation is employed to increase the sizes of both upper-bound and lower-bound filters; whereas a shrinking operation is used to decrease the filter sizes.

[45] Optimizing Filter Size in Convolutional Neural Networks for Facial Action Unit Recognition

Shizhong Han, Zibo Meng, Zhiyuan Li, James O'Reilly, Jie Cai, Xiaofeng Wang, Yan Tong

Problem
: Facial Action Unit Recognition

Strategy
- OFS-CNN(Optimal Filter Size –CNN) → Filter size도 함께 학습
  → optimal한 filter size 찾기

Background technique
: CNN

Dataset
: DISFA

[46] Classifier Learning With Prior Probabilities for Facial Action Unit Recognition

Yong Zhang, Weiming Dong, Bao-Gang Hu, Qiang Ji

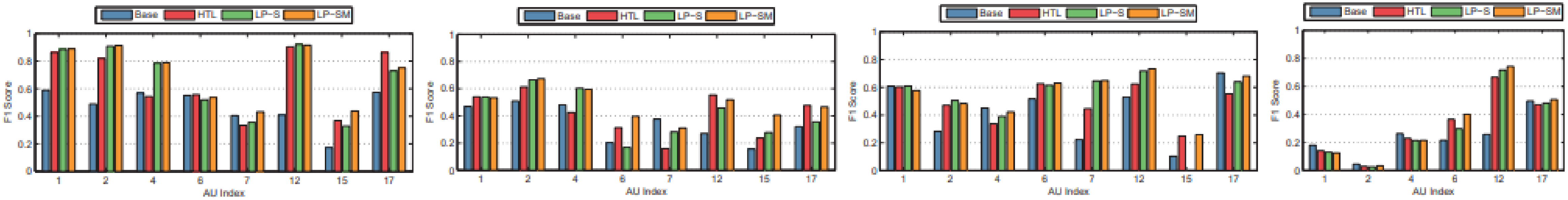


Figure 1. Within database performance for each AU. From left to right are the results on CK+, MMI, BP4D, and ENet.

201810957 염지현

[46] Classifier Learning With Prior Probabilities for Facial Action Unit Recognition

Yong Zhang, Weiming Dong, Bao-Gang Hu, Qiang Ji

Problem

: Facial Unit Recognition without AU annotation

Strategy

- AU 사전확률(expression-independent & expression dependent 포함) 활용하여 AU annotation 없이 AU classifier 학습

Background technique

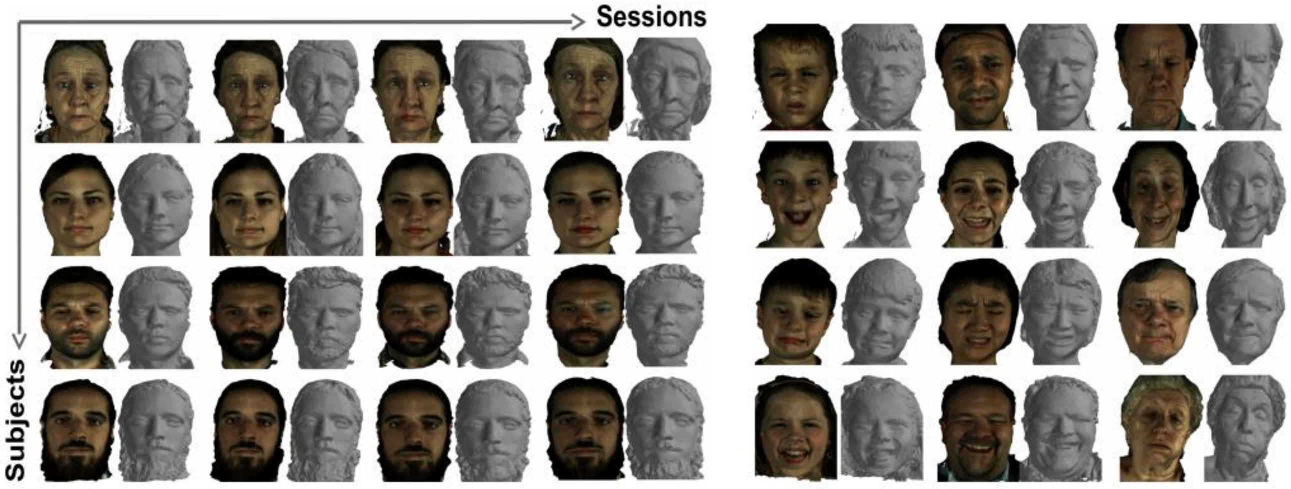
: CNN

Dataset

: CK+, MIMI, BP4D

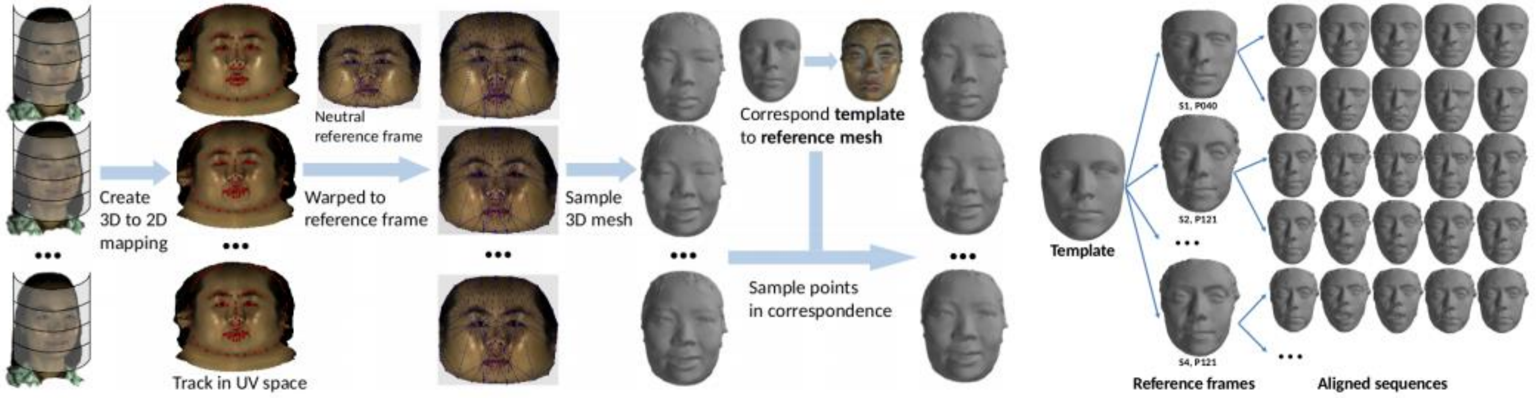
[47] 4DFAB: A Large Scale 4D Database for Facial Expression Analysis and Biometric Applications

Shiyang Cheng, Irene Kotsia, Maja Pantic, Stefanos Zafeiriou



(a) Examples of subjects in four sessions. (b) Examples of spontaneous expressions.

Figure 2: Examples of 3D faces in the database.



(a) 4D alignment pipeline. (b) Overview of correspondences.

Figure 3: Our framework for establishing 4D dense correspondences.

201810957 염지현

[47] 4DFAB: A Large Scale 4D Database for Facial Expression Analysis and Biometric Applications

Shiyang Cheng, Irene Kotsia, Maja Pantic, Stefanos Zafeiriou

Problem
: Computer Vision 분야의 data 부족

Strategy
- Dynamic high-resolution 3D face의 database 생성

Background technique
: Dataset

Dataset
: B3D, BP4D-Spontaeous, BP4D+

[48] Bilateral Ordinal Relevance Multi-Instance Regression for Facial Action Unit Intensity Estimation

Yong Zhang, Rui Zhao, Weiming Dong, Bao-Gang Hu, Qiang Ji

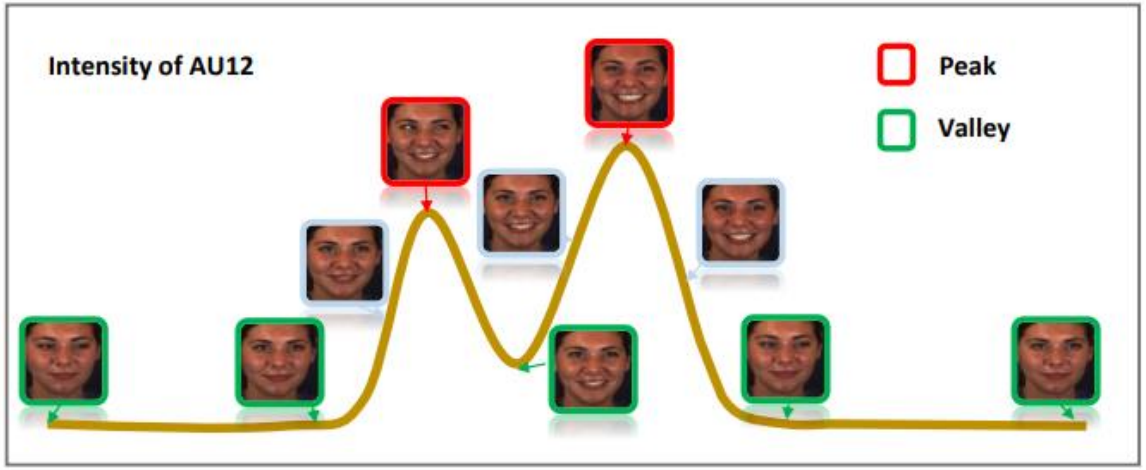


Figure 2. Illustration of AU in a sequence.

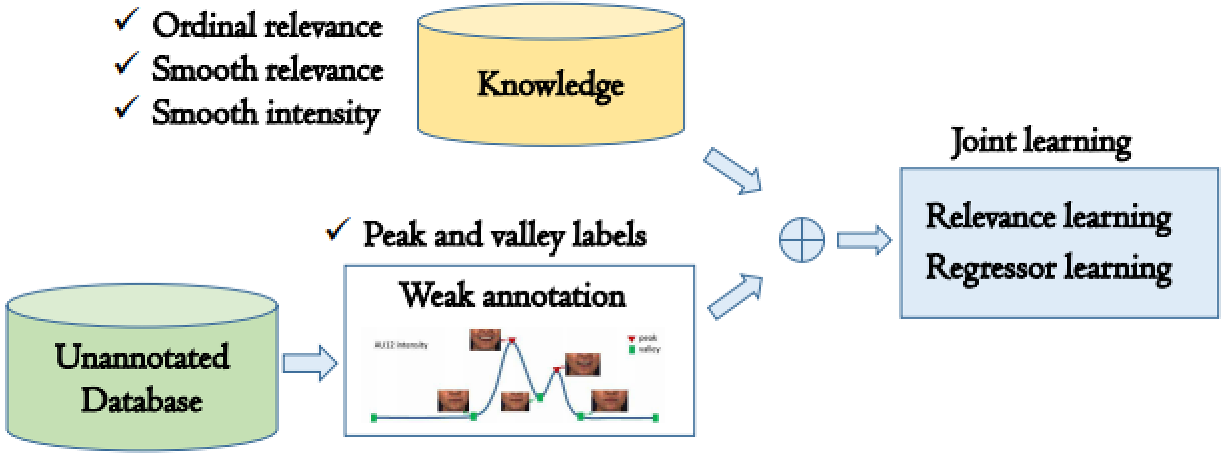


Figure 1. The pipeline of the proposed method. BORMIR combines the weakly annotated sequences and the domain knowledge to jointly learn the frame-level relevance and intensity regressor.

201810957 염지현

[48] Bilateral Ordinal Relevance Multi-Instance Regression for Facial Action Unit Intensity Estimation

Yong Zhang, Rui Zhao, Weiming Dong, Bao-Gang Hu, Qiang Ji

Problem

: Facial Action Unit Intensity Estimation
(subtle changes of facial appearance & the lack of AU's annotation)

Strategy

- BORMIR(model-Bilateral Ordinal Relevance Multi-instance Regression)
  → weakly labeled sequence를 사용하여 intensity estimation

Background technique

: Deep models

Dataset

: FERA2015, DISFA, PAIN

[49] Mesoscopic Facial Geometry Inference Using Deep Neural Networks

Loc Huynh, Weikai Chen, Shunsuke Saito, Jun Xing, Koki Nagano, Andrew Jones, Paul Debevec, Hao Li;

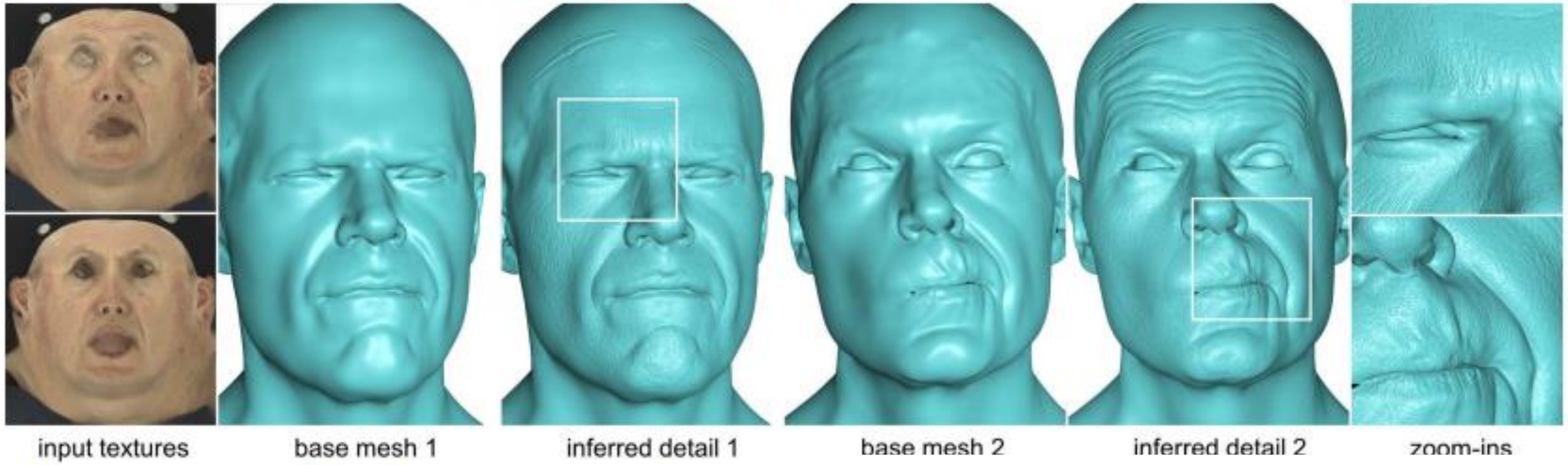


Figure 1: Given a flat-lit facial input textures and a base mesh,

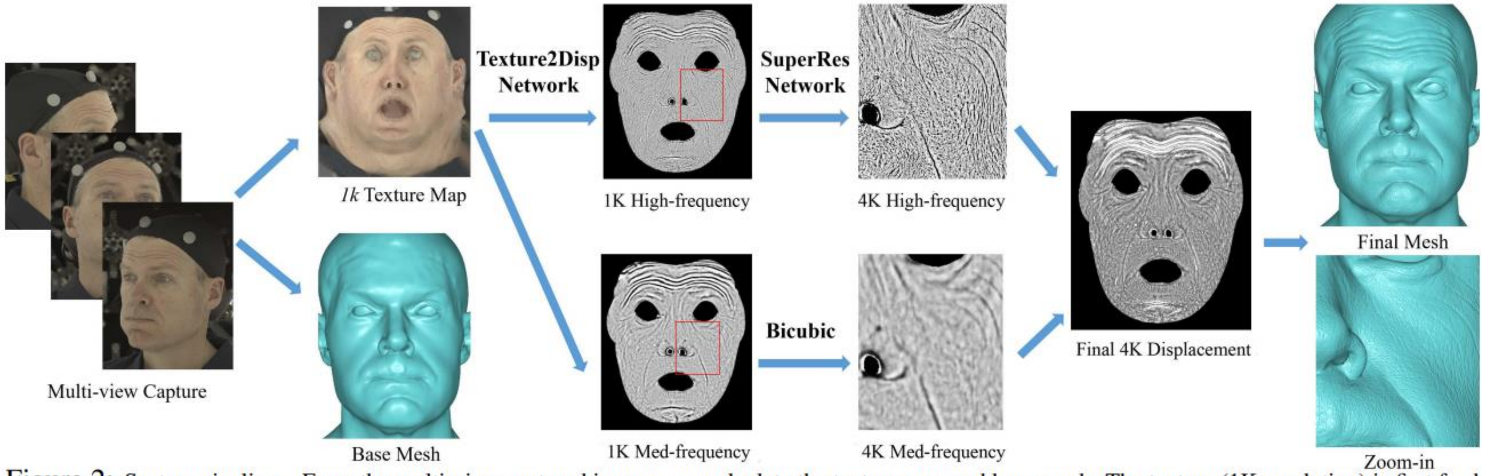


Figure 2: System pipeline. From the multi-view captured images, we calculate the texture map and base mesh. The texture (1K resolution) is first feed into our trained Texture2Disp network to produce a 1K-high and 1K-middle frequency displacement maps, followed by up-sampling them to 4K resolution using our trained SuperRes Network and bicubic interpolation, respectively. The combined 4K displacement map can be embossed to the base mesh to produce the final high detailed mesh.

201810957 염지현

[49] Mesoscopic Facial Geometry Inference Using Deep Neural Networks

Loc Huynh, Weikai Chen, Shunsuke Saito, Jun Xing, Koki Nagano, Andrew Jones, Paul Debevec, Hao Li;

Problem
: Facial Geometry Inference

Strategy
- Light stage에서 polarized gradient illumination 사용
  → 수염, 모공과 같은 어두운 부분을 오목한 부분으로 해석하는 것을 방지
- Hybrid Network(image-to-image translation, super resolution network)
  → 미세한 detail encoding

Background technique
: CNN, GAN

Dataset
: 언급(X)

Face & Facial

NIPS

201810957 염지현

[50] Learning a High Fidelity Pose Invariant Model for High-resolution Face Frontalization

Jie Cao, Yibo Hu, Hongwen Zhang, Ran He, Zhenan Sun

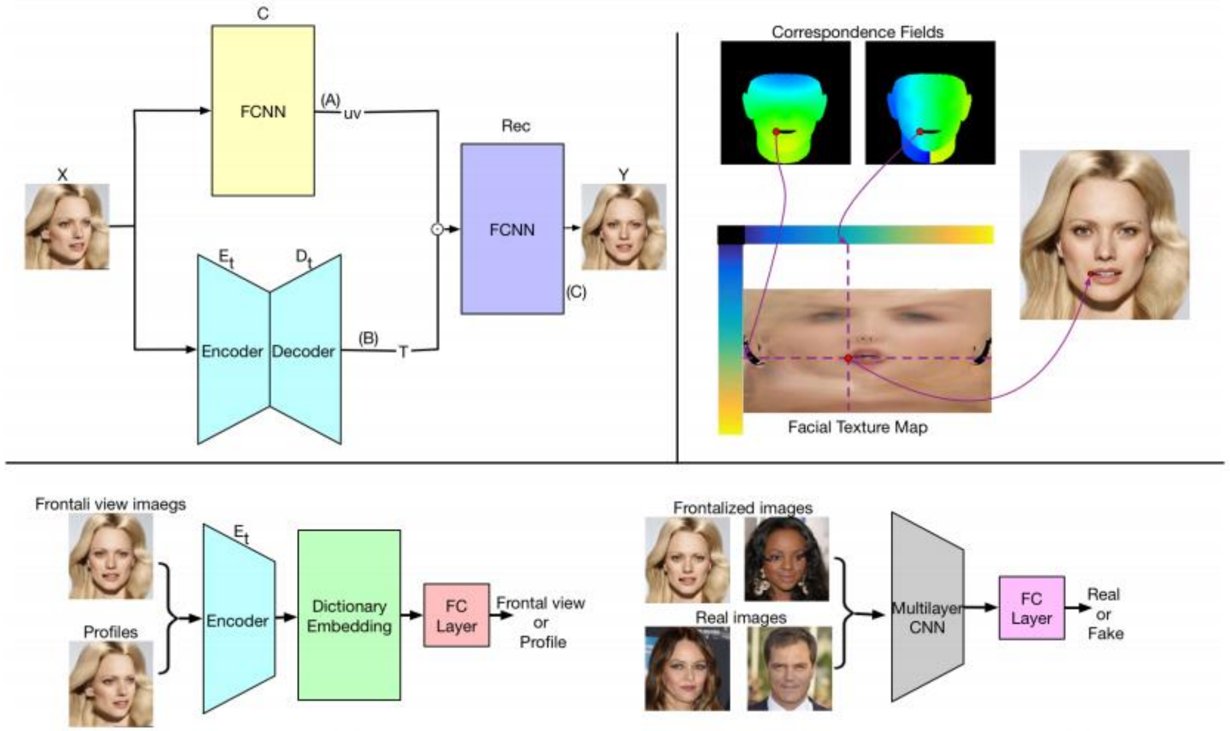


Figure 1: Left side on top: the framework of our HF-PIM to frontalize face images. The procedure consists of correspondence field estimation (A), facial texture feature map recovering (B) and frontal view warping (C). The right side on top: an illustration about the warping procedure discussed in Eq. 1. Those red dots and purple lines indicate the relationships between the facial texture map, correspondence field, and the RGB color image. Bottom side: the discriminators employed for ARDL (on the left) and ordinary adversarial learning (on the right).

201810957 염지현

[50] Learning a High Fidelity Pose Invariant Model for High-resolution Face Frontalization

Jie Cao, Yibo Hu, Hongwen Zhang, Ran He, Zhenan Sun

Problem
: High-resolution Face Frontalization

Strategy
- HF-PIM(High Fidelity Pose Invariant Model)
- Identity-preserving
  - correspondence fields → facial texture map, correspondence field, RGB color image
  - ARDL – 3D data 없이 facial texture map을 revering
    (Adversarial Residual Dictionary Learning)

Background technique
: CNN, GAN

Dataset
: IJB-A, LFW, Multi-PIE, CelebA-HQ

201810957 염지현

[51] Unsupervised Depth Estimation, 3D Face Rotation and Replacement

Joel Ruben Antony Moniz, Christopher Beckham, Simon Rajotte, Sina Honari, Chris Pal

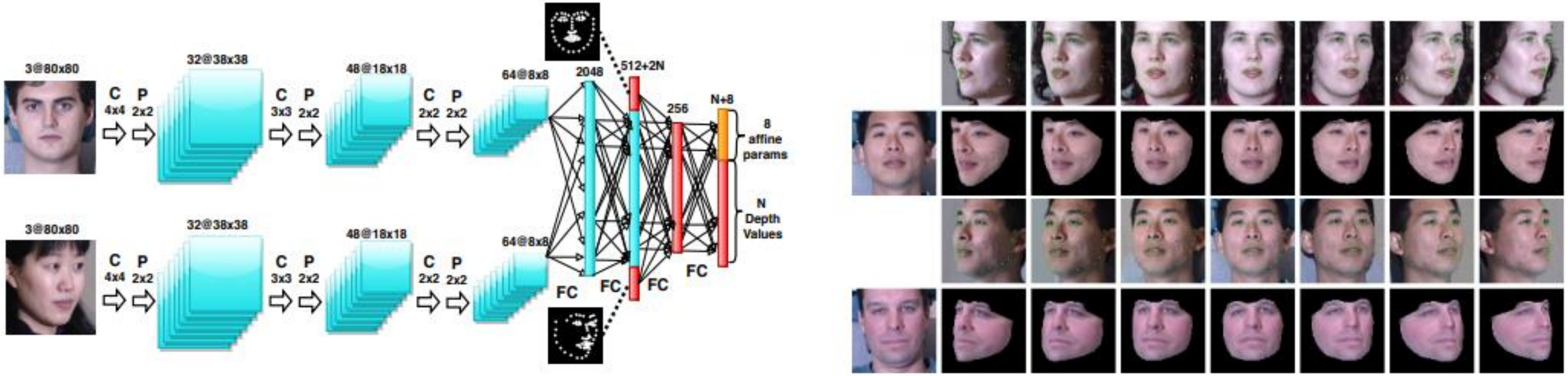


Figure 1: (Left) DepthNet architecture. The blue region is only used in case (A) and the red part is used in both cases (A) and (B), described in Section 2. The orange output (the 8 affine transformation parameters) is predicted only by model variations described in Sections 2.1 and 2.2, and not the model described in section 2.3. All three models predict the N depth values of the source keypoints. C, P, and FC correspond to valid conv, pool and fully-connected layers. The two paths of Siamese network share parameters and the black dots indicate concatenating keypoint values to FC units. (Right) Visualizing face rotation by re-projecting a frontal face (far left) to a range of other poses defined by the faces in the row above (in each pair of rows). In this experiment, we only use keypoints from the top-row in the DepthNet model (Model 7 in Table 1).

[51] Unsupervised Depth Estimation, 3D Face Rotation and Replacement

Joel Ruben Antony Moniz, Christopher Beckham, Simon Rajotte, Sina Honari, Chris Pal

Problem
: 3D Face Rotation & Replacement(Unsupervised)

Strategy
- DepthNet
  - 입력된 face로부터 추론한 3D key points
  - 원하는 target facial geometry or pose
    → mapping parameters of a 3D affine transform matrix 예측
  - Depth * Geometric transformation

Background technique
: CNN, GAN

Dataset
: 3DFAW, 300W, Multi-PIE, CelebA

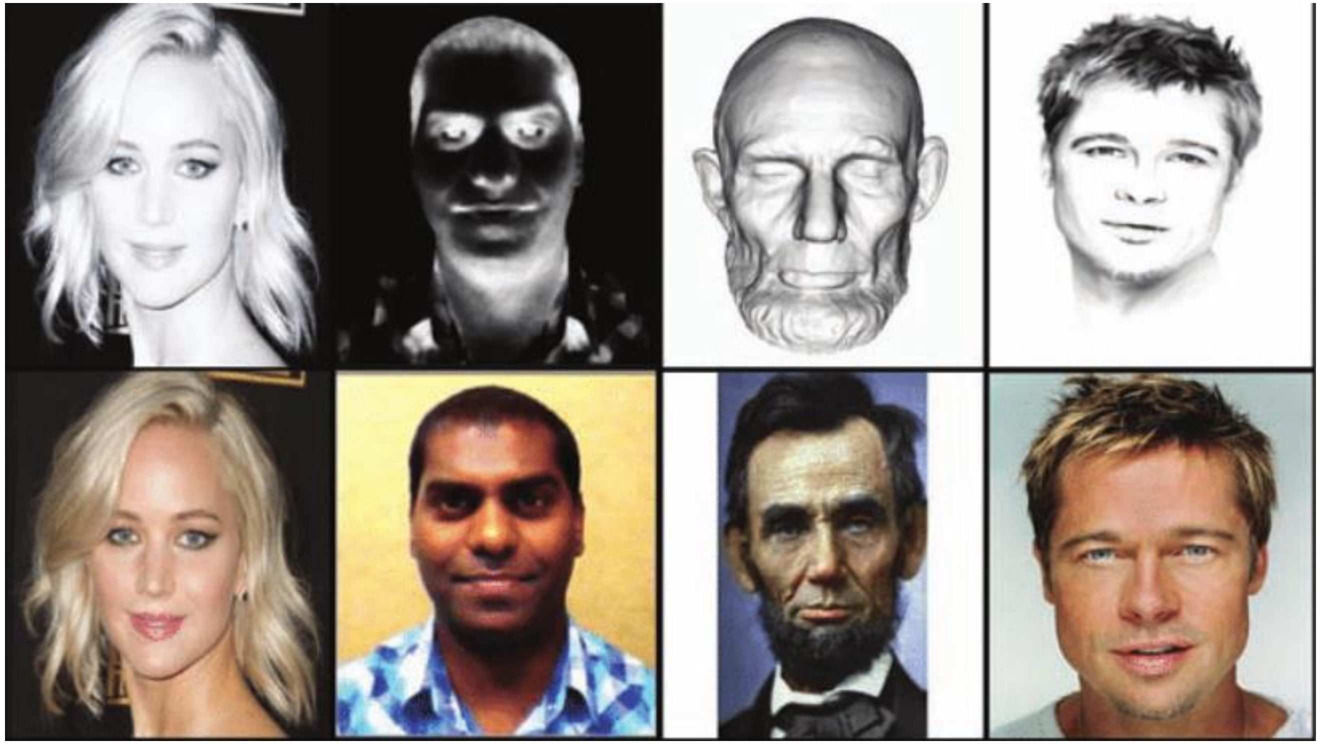
Face & Facial

AAAI

201810957 염지현

[52] Coupled Deep Learning for Heterogeneous Face Recognition

Xiang Wu, Lingxiao Song, Ran He, Tieniu Tan



[52] Coupled Deep Learning for Heterogeneous Face Recognition

Xiang Wu, Lingxiao Song, Ran He, Tieniu Tan

Problem

: Heterogeneous Face Recognition → paired image 불충분

Strategy

- CDL(Coupled Deep Learning) --> objective function
  - 짝을 이루지 않은 이미지 clustering & correlating
  - Regularize parameter → overfitting 완화
  - 다른 Identity와의 마진 최대화 및 sampled data 늘리기 위해 cross modal ranking 이 포함

Background technique

: CNN, GAN

Dataset

: CASIA NIR-VIS 2.0, IIIT-D, CUHK, CUFS, CUFSF

[53] r-BTN: Cross-Domain Face Composite and Synthesis From Limited Facial Patches

Yang Song, Zhifei Zhang, Hairong Qi

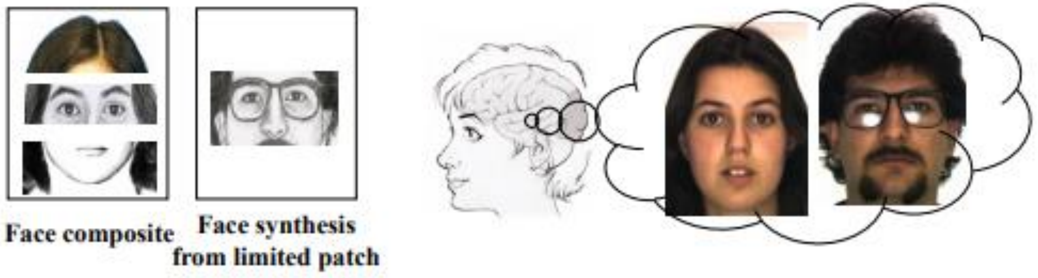


Figure 1: Illustration of face composite based on cross-domain patches and face synthesis from limited facial patch.

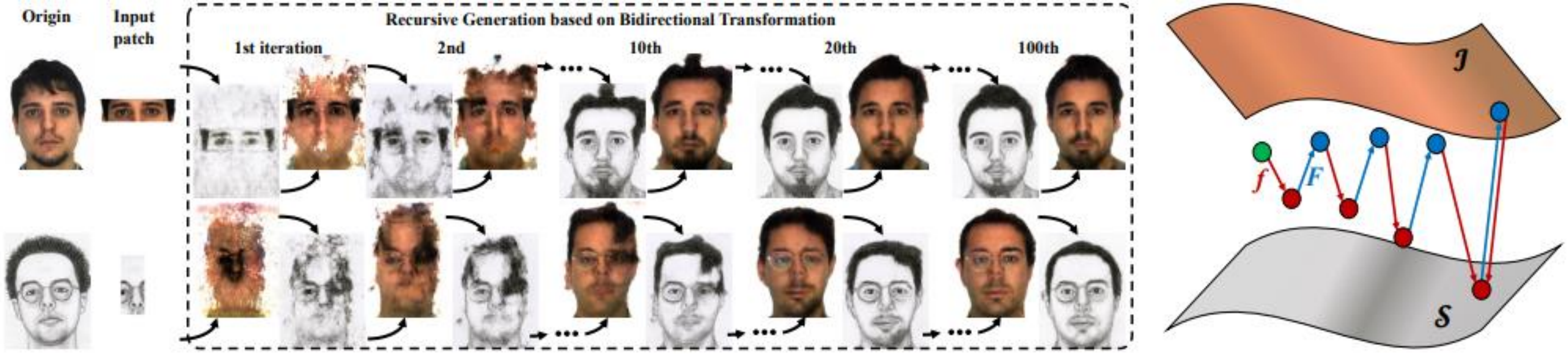


Figure 2: Examples of recursive generation from small patches by the bidirectional transformation network. Left: Original face/sketch and the corresponding input patches extracted from them. Inside of the dashed box demonstrates the generated face/sketch at different iteration steps. Right: Illustration of transformation between the face and sketch manifolds $\mathcal{I}$ and $\mathcal{S}$, respectively. The green dot denotes a given face patch. The red and blue arrows are the learned mapping f and F, respectively. The red and blue dots are generated sketches and faces through corresponding mapping.

201810957 염지현

[53] r-BTN: Cross-Domain Face Composite and Synthesis From Limited Facial Patches

Yang Song, Zhifei Zhang, Hairong Qi

Problem

: Face Composite and Synthesis from Limited Facial Patches

Strategy

- r-BTN(recursive generation by Bidirectional Transformation Network)
  → bidirectional transformation network를 사용하여 forward & backward error를 최소화하여
  일관된 결과 도출

Background technique

: CNN, GAN

Dataset

: CHUK, CUFSF, AR, FERET, IIIT-D, CFD, siblingsDB, PUT

[54] Asymmetric Joint Learning for Heterogeneous Face Recognition

Bing Cao, Nannan Wang, Xinbo Gao, Jie Li

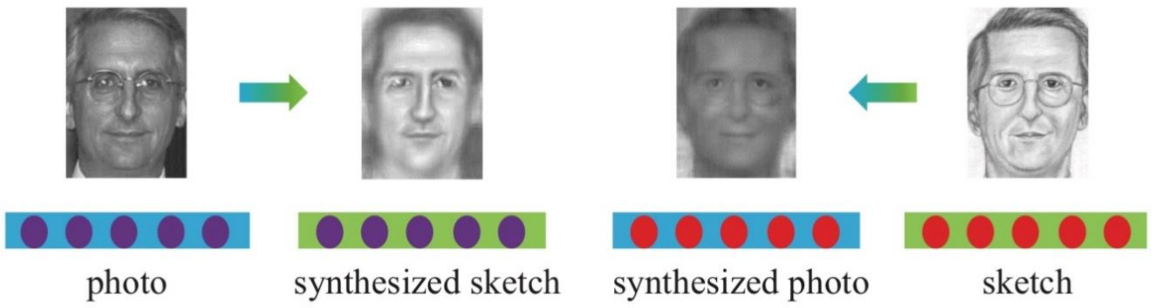


Figure 1: The latent information in synthesized images. Different colors of circles and backgrounds represent different texture information and domains respectively. For example, The purple circles with the blue rectangle background represent the texture information of one subject in photo domain and the purple circles with the green rectangle background represent the same texture information of the synthesized sketch in sketch domain.

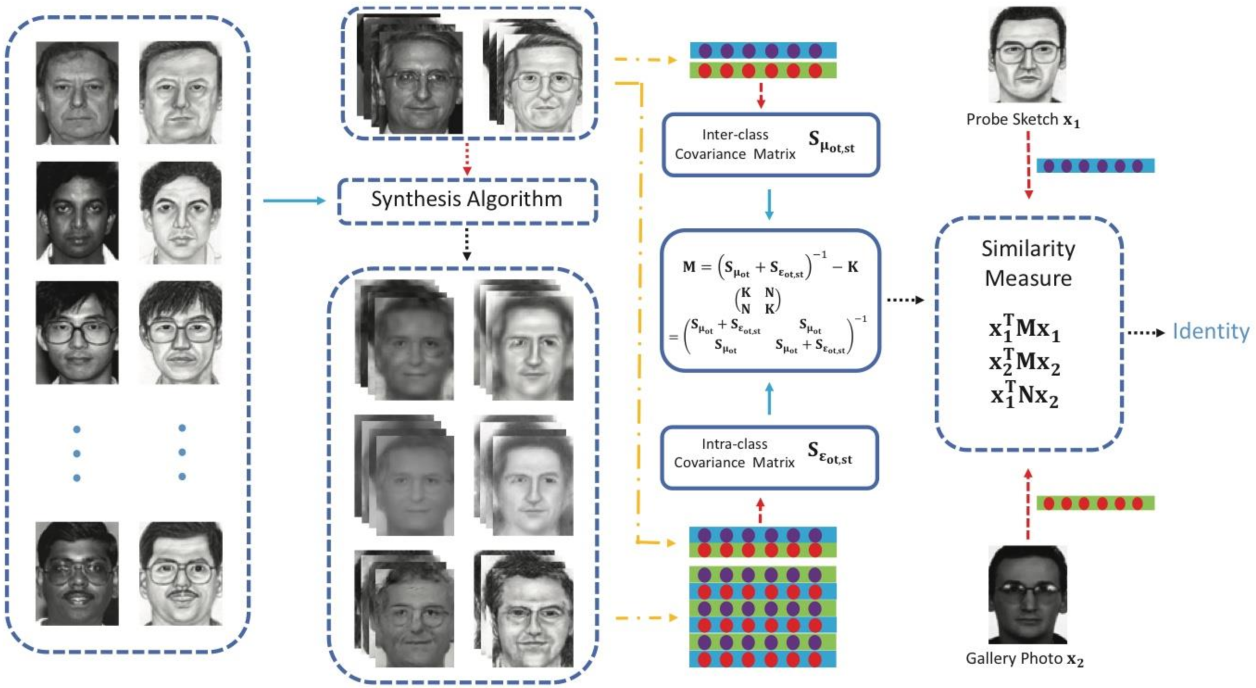


Figure 2: Framework of the proposed asymmetric joint learning method for heterogeneous face recognition.

201810957 염지현

[54] Asymmetric Joint Learning for Heterogeneous Face Recognition

Bing Cao, Nannan Wang, Xinbo Gao, Jie Li

Problem
: HFR(Heterogeneous Face Recognition) – cross-modality image 간의 불일치(paired data (x))

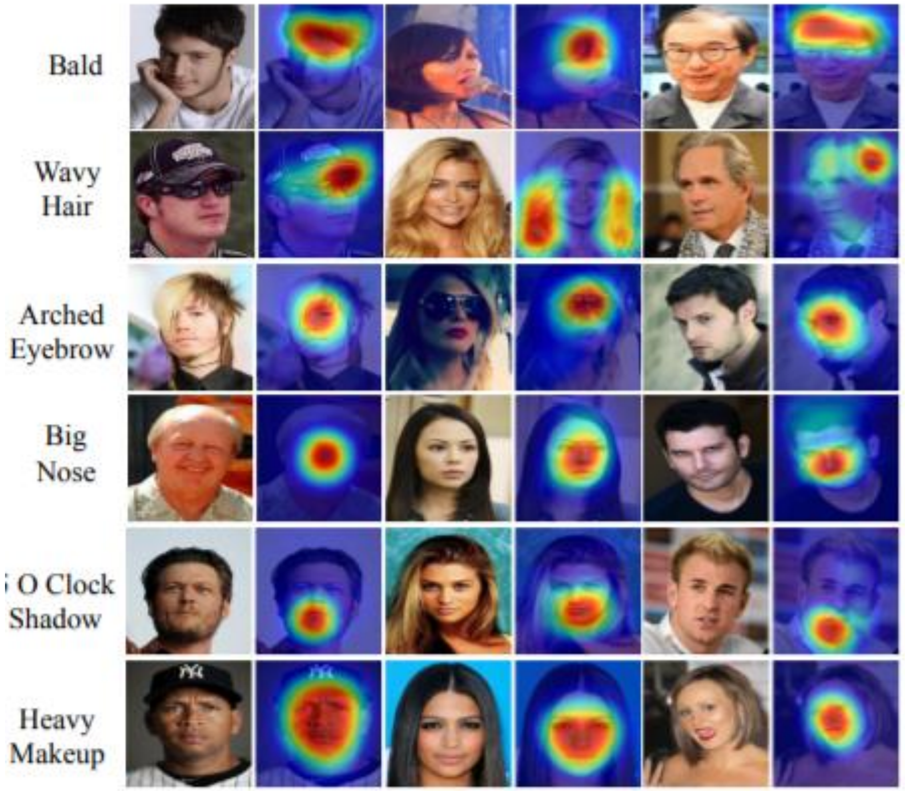
Strategy
- AJL(Asymmetric Joint Learning)
  - 합성된 이미지에 discriminative information을 learning process애 합쳐서 상호 양식 차이를
    상호 변화

Background technique
: GAN

Dataset
: CUFSF, IIIT-D, CUHK VIS-NIR

[55] A Deep Cascade Network for Unaligned Face Attribute Classification

Hui Ding, Hao Zhou, Shaohua Kevin Zhou, Rama Chellappa



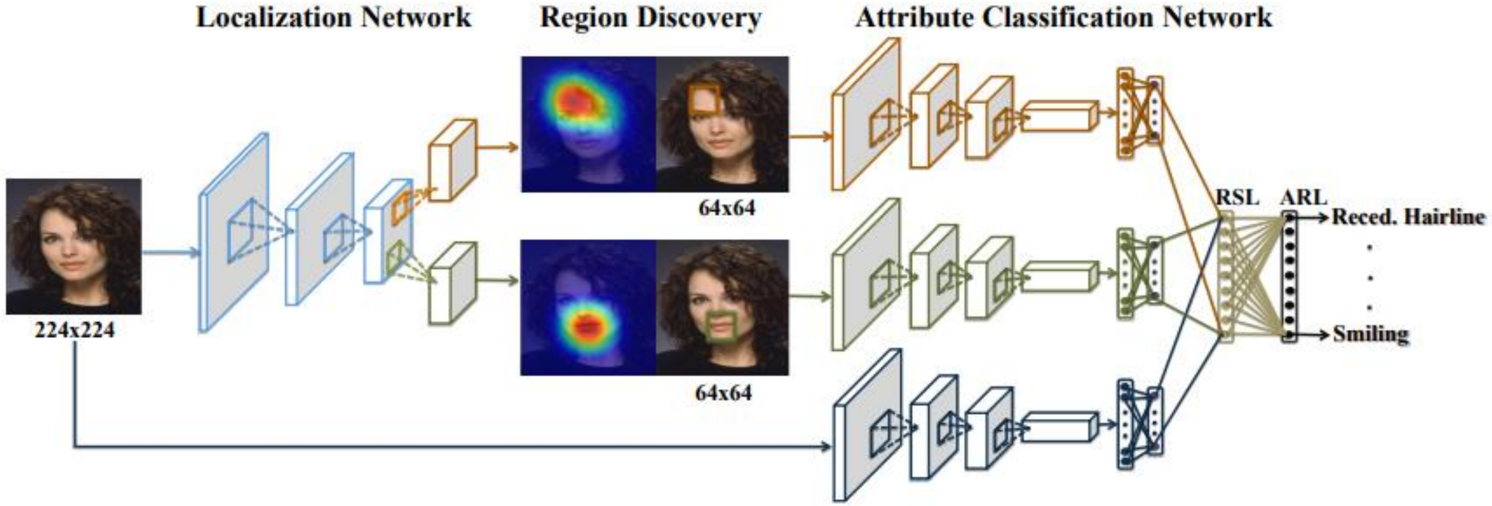


Figure 1: Overview of our face attribute recognition framework. It consists of a facial region localization (FRL) network and a Parts and Whole (PaW) classification network. The localization network detects a discriminative part for each attribute. Then the detected face regions and the whole face image are fed into the PaW classification network. The region switch layer (RSL) selects the relevant subnet for predicting the attribute, while the attribute relation layer (ARL) models the attribute relationships.

[55] A Deep Cascade Network for Unaligned Face Attribute Classification

Hui Ding, Hao Zhou, Shaohua Kevin Zhou, Rama Chellappa

Problem

: Face Attribute Classification
→ learns to localize face regions specific to attribute without alignment face attribute classification

Strategy

- Cascade Network
  - Weakly-supervised face region localization network – 자동으로 region detect
  - Multiple port-based networks & whole-image-based network → 개별적으로 구성
  - 최종 분류 단계 = 영역 스위치 계층 + 속성 관계 계층

Background technique

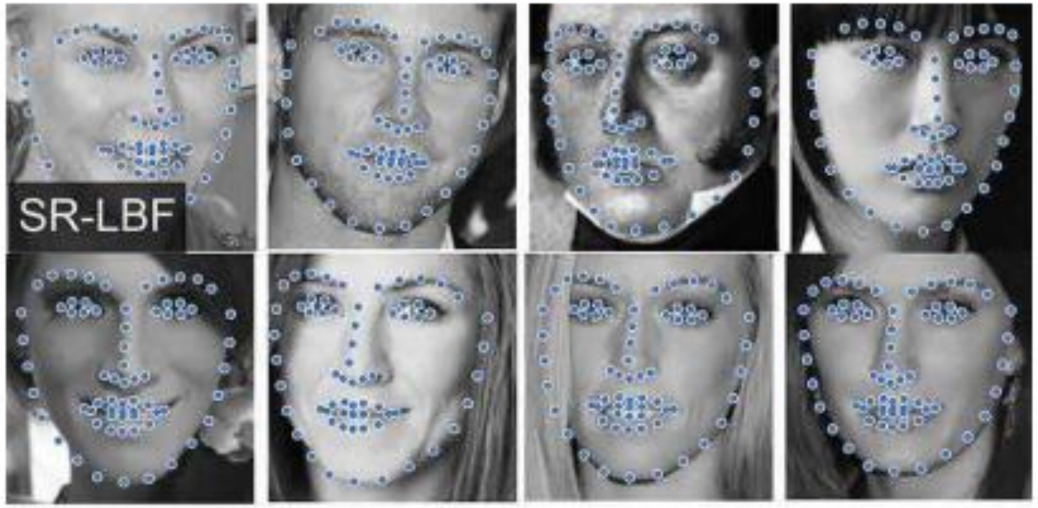
: CNN

Dataset

: CelebA

[56] Self-Reinforced Cascaded Regression for Face Alignment

Xin Fan, Risheng Liu, Kang Huyan, Yuyao Feng, Zhongxuan Luo



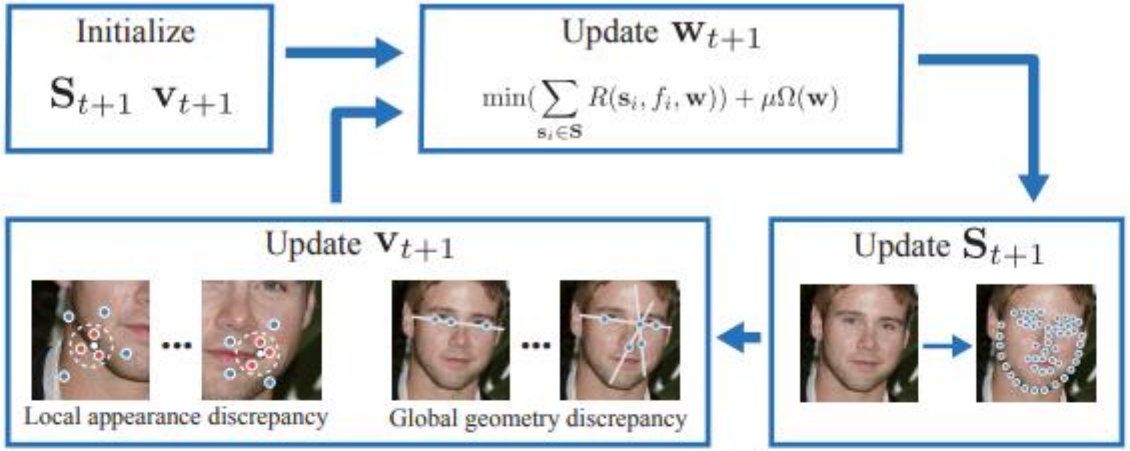


Figure 1: Overview of our self-reinforced cascaded regression, forming a closed loop with label prediction $\mathbf{S}$ and survival $\mathbf{v}$ as well as regression upgrading $\mathbf{W}$.

201810957 염지현

[56] Self-Reinforced Cascaded Regression for Face Alignment

Xin Fan, Risheng Liu, Kang Huyan, Yuyao Feng, Zhongxuan Luo

Problem

: Face alignment

: cascaded regression은 accuracy, robustness 덕분에 face alignment에서 우세하지만 shapeindexed features와 shape updates 사이의 불일치가 낮은 annotated examples 필요

Strategy

- Self-Reinforced Cascaded Regression
    - 반복적으로 training examples 수량 확장 및 질 향상을 통해 cascaded regression 자체의 성능 향상
    - Human face's local appearance and global geometry에 대한 일관성 → 품질 평가

Background technique

: GAN

Dataset

: FRGC v2.0, LFPW, HELEN, AFW, iBUG, 300W

[57] Unravelling Robustness of Deep Learning Based Face Recognition Against Adversarial Attacks

Gaurav Goswami, Nalini Ratha, Akshay Agarwal, Richa Singh, Mayank Vatsa

Figure 1: We show that deep learning based OpenFace (OF) and VGG-Face can be deceived even by image processing operations that mimic real world distortions.

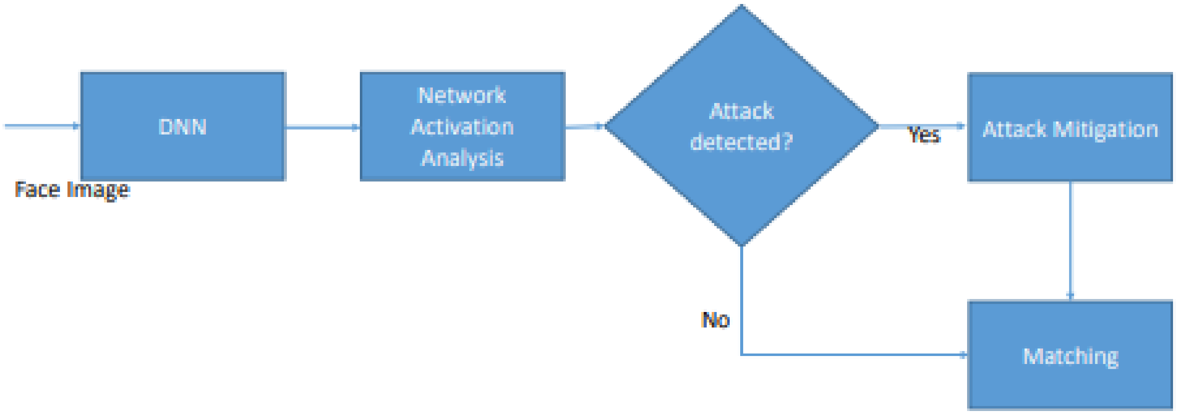


Figure 4: Flow chart for the proposed detection and mitigation methodology.

201810957 염지현

[57] Unravelling Robustness of Deep Learning Based Face Recognition Against Adversarial Attacks

Gaurav Goswami, Nalini Ratha, Akshay Agarwal, Richa Singh, Mayank Vatsa

Problem
: Face Recognition을 위해 학습시키는 DNN은 Black box method

Strategy
- Face recognition을 위한 DNN의 견고성

Background technique
: DNN

Dataset
: MEDS, PaSC

[58] Merge or Not? Learning to Group Faces via Imitation Learning

ue He, Kaidi Cao, Cheng Li, Chen Change Loy

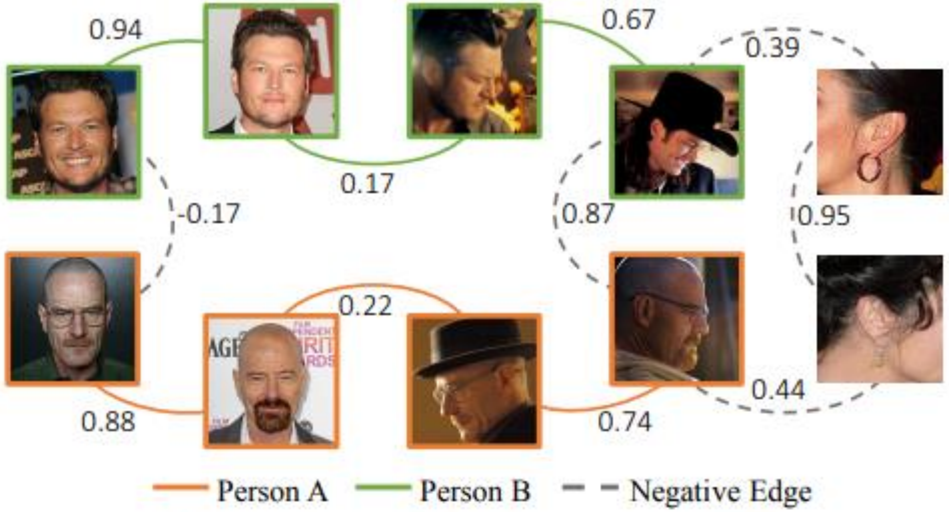


Figure 1. **Cosine angle in a deep feature space.** We measure the cosine angle between the deep feature vector of two faces. It is noted that even for two men with significantly different appearances, the angle between their profiles and noise faces (gray dash lines with 0.39∼0.44) is much larger than one's frontal and his own profile (0.17 and 0.22).

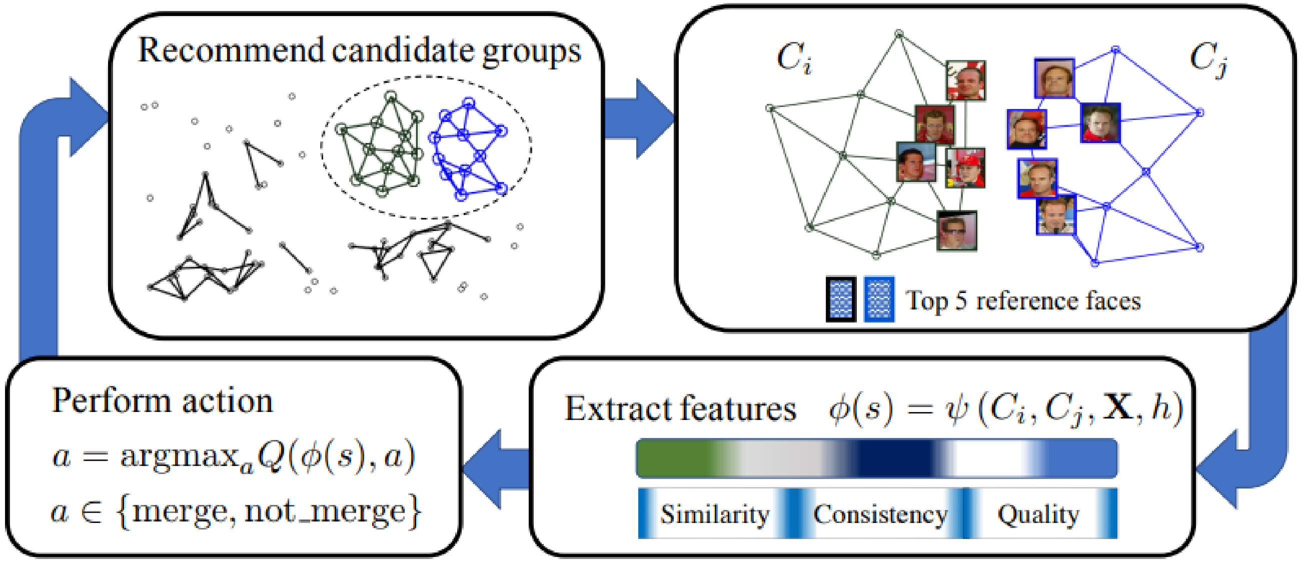


Figure 2. Face grouping by the proposed framework.

[58] Merge or Not? Learning to Group Faces via Imitation Learning

ue He, Kaidi Cao, Cheng Li, Chen Change Loy

Problem

: Clustering-Profile과 uninteresting faces & noise detection → face clustering이 어려움

Strategy

- IRL(Inverse Reinforcement Learning)
  - 단기, 장기 보상에 의해 2개 faces instance/group을 병합할 시기를 동적으로 결정(순차적으로)

Background technique

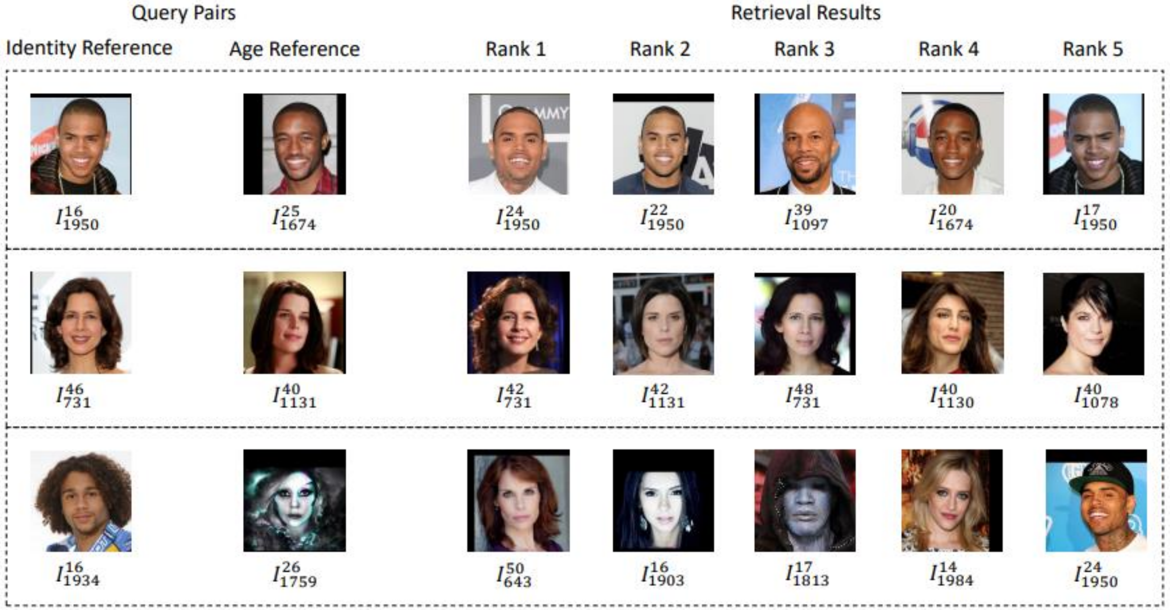
: Reinforcement Learning

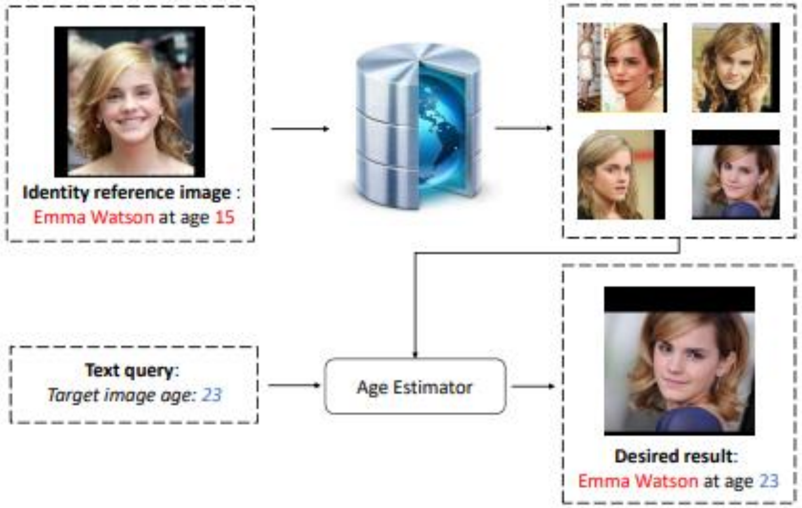
Dataset

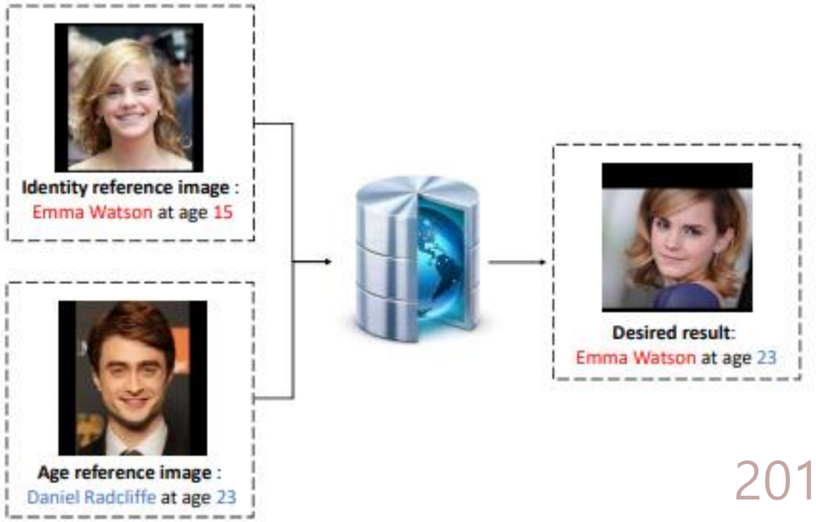
: GFW, LFW, ACCIO, GDL

[59] Dual-Reference Face Retrieval

BingZhang Hu, Feng Zheng, Ling Shao







201810957 염지현

[59] Dual-Reference Face Retrieval

BingZhang Hu, Feng Zheng, Ling Shao

Problem
: Face Retrieval – 특정 나이 사람의 얼굴 이미지 검색 불가(다른 사람의 이미지를 보여줌)

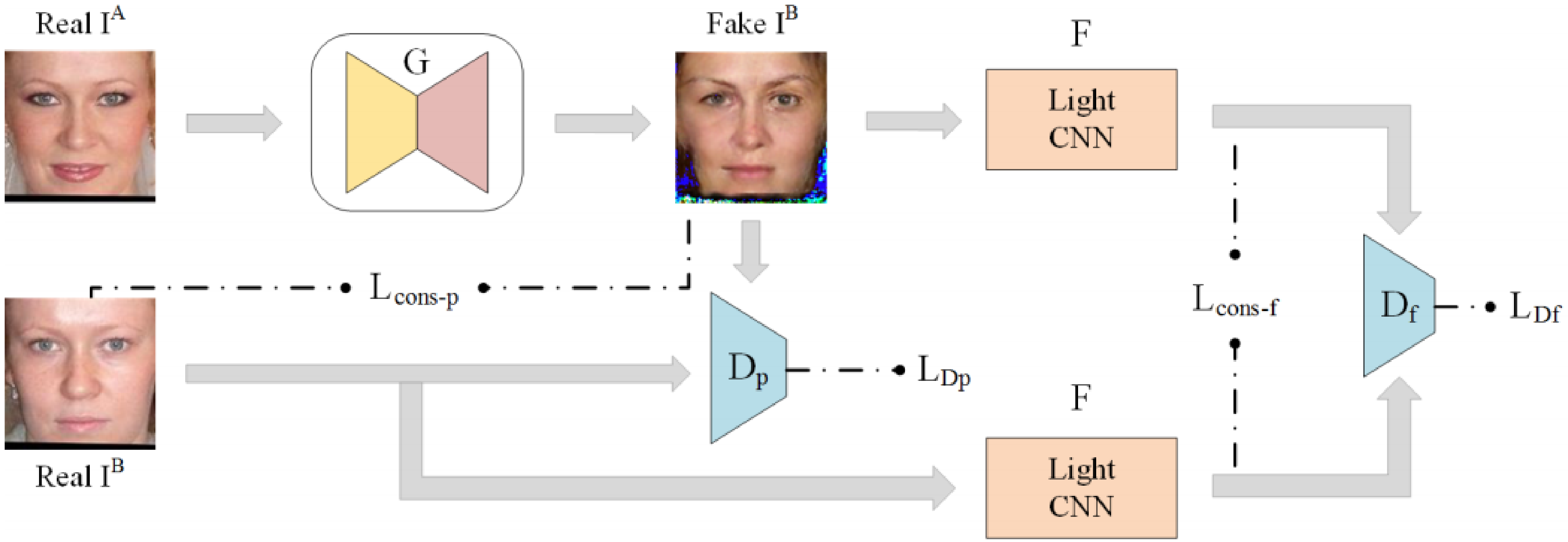
Strategy
- Dual-reference Face Retrieval
    - Input 2개
        - Target identity를 나타내는 identity reference image
        - Target age를 나타내는 age reference image
    - Raw image로부터 age, identity를 보존하여 joint manifold에 투영
    - Quarter-based model을 사용하여 age, identity 각각 유사성 metric 활용하여 최적화

Background technique
: DNN

Dataset
: DRFB, CACD, MORPH

[60] Anti-Makeup: Learning A Bi-Level Adversarial Network for Makeup-Invariant Face Verification

Yi Li, Lingxiao Song, Xiang Wu, Ran He, Tieniu Tan



201810957 염지현

[60] Anti-Makeup: Learning A Bi-Level Adversarial Network for Makeup-Invariant Face Verification

Yi Li, Lingxiao Song, Xiang Wu, Ran He, Tieniu Tan

Problem
: makeup한 face와 하지 않은 face 일치시키기

Strategy
- BLAN(bi-level adversarial network)
  - Makeup image에서 makeup 안 한 이미지 생성 후 합성된 메이크업 이미지를 사용하여 추가 확인
  - 2개의 adversarial network
    - 메이크업 한 face 재구성
    - Identity information preserving

Background technique
: GAN

Dataset
: FAM

[61] Multi-Scale Face Restoration with Sequential Gating Ensemble Network

Jianxin Lin, Tiankuang Zhou, Zhibo Chen

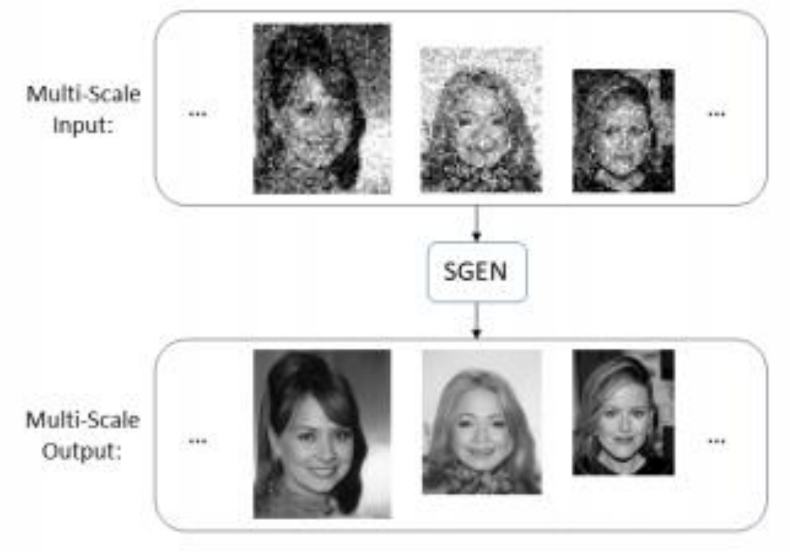
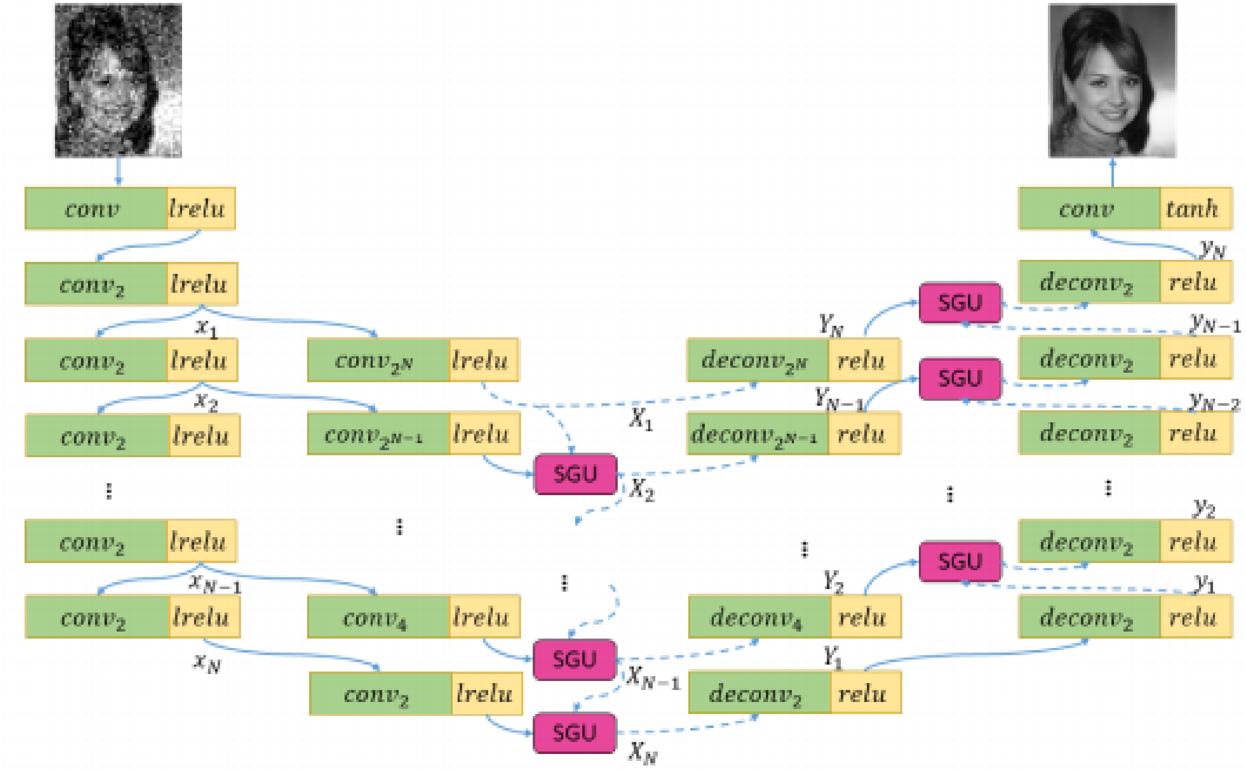


Figure 1: Illustration of our SGEN on multi-scale face restoration. The multi-scale noise corrupted LR face images are up-sampled to the size of ground truth before being fed into network.



[61] Multi-Scale Face Restoration with Sequential Gating Ensemble Network

Jianxin Lin, Tiankuang Zhou, Zhibo Chen

Problem
: Face Restoration(from distortions)

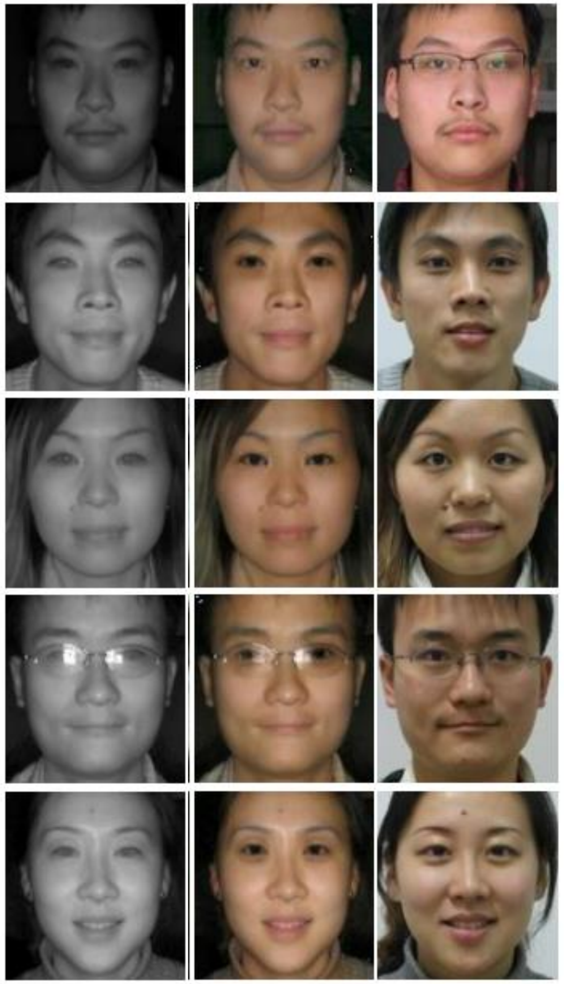
Strategy
- SGEN(Sequential Gating Ensemble Network)
  - 여러 수준의 base-ecoder/decoder에 순차적으로 data를 가져옴
  - 상향식 – base-encoder에서 고수준 정보를 순차적으로 추출
  - 하향식 – based-decoder에서 저수준 정보를 복원
  - SGU(Seqential Gating Unit)
    - 상향식, 하향식 정보 조합 및 선택
    - 2개의 서로 다른 level 입력을 받고 active input에 의해 출력 결정

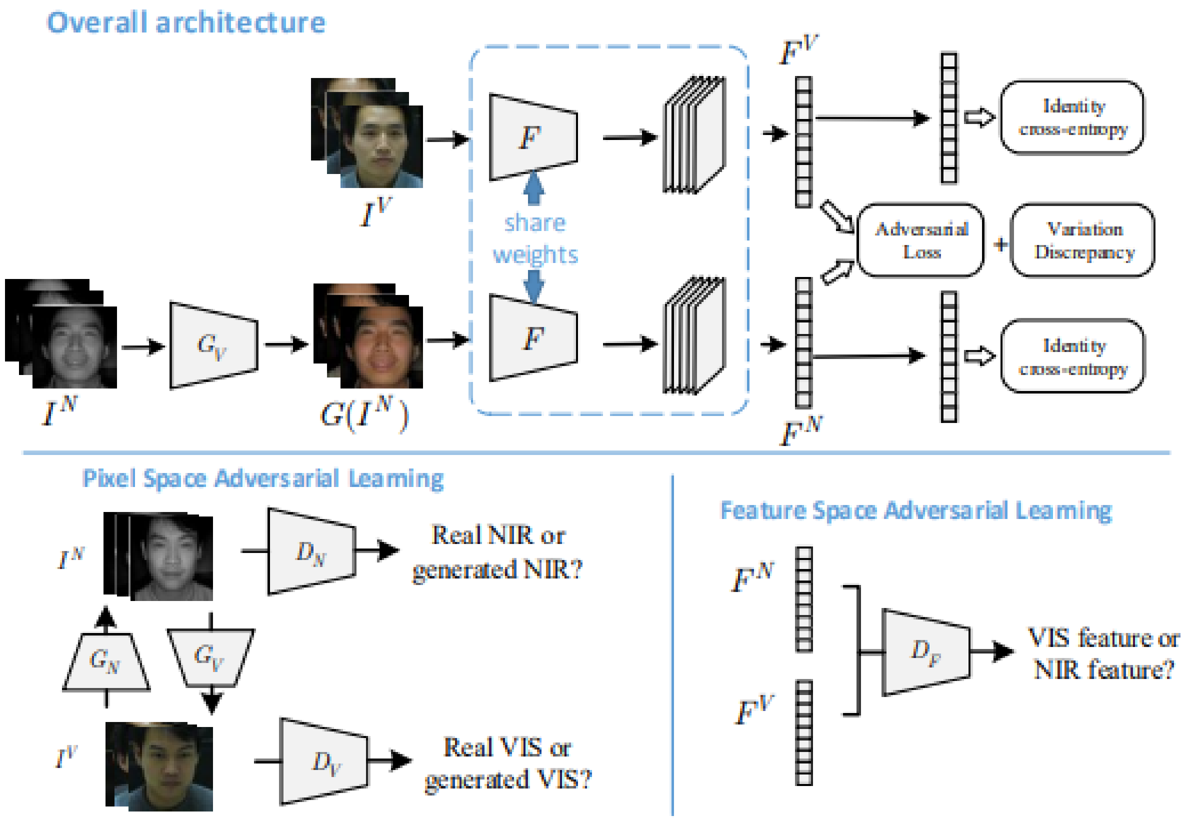
Background technique
: CNN

Dataset
: CelebA

[62] Adversarial Discriminative Heterogeneous Face Recognition

Lingxiao Song, Man Zhang, Xiang Wu, Ran He





201810957 염지현

[62] Adversarial Discriminative Heterogeneous Face Recognition

Lingxiao Song, Man Zhang, Xiang Wu, Ran He

Problem
: Heterogeneous Face Recognition

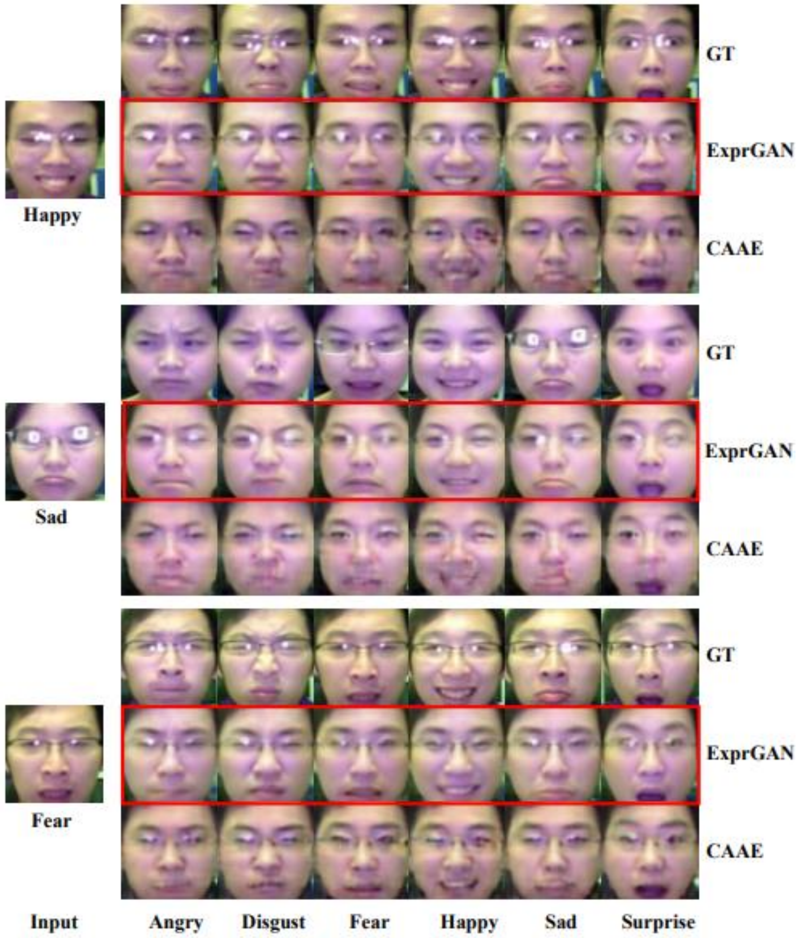
Strategy
- Adversarial Discriminative Feature Learning Framework
  - Row-pixel space & compact feature space를 adversarial learning을 통해 sensing gap 줄이기
  - Pixel space : GAN 사용(cross-spectral face hallucination 수행하기 위해)
  - Feature space : adversarial loss & high-order variance discrepancy loss
    → 두 분포 간의 불일치 측정

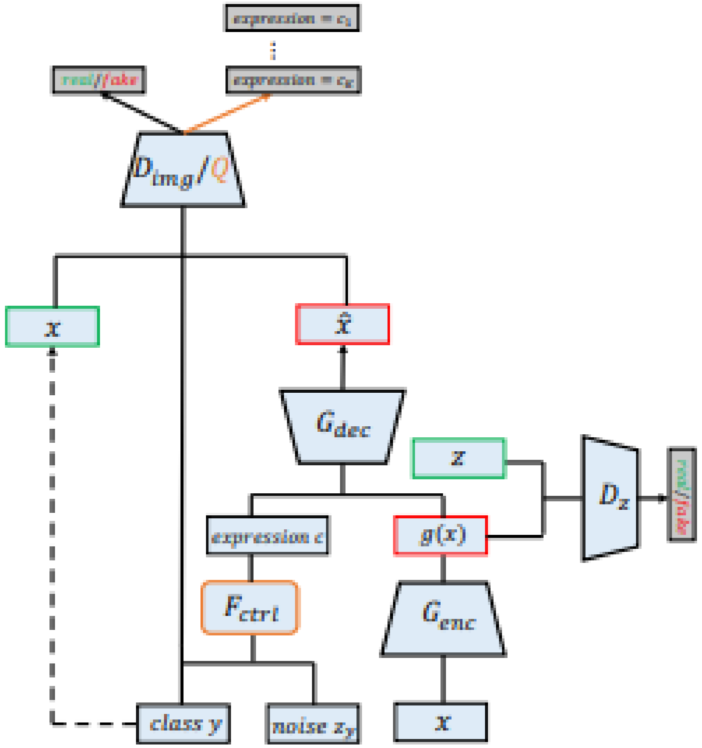
Background technique
: CNN, GAN

Dataset
: NIR-VIS

[63] ExprGAN: Facial Expression Editing with Controllable Expression Intensity

Hui Ding, Kumar Sricharan, Rama Chellappa





(c) ExprGAN

201810957 염지현

[63] ExprGAN: Facial Expression Editing with Controllable Expression Intensity

Hui Ding, Kumar Sricharan, Rama Chellappa

Problem
: Facial Expression Editing

Strategy
- ExprGAN(Expression GAN)
    - Expression controller module은 encode/decode network에 추가되어 표현적이고 간결한 코드를 학습하도록 설계

Background technique
: GAN

Dataset
: Oulu-CASIA

[64] Facial Landmarks Detection by Self-Iterative Regression based Landmarks-Attention Network

Tao Hu, Honggang Qi, Jizheng Xu, Qingming Huang

Figure 9: Several facial landmarks detection results in 300-W public testing and competition testing set. Blue dot in each sub-picture indicates ground truth landmarks location and yellow dot indicates the predicted location of SIR. Pictures for the five rows are from HELEN testing set, LFPW testing set, IBUG set, 300-W competition testing Indoor and Outdoor set respectively.

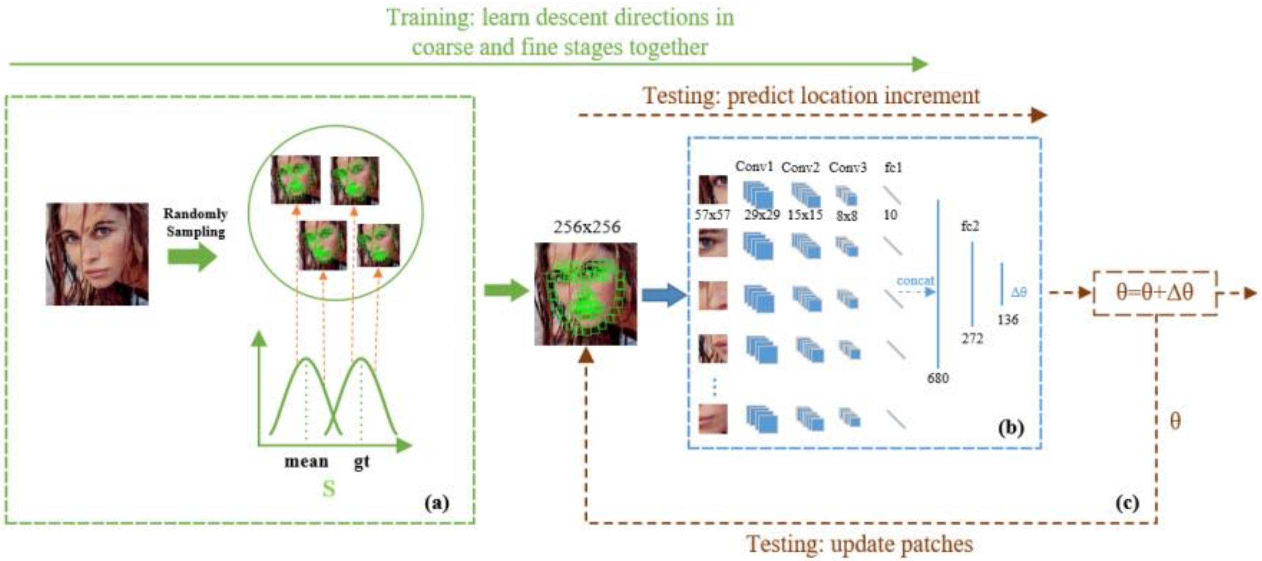


Figure 2: Training and testing process of the proposed SIR. (a) random sampling process. (b) Landmarks-Attention Network. (c) Iterative predicting and updating process. The training process consists of (a) and (b), while the testing process consists of (b) and (c). In the figure, one of the dimension of facial Landmarks Model parameter S is showed, and θ is landmarks' location parameter.

[64] Facial Landmarks Detection by Self-Iterative Regression based Landmarks-Attention Network

Tao Hu, Honggang Qi, Jizheng Xu, Qingming Huang

Problem

: Facial Landmarks Detection
: Cascaded Regression의 한계 : 각 단계 regressor's 학습 데이터 = 이전 단계의 출력 → robust하지 못함

Strategy

- SIR(Self-Iterative Rgression)
  - 단 하나의 regressor가 거친 단계에서 미세한 단계까지 samples의 하강 방법 학습
  - Parameters 반복적으로 update
- LAN(Landmarks-Attention Network)
  - Landmark 주변의 feature 학습
  - Obtain holistic location increment

Background technique

: CNN

Dataset

: 300-W, AFW, HELEN, LFPW

[65] Brute-Force Facial Landmark Analysis With A 140,000-Way Classifier

Mengtian Li, Laszlo Jeni, Deva Ramanan

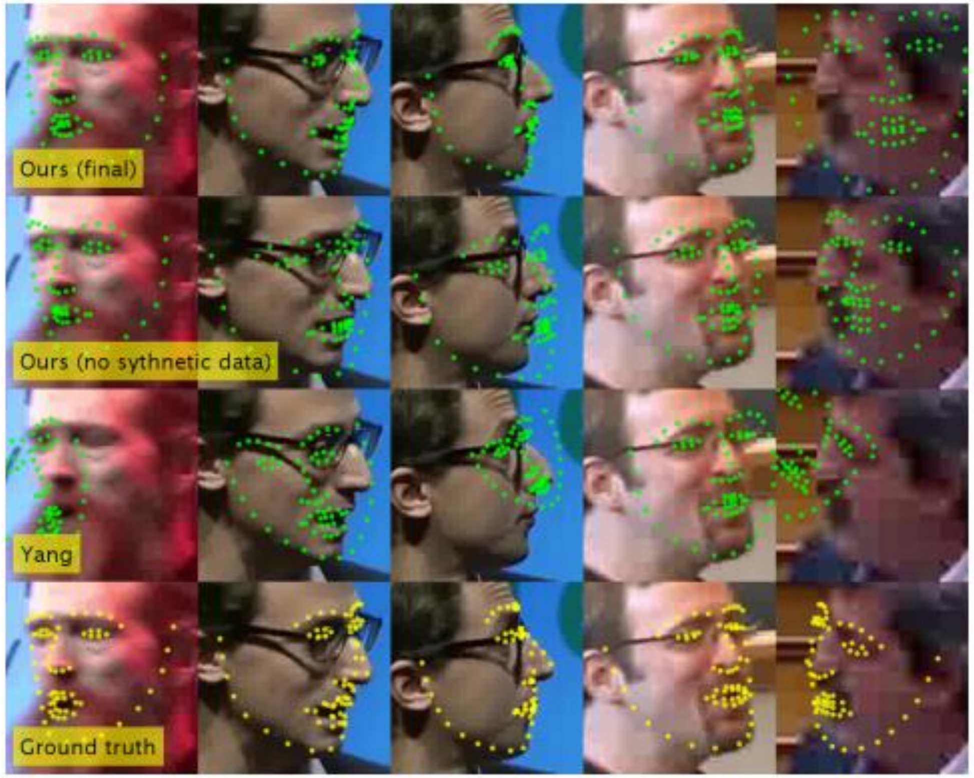


Figure 9: Qualitative results for Category 3-Hard frames on 300VW. We can see that by training on synthetic images (comparing Row 1 with Row 2), our method is robust to large pose variation. The last column shows a failure case.

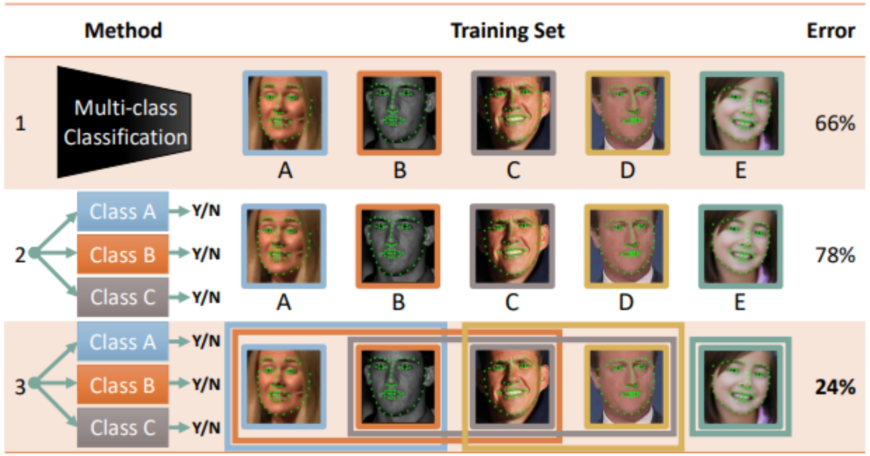


Figure 5: Scaling up the number of classes in a classification network. This figure shows three different ways of training a multi-class classifier, with the mean validation error of the landmarks shown in the last column. The error is shown as a percentage of the error of a random classifier. We adopt the third method for our approach, where we train independent binary classifiers with *example sharing*. The colors in the figure denote classes and the boxes in the last row circle the training examples used for a particular class (see the matching color). For example, the blue box denotes that the images of class A and B are used as the *positive* examples for training class A.

201810957 염지현

[65] Brute-Force Facial Landmark Analysis With A 140,000-Way Classifier

Mengtian Li, Laszlo Jeni, Deva Ramanan

Problem

: Facial Landmark Analysis

Strategy

- 140,000 Way classifier
  - Class의 수 = 훈련 예제의 수
  - Multi-label loss function 사용 – training examples을 개별 class에서 균일하지 않게 공유

Background technique

: Classifier

Dataset

: 300-W, AFW, HELEN, LFPW, IBUG, 300W-LP

Face & Facial

ECCV

201810957 염지현

[66] Face Recognition with Contrastive Convolution

Chunrui Han, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen

Fig. 1. Illustration of how we humans do face verification by focusing on distinct face characteristics when the same face A is compared with different persons. (a) When comparing A with B_1 who features small eyes, our focus is attracted to regions around the eyes of A; (b) when comparing A with B_2 whose face is round, we pay more attention to the contour of A. This reveals that a face should be described differently by using contrastive charateristics for example, when being compared with different persons.

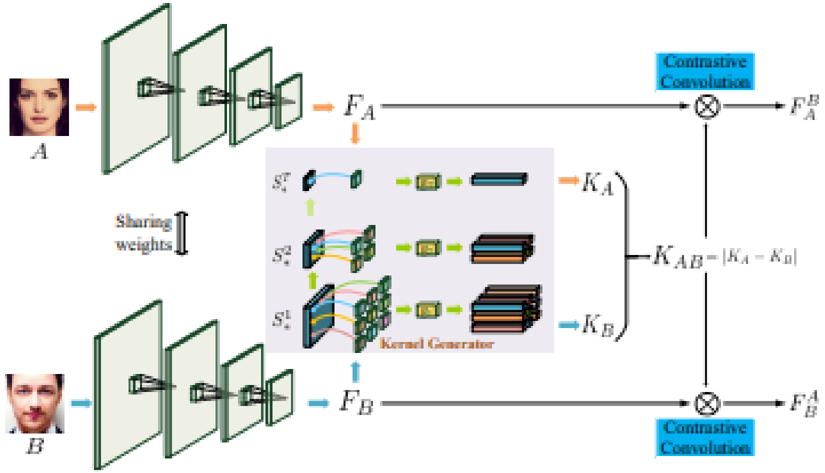


Fig. 2. The pipeline of our contrastive CNN. Given a pair of face images, A and B, a common feature extractor C consisting of several cascaded convolution layers is firstly used to obtain expressive feature representations F_A and F_B of them. Then, the kernel generator G consisting of several sub-generators generates personalized kernels for A and B respectively, based on which the contrastive kernels are achieved as $|K_A - K_B|$. Finally, with those contrastive kernels, the contrastive features of A and B are extracted via convolution operations respectively for the final similarity calculation. Note the subscript $*$ of S in kernel generator can be A or B.

[66] Face Recognition with Contrastive Convolution

Chunrui Han, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen

<u>Problem</u>

: Face Recognition
: 기존 Face Recognition은 비교할 이미지 두 장에 동일한 filter 적용 → 얼굴의 표현은 고정된 상태로 유지

<u>Strategy</u>

- Contrastive Convolution(두 얼굴 사이 대조적인 특징 비교)

<u>Background technique</u>

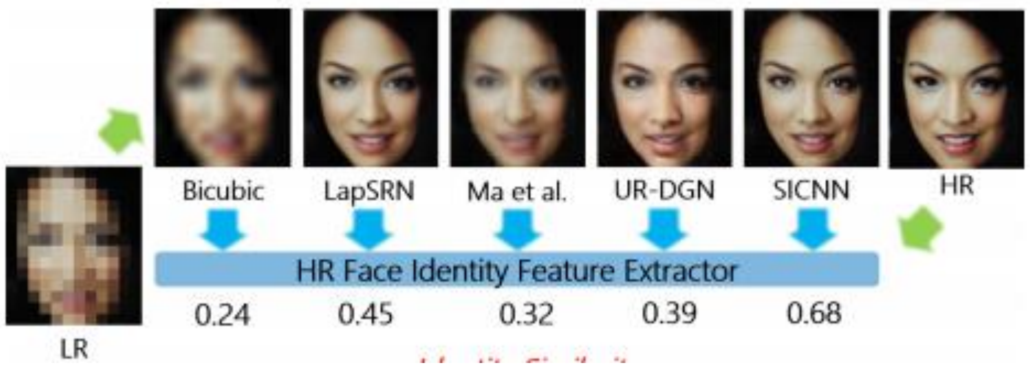
: CNN

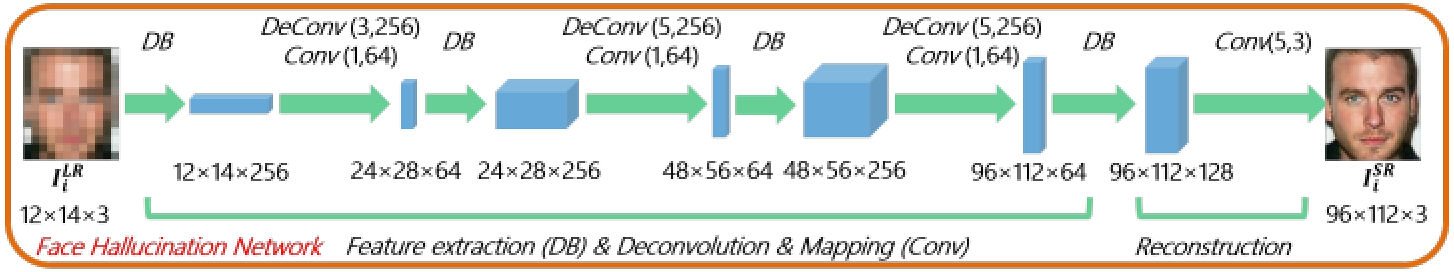
<u>Dataset</u>

: LFW, IJB-A

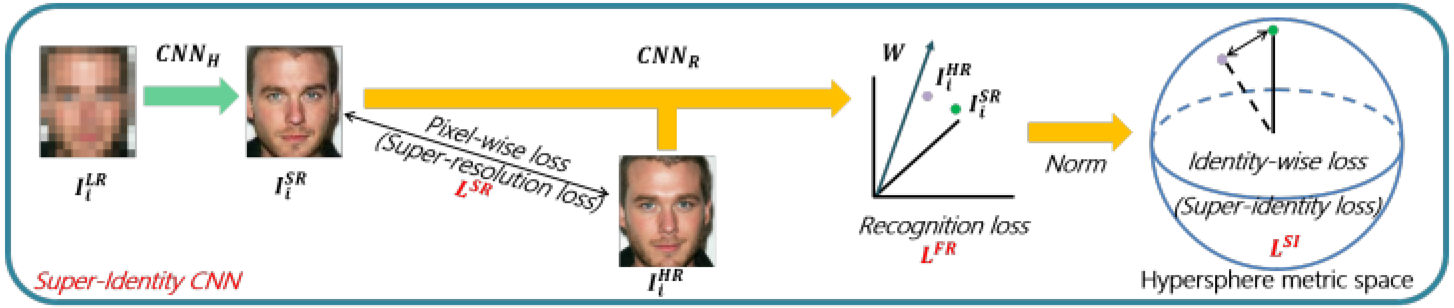
[67] Super-Identity Convolutional Neural Network for Face Hallucination

Kaipeng Zhang, Zhanpeng Zhang, Chia-Wen Cheng, Winston H. Hsu, Yu Qiao, Wei Liu, Tong Zhang





(a) Network architecture of hallucination model (CNN_H)



(b) Illustration of the proposed super-identity CNN

[67] Super-Identity Convolutional Neural Network for Face Hallucination

Kaipeng Zhang, Zhanpeng Zhang, Chia-Wen Cheng, Winston H. Hsu, Yu Qiao, Wei Liu, Tong Zhang

Problem
: Identity를 preserve하면서 Face Hallucination

Strategy
- SICNN(Super-Identity CNN)
  - Super-identity loss : hallucinated face와 high-resolution face 사이의 차이 측정
  - Domain-integrated Training : 두 도메인 간의 margin이 엄청 커서 발산 가능성이 있음
    → 두 도메인 얼굴 ID metric을 구성하여 발산 방지

Background technique
: CNN

Dataset
: LFW, YTF

[68] Face Super-resolution Guided by Facial Component Heatmaps

Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli, Richard Hartley

Fig. 1. Comparison of state-of-the-art face super-resolution methods on very low-resolution (LR) face images. Columns: (a) Unaligned LR inputs. (b) Original HR images. (c) Nearest Neighbors (NN) of aligned LR faces. Note that image intensities are used to find NN. (d) CBN [6]. (e) TDAE [7]. (f) TDAE†. We retrain the original TDAE with our training dataset. (g) Our results.

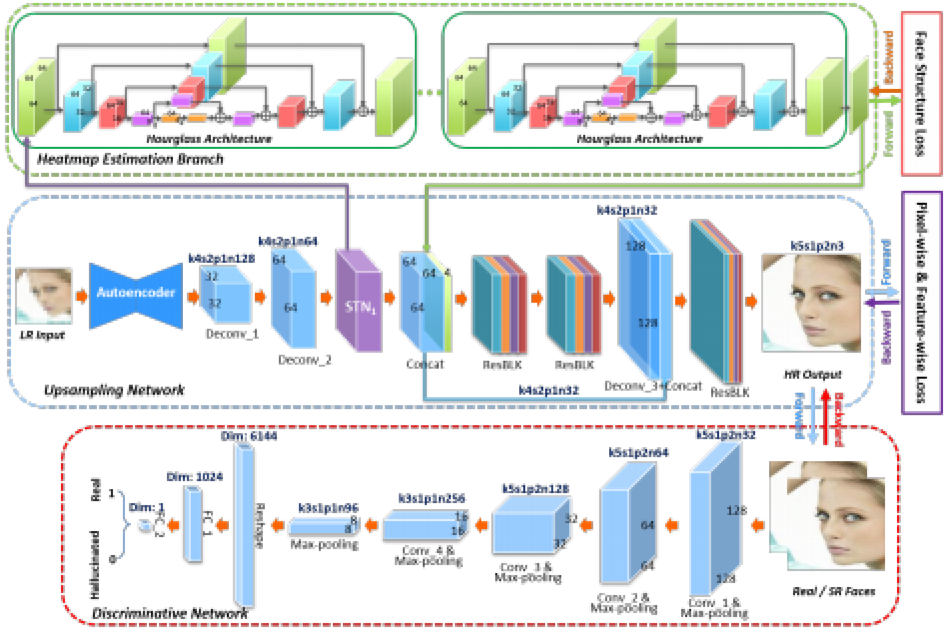


Fig. 2. The pipeline of our multi-task upsampling network. In the testing phase, the upsampling branch (blue block) and the heatmap estimation branch (green block) are used.

[68] Face Super-resolution Guided by Facial Component Heatmaps

Xin Yu, Basura Fernando, Bernard Ghanem, Fatih Porikli, Richard Hartley

Problem
: Face Super-resolution

Strategy
- Multi-task CNN
  - two branches
    - Super-resolving face image
    - 얼굴의 두드러진 영역 예측(heatmap)
  - 얼굴 구성 요소의 공간적 제약 추가 탐지

Background technique
: CNN

Dataset
: CelebA, Menpo

[69] Learning Warped Guidance for Blind Face Restoration

Xiaoming Li, Ming Liu, Yuting Ye, Wangmeng Zuo, Liang Lin, Ruigang Yang

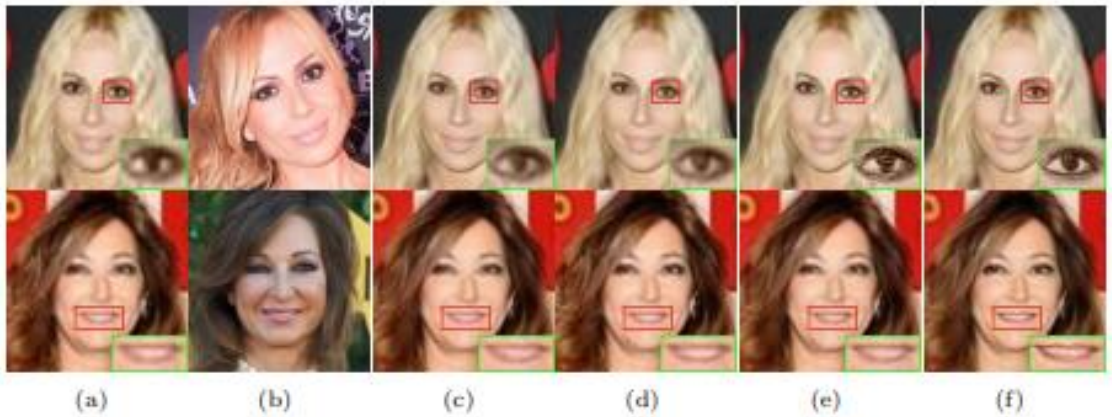


Fig. 1. Restoration results on real low quality images: (a) real low quality image, (b) guided image, and the results by (c) U-Net [1] by taking low quality image as input, (d) U-Net [1] by taking both guided image and low quality image as input, (e) our GFRNet without landmark loss, and (f) our full GFRNet model. Best viewed by zooming in the screen.

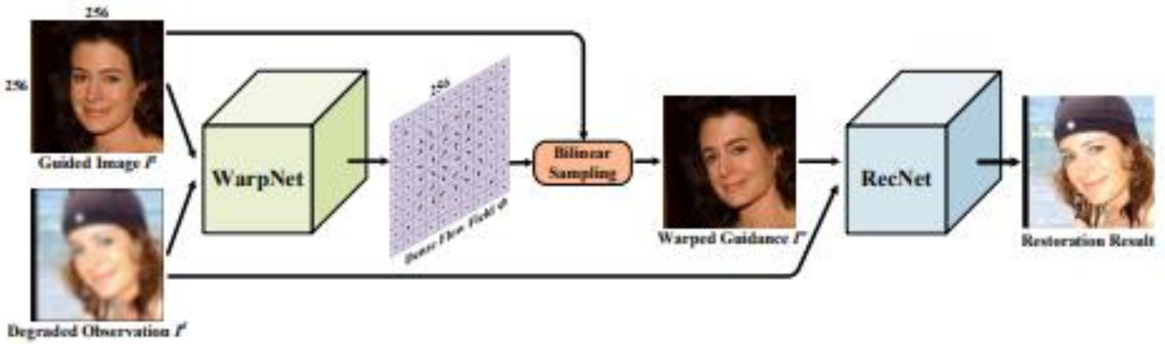


Fig. 2. Overview of our GFRNet. The WarpNet takes the degraded observation I^d and guided image I^g as input to predict the dense flow field Φ, which is adopted to deform I^g to the warped guidance I^w. I^w is expected to be spatially well aligned with ground-truth I. Thus the RecNet takes I^w and I^d as input to produce the restoration result $\hat{I}$.

201810957 염지현

[69] Learning Warped Guidance for Blind Face Restoration

Xiaoming Li, Ming Liu, Yuting Ye, Wangmeng Zuo, Liang Lin, Ruigang Yang

Problem
: Face Restoration

Strategy
- GFRNet(Guidance Face Restoration Network)
    - GFRNet = WrapNet + RecNet
    - WrapNet : 가이드 이미지를 왜곡하여 자세와 표현을 수정하여 왜곡된 유도 예측
    - RecNet : 복원 결과 생성

Background technique
: GAN

Dataset
: CASIA-WebFace, VggFace2

[70] Face De-Spoofing: Anti-Spoofing via Noise Modeling

Amin Jourabloo, Yaojie Liu, Xiaoming Liu

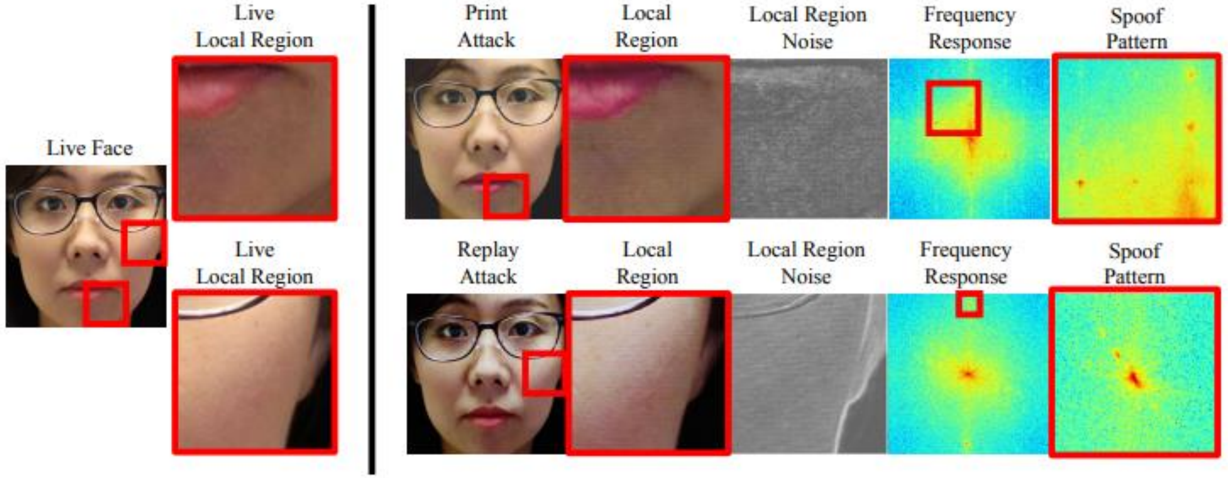


Fig. 2. The illustration of the spoof noise pattern. **Left:** live face and its local regions. **Right:** Two registered spoofing faces from print attack and replay attack. For each sample, we show the local region of the face, intensity difference to the live image, magnitude of 2D FFT, and the local peaks in the frequency domain that indicates the spoof noise pattern. Best viewed electronically.

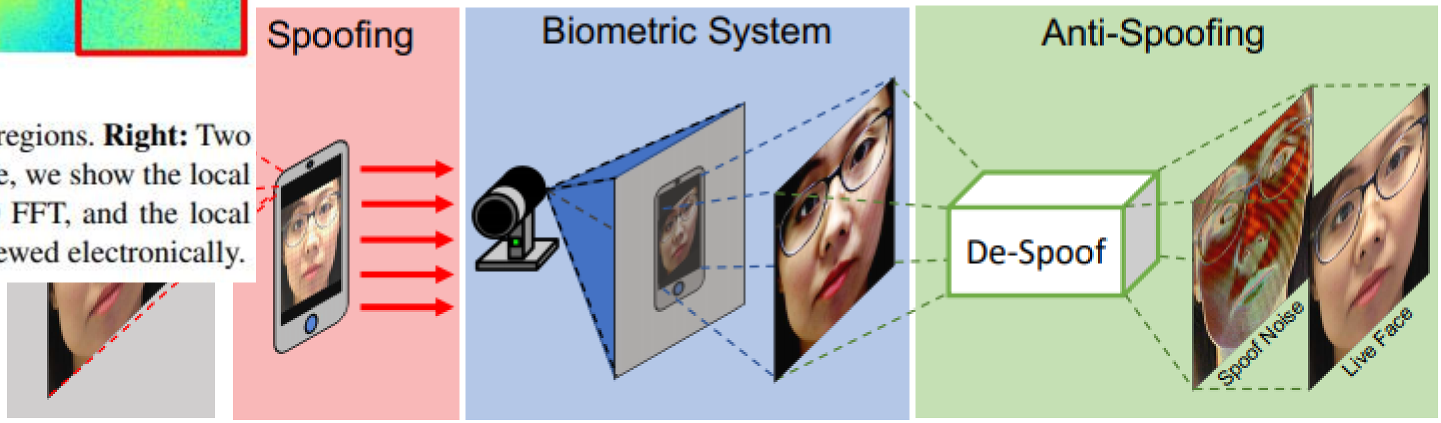


Fig. 1. The illustration of face spoofing and anti-spoofing processes. De-spoofing process aims to estimate a spoof noise from a spoof face and reconstruct the live face. The estimated spoof noise should be discriminative for face anti-spoofing.

201810957 염지현

[70] Face De-Spoofing: Anti-Spoofing via Noise Modeling

Amin Jourabloo, Yaojie Liu, Xiaoming Liu

Problem
: Face Anti-Spoofing(live faces와 spoof faces 간의 미묘한 차이)

Strategy
- Nose Modeling
- 이미지를 분할하지 않고 전체적으로 처리

Background technique
: CNN

Dataset
: 언급 X

[71] Visual Psychophysics for Making Face Recognition Algorithms More Explainable

Brandon RichardWebster, So Yon Kwon, Christopher Clarizio, Samuel E. Anthony, Walter J. Scheirer

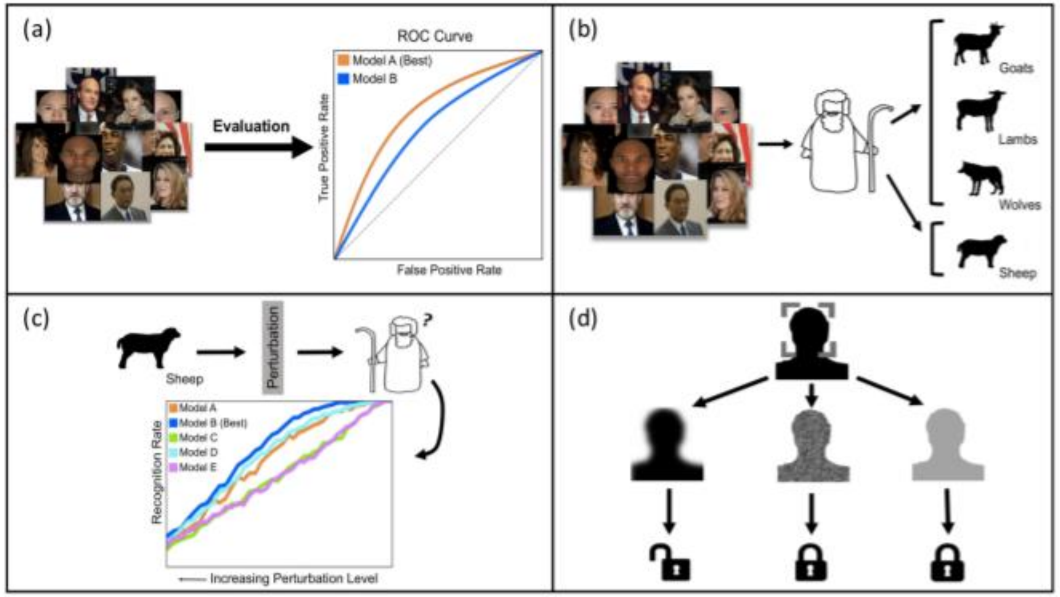


Fig. 1. Visual Pyschophysics [3,4,5] helps us explain algorithm behavior in a way that traditional dataset evaluation (a) cannot. Our proposed methodology introduces a theoretical mapping between elements of psychopysical testing and the biometric menagerie paradigm [6], where a shepherd function first isolates cooperative users ("sheep") from all others (b). From a perfect matching scenario, the images of the sheep are incrementally perturbed using a chosen image transformation, and item-response curves are plotted so that points of failure can be identified (c). The results can then be used to explain why matching works for some input images, but not others (d).

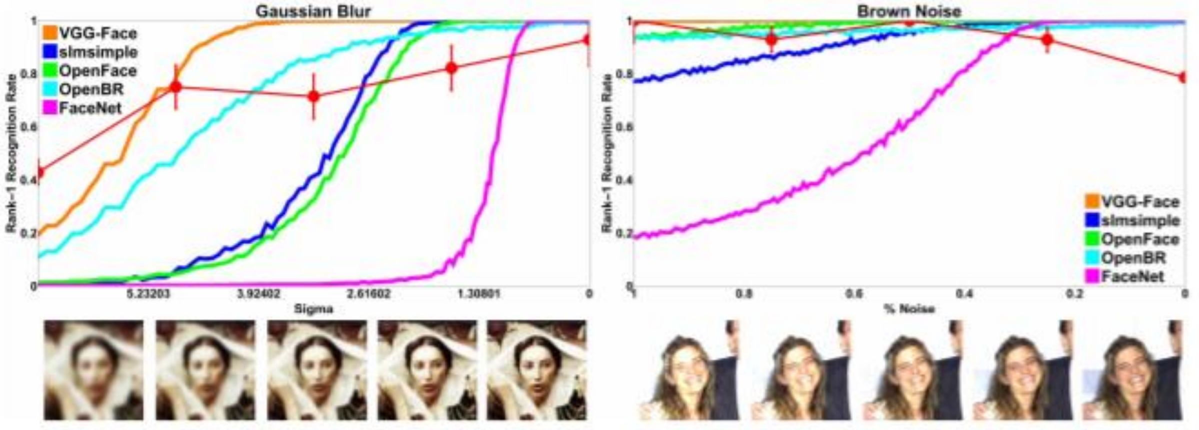


Fig. 2. A selection of item-response curves for the M-AFC task using data from the LFW dataset [64]. Each experiment used five different face recognition algorithms [59,60,61,62,63]. A perfect curve would be a flat line at the top of the plot. The images at the bottom of each curve show how the perturbations increase from right to left, starting with no perturbation (*i.e.*, the original image) for all conditions. The red dots indicate mean human performance for a selected stimulus level; error bars are standard error. Curves are normalized so chance is 0 on the y-axis. All plots are best viewed in color.

201810957 염지현

[71] Visual Psychophysics for Making Face Recognition Algorithms More Explainable

Brandon RichardWebster, So Yon Kwon, Christopher Clarizio, Samuel E. Anthony, Walter J. Scheirer

Problem

: Face Recognition 방법 → 특정 얼굴이 인식되지 않거나 사기꾼이 인식되는 이유가 명확하지 않음

Strategy

- Controlled manipulation of stimuli + model system에서 유발하는 반응에 대한 연구

Background technique

: CNN

Dataset

: 언급 X

[72] Semi-supervised Adversarial Learning to Generate Photorealistic Face Images of New Identities from 3D Morphable Model

Baris Gecer, Binod Bhattarai, Josef Kittler, Tae-Kyun Kim

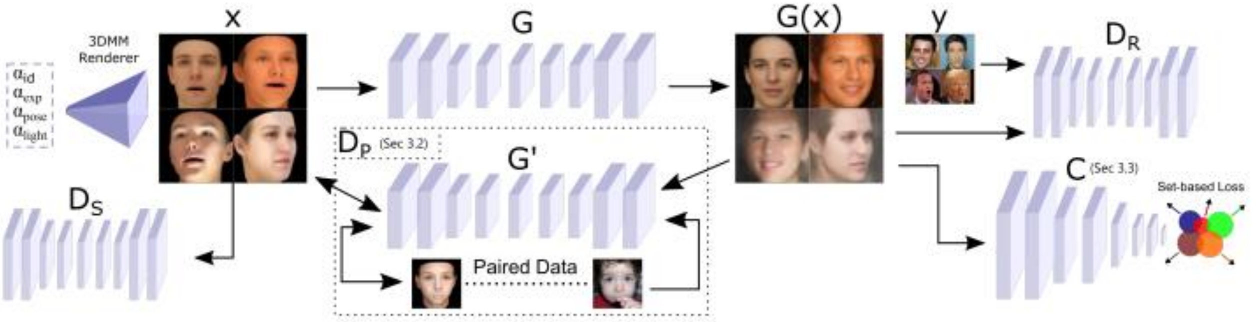


Fig. 1: Our approach aims to synthesize photorealistic images conditioned by a given synthetic image by 3DMM. It regularizes cycle consistency [63] by introducing an additional adversarial game between the two generator networks in an unsupervised fashion. Thus the under-constraint cycle loss is supervised to have correct matching between the two domains by the help of a limited number of paired data. We also encourage the generator to preserve face identity by a set-based supervision through a pretrained classification network.

[72] Semi-supervised Adversarial Learning to Generate Photorealistic Face Images of New Identities from 3D Morphable Model

Baris Gecer, Binod Bhattarai, Josef Kittler, Tae-Kyun Kim

Problem
: 3DMM에 의해 주어진 합성 이미지로부터 photorealistic image 합성

Strategy
- Semi-supervised Adversarial Learning Framework
  - Non-paired한 대량의 데이터 + paired한 소량의 데이터 → semi-supervised

Background technique
: GAN

Dataset
: Oxford VGG Face

[73] A Hybrid Model for Identity Obfuscation by Face Replacement

Qianru Sun, Ayush Tewari, Weipeng Xu, Mario Fritz, Christian Theobalt, Bernt Schiele

(a) Original hair (b) Blurred hair (c) Blacked-out hair

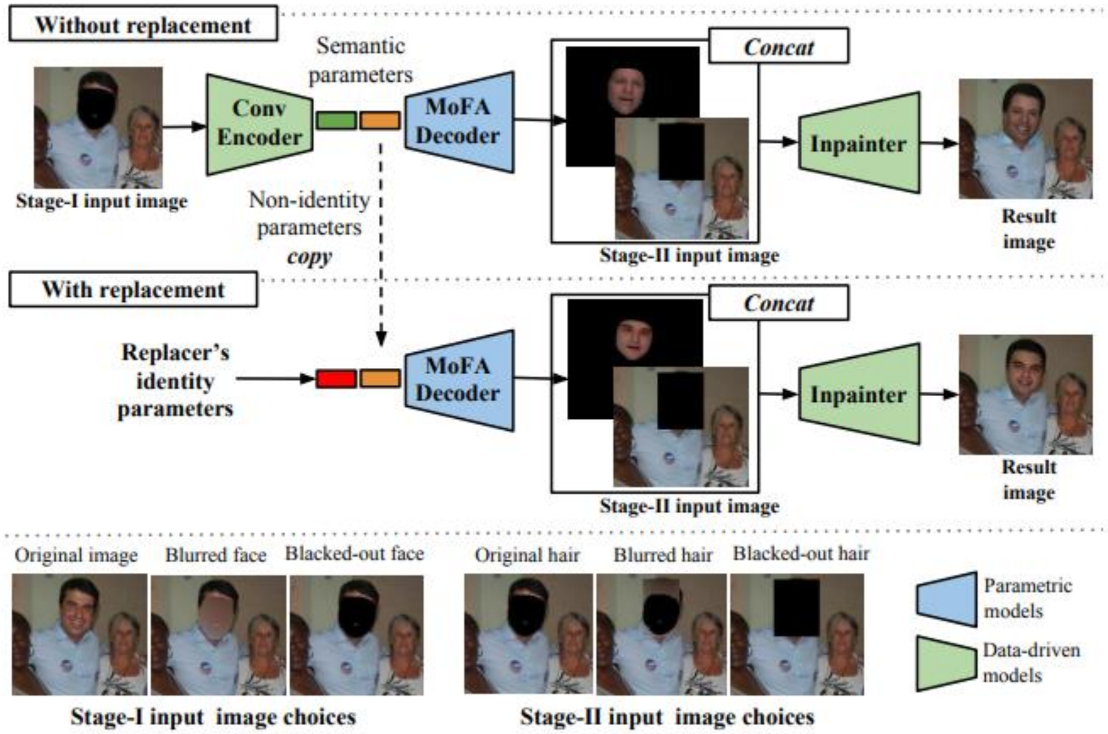


Fig. 1: Our obfuscation method based on data-driven deep models and parametric face models. The bottom row shows the input image choices for stage-I and stage-II. Different input combination results in different levels of obfuscation.

[73] A Hybrid Model for Identity Obfuscation by Face Replacement

Qianru Sun, Ayush Tewari, Weipeng Xu, Mario Fritz, Christian Theobalt, Bernt Schiele

Problem
: Identity Obfuscation ← Face Replacement를 통해 obfuscate identity

Strategy
- Hybrid Model
  - Parameteric : facial parameter control, identity 조작
  - Data 기반 측면 : 미세한 세부 사항 + 전반적인 사실감 + 장면 혼합

Background technique
: GAN

Dataset
: PIPA

[74] Consensus-Driven Propagation in Massive Unlabeled Data for Face Recognition

Xiaohang Zhan , Ziwei Liu, Junjie Yan , Dahua Lin , Chen Change Loy

Fig. 6: This figure shows two groups of faces in the unlabeled data. All faces in a group has the same identity according to the original annotations. The number on the top-left conner of each face is the label assigned by our proposed method, and the faces in red boxes are discarded by our method. The results suggest the high precision of our method in identifying persons of the same identity. Interestingly, our method is robust in pinpointing wrongly annotated faces (group 1), extremely low-quality faces (e.g., heavily blurred face, cartoon in group 2), which do not help training. See supplementary materials for more visual results.

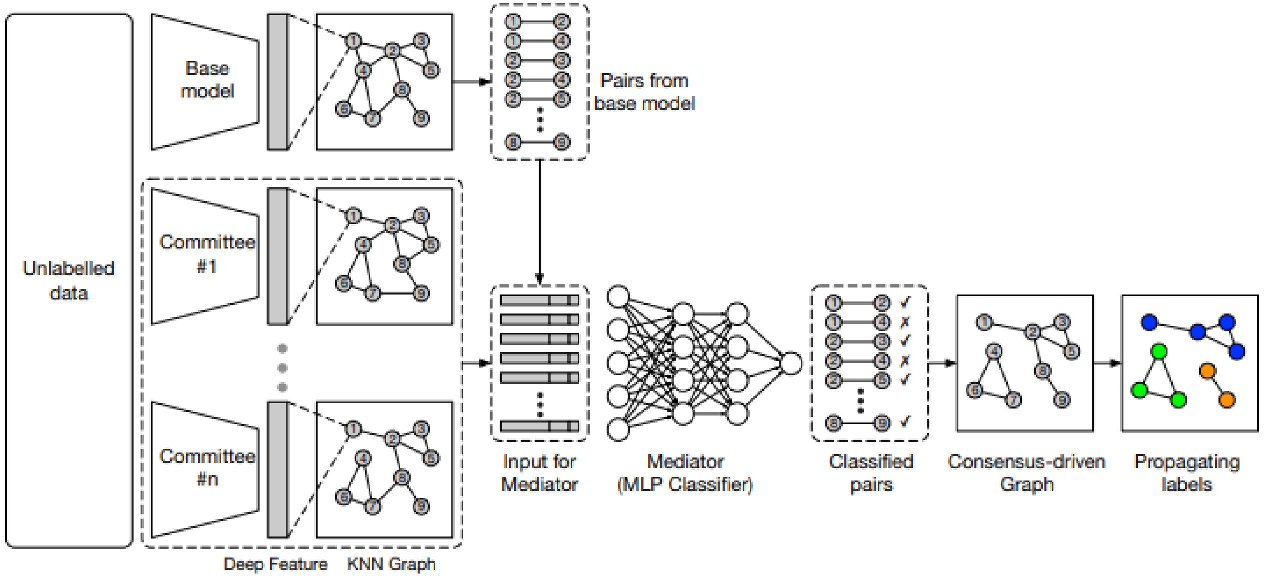


Fig. 1: **Consensus-Driven Propagation.** We use a base model and committee models to extract features from unlabeled data and create k-NN graphs. The input to the mediator is constructed by various local statistics of the k-NN graphs of the base model and committee. Pairs that are selected by the mediator compose the "consensus-driven graph". Finally, we propagate labels in the graph, and the propagation for each category ends by recursively eliminating low-confidence edges.

[74] Consensus-Driven Propagation in Massive Unlabeled Data for Face Recognition

Xiaohang Zhan , Ziwei Liu, Junjie Yan , Dahua Lin , Chen Change Loy

Problem

: Face Recognition – ID annotation 확장의 어려움(label이 없는 data가 많다)

Strategy

- CDP(Consensus-Driven Propagation)
  - Committee : 위원회
  - Mediator : 중재자

Background technique

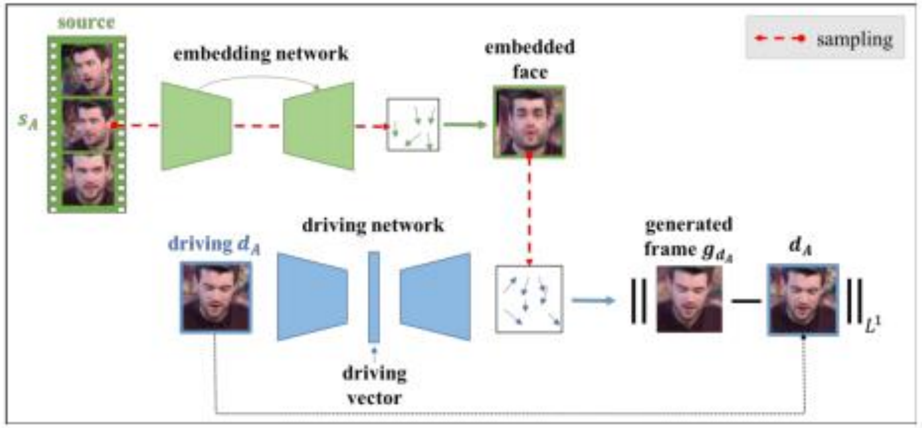
: CNN

Dataset

: 언급 X

[75] X2Face: A network for controlling face generation using images, audio, and pose codes

Olivia Wiles, A. Sophia Koepke, Andrew Zisserman



2: An overview of X2Face during the initial training stage. Given multiple nes of a video (here 4 frames), one frame is designated the *source* frame and ther the *driving* frame. The *source* frame is input to the *embedding network*, which learns a sampler to map pixels from the *source* frame to the *embedded* face. The *driving* frame is input to the *driving network*, which learns to map pixels from the *embedded* face to the *generated* frame. The *generated* frame should have the identity of the *source* frame and the pose/expression of the *driving* frame. In this training stage, as the frames are from the same video, the *generated* and *driving* frames should match. However, at test time the identities of the *source* and *driving* face can differ.

[75] X2Face: A network for controlling face generation using images, audio, and pose codes

Olivia Wiles, A. Sophia Koepke, Andrew Zisserman

Problem
: Controlling face generation

Strategy
- X2Face – 주어진 얼굴의 pose와 expression을 control
  - Source frame ID + 구동 frame 표정/포즈

Background technique
: GAN

Dataset
: VGG-Face, VoxCeleb, AFLW, LRW

[76] Dependency-aware Attention Control for Unconstrained Face Recognition with Image Sets

Xiaofeng Liu, B.V.K Vijaya Kumar, Chao Yang, Qingming Tang, Jane You

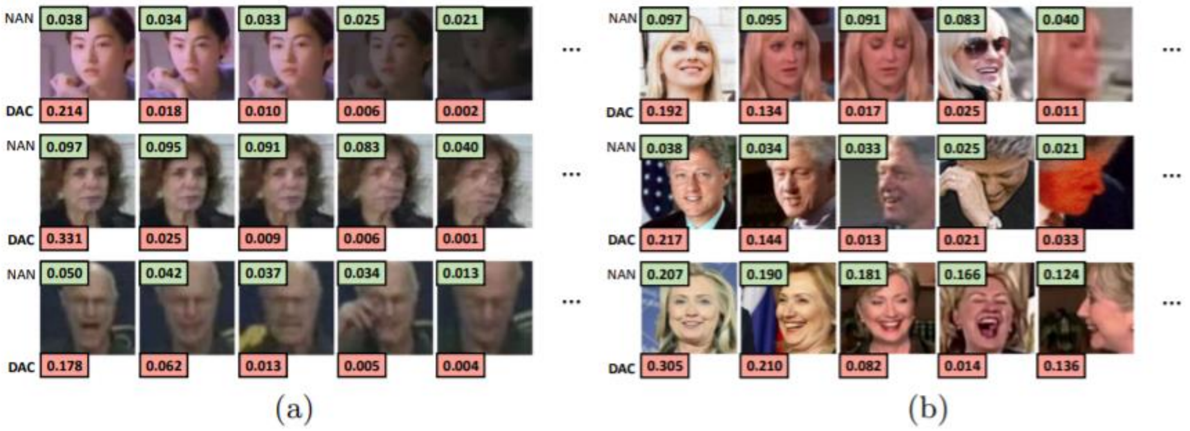


Fig. 2. Typical examples on the test set of (a) YTF and (b) IJB-A dataset showing the weights of images calculated by the previous method NAN [16], and proposed DAC.

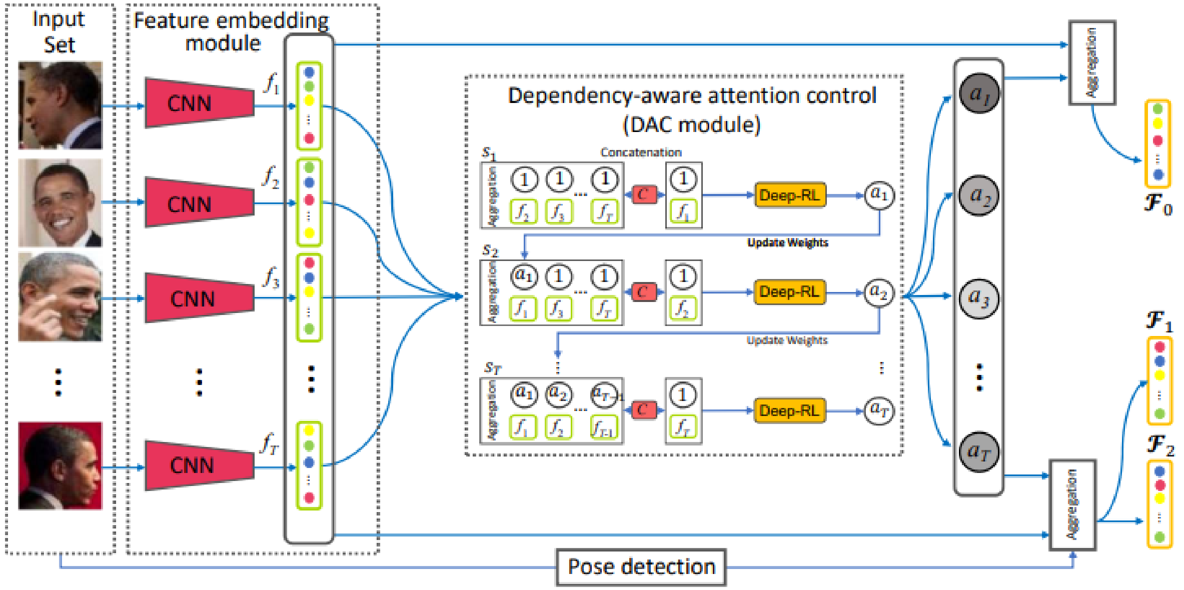


Fig. 3. Our network architecture for image set-based face recognition.

201810957 염지현

[76] Dependency-aware Attention Control for Unconstrained Face Recognition with Image Sets

Xiaofeng Liu, B.V.K Vijaya Kumar, Chao Yang, Qingming Tang, Jane You

Problem

: Unconstrained Face Recognition

Strategy

- DAC(Dependency-Aware Control)

Background technique

: CNN

Dataset

: YTFM LFW, IJB-A

[77] Using LIP to Gloss Over Faces in Single-Stage Face Detection Networks

Siqi Yang, Arnold Wiliem, Shaokang Chen, Brian C. Lovell

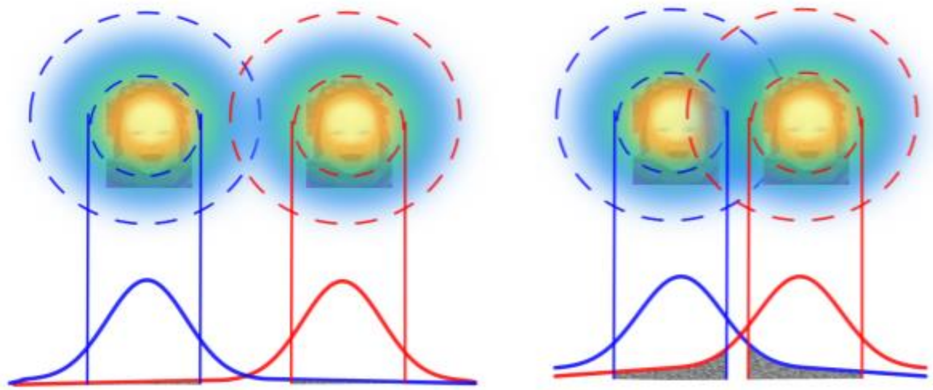
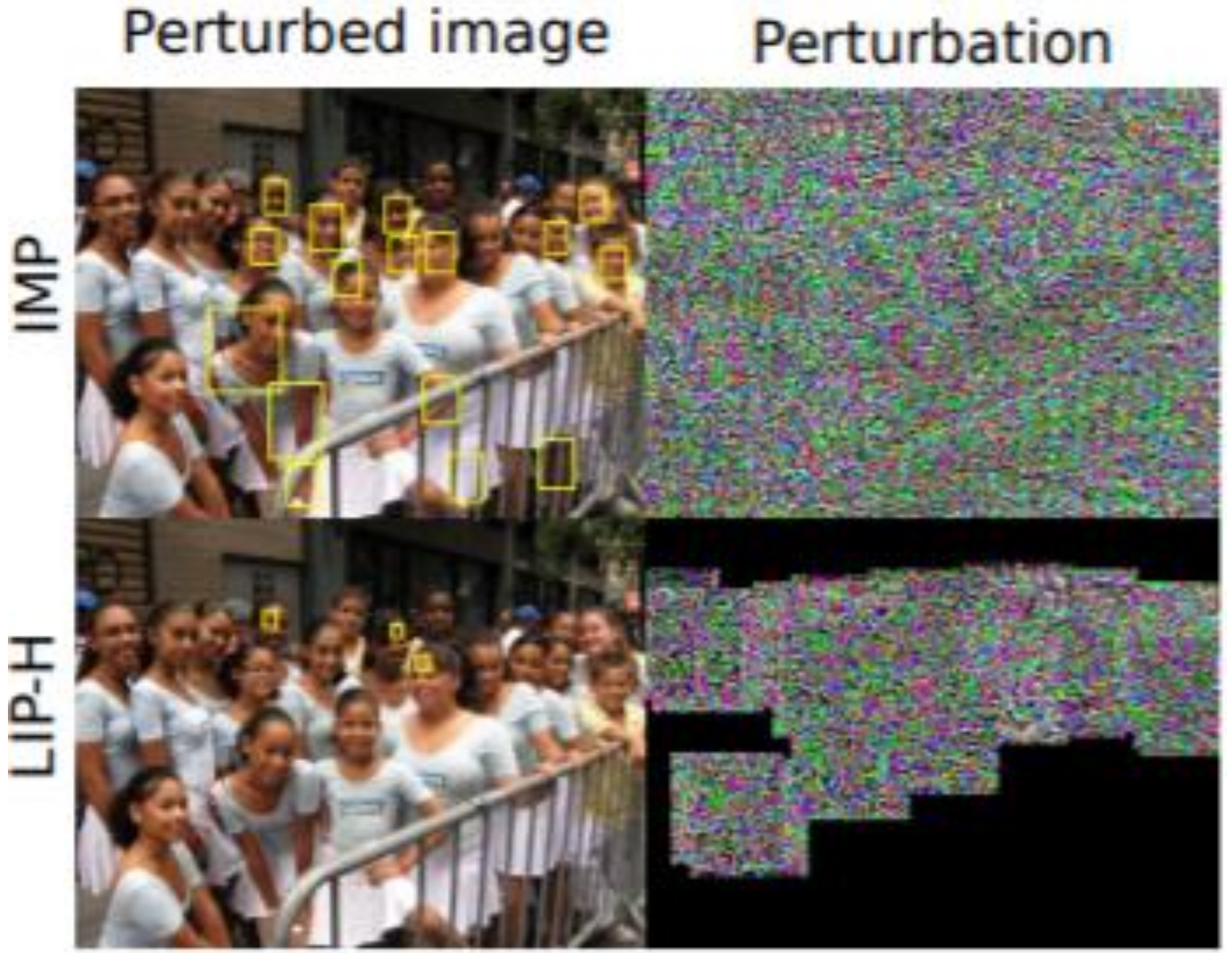


Figure 1: An illustration of the Instance Perturbation Interference (IPI) problem. Upper row: two instances with their generated adversarial perturbations. The outer and inner circles indicate the Theoretical Receptive Field (TRF) and Effective Receptive Field (ERF), respectively. Lower row: one dimensional representation of the perturbations. IPI problem refers to the perturbation generated for one instance significantly disrupting the perturbation generated for the other instance. The disruption does not have significant effect on the left case, whereas on the right case, it will reduce the effectiveness of the attack.



[77] Using LIP to Gloss Over Faces in Single-Stage Face Detection Networks

Xiaofeng Liu, B.V.K Vijaya Kumar, Chao Yang, Qingming Tang, Jane You

Problem
: IPI

Strategy
- LIP(Localized Instance Perturbation)

Background technique
: Perturbation

Dataset
: WIDER FACE

[78] Deep Adaptive Attention for Joint Facial Action Unit Detection and Face Alignment

Zhiwen Shao, Zhilei Liu, Jianfei Cai, Lizhuang Ma

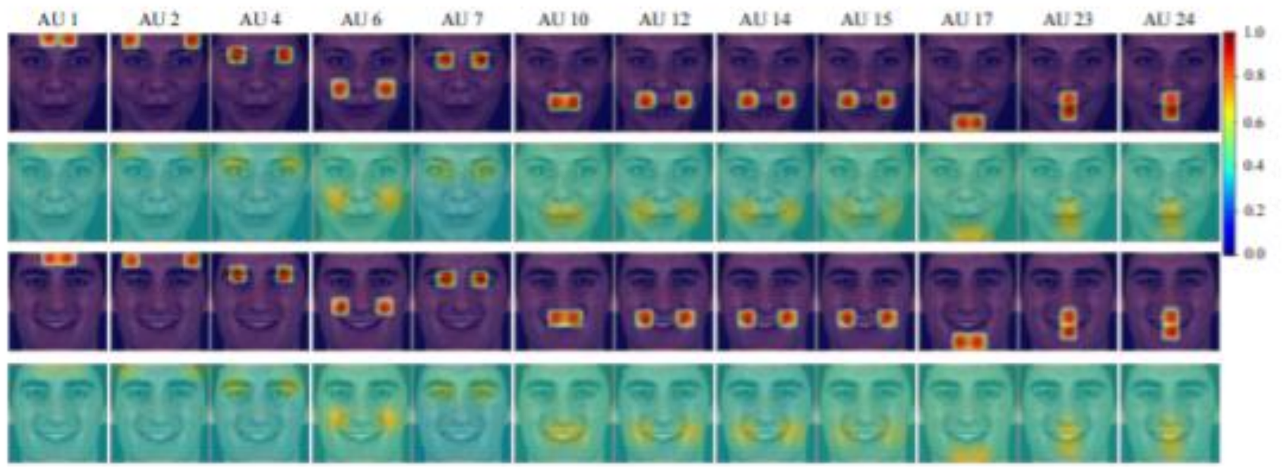


Fig. 4. Visualization of attention maps of JAA-Net. The first and third rows show the predefined attention maps, and the second and fourth rows show the refined attention maps. Attention weights are visualized with different colors as shown in the color bar

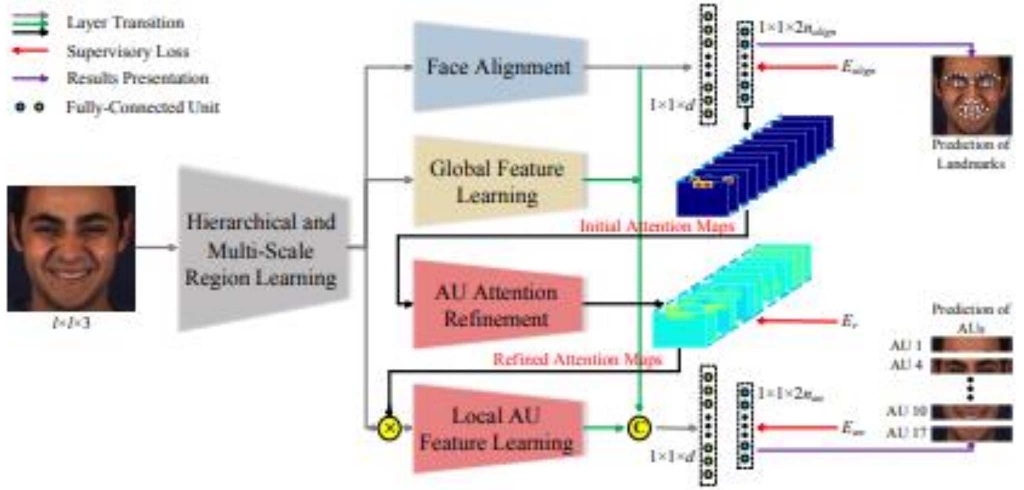


Fig. 1. The proposed JAA-Net framework, where "C" and "×" denote concatenation and element-wise multiplication, respectively

201810957 염지현

[78] Deep Adaptive Attention for Joint Facial Action Unit Detection and Face Alignment

Zhiwen Shao, Zhilei Liu, Jianfei Cai, Lizhuang Ma

Problem

: Face Detection and Face Alignment

Strategy

- End-to-end deep learning framework
  - Multi-scale shared features 학습
  - High-level features of face alignment를 AU Detection에 사용
  - Adaptive attention learning module → 정확한 지역적 특징 추론에 사용

Background technique

: CNN

Dataset

: BP4D, DISFA

[79] Dual-Agent Deep Reinforcement Learning for Deformable Face Tracking

Minghao Guo, Jiwen Lu, Jie Zhou

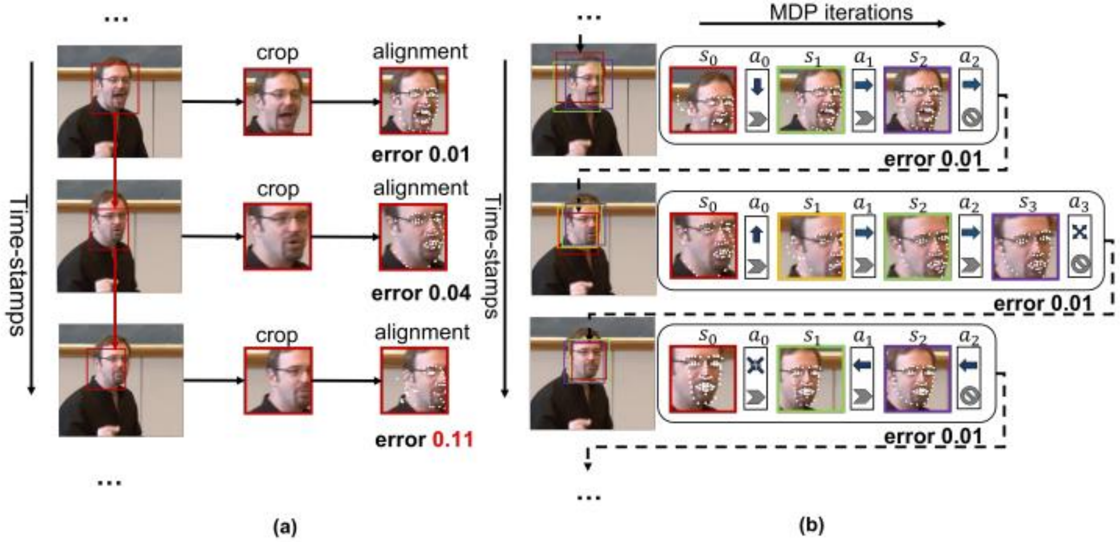


Fig. 1. (a) Existing "tracking-by-detection" methods [1–3] produce deformable face tracking in a *serial* manner. (b) Our DADRL method formulates deformable face tracking as a Markov decision process (MDP) problem, and produces bounding box tracking and landmark detection in an *interactive* manner. Here s_i denotes the MDP state, a_i denotes the MDP action. The dash line represents that initial bounding box of the current frame is the tracked box of previous frame. The blue color and the gray color denote the tracking agent action and the alignment agent action respectively (best viewed in color).

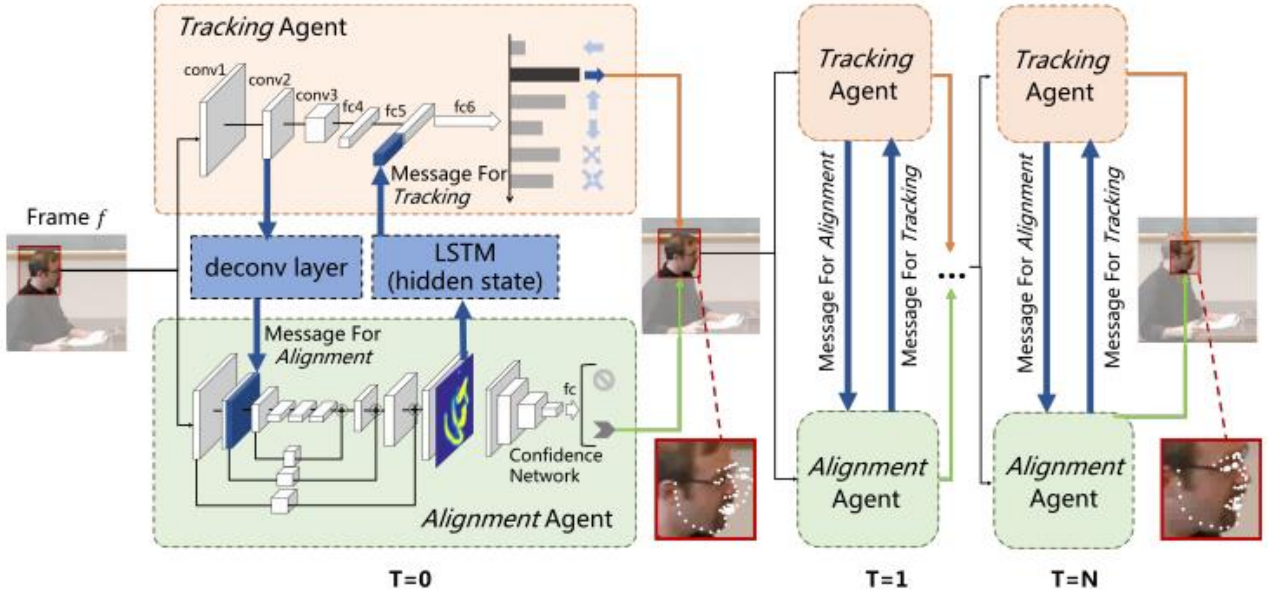


Fig. 2. The architecture of our proposed DADRL. Our DADRL consists of two agents: a tracking agent and an alignment agent. Each agent has a discrete action set. The communicated messages are encoded by a deconvolution layer and a LSTM unit, respectively. These two agents decide a sequence of actions to adjust the target face's bounding box and regress the facial landmarks simultaneously. The agents go to the next frame until the detected facial landmarks are finalized. Note that, T denotes the iteration number of MDP, rather than time-stamps number of the video.

201810957 염지현

[79] Dual-Agent Deep Reinforcement Learning for Deformable Face Tracking

Minghao Guo, Jiwen Lu, Jie Zhou

Problem
: Deformal Face Tracking

Strategy
- DADRL(Dual-agent deep RL)
  - Tracking agent
  - Alignment agent

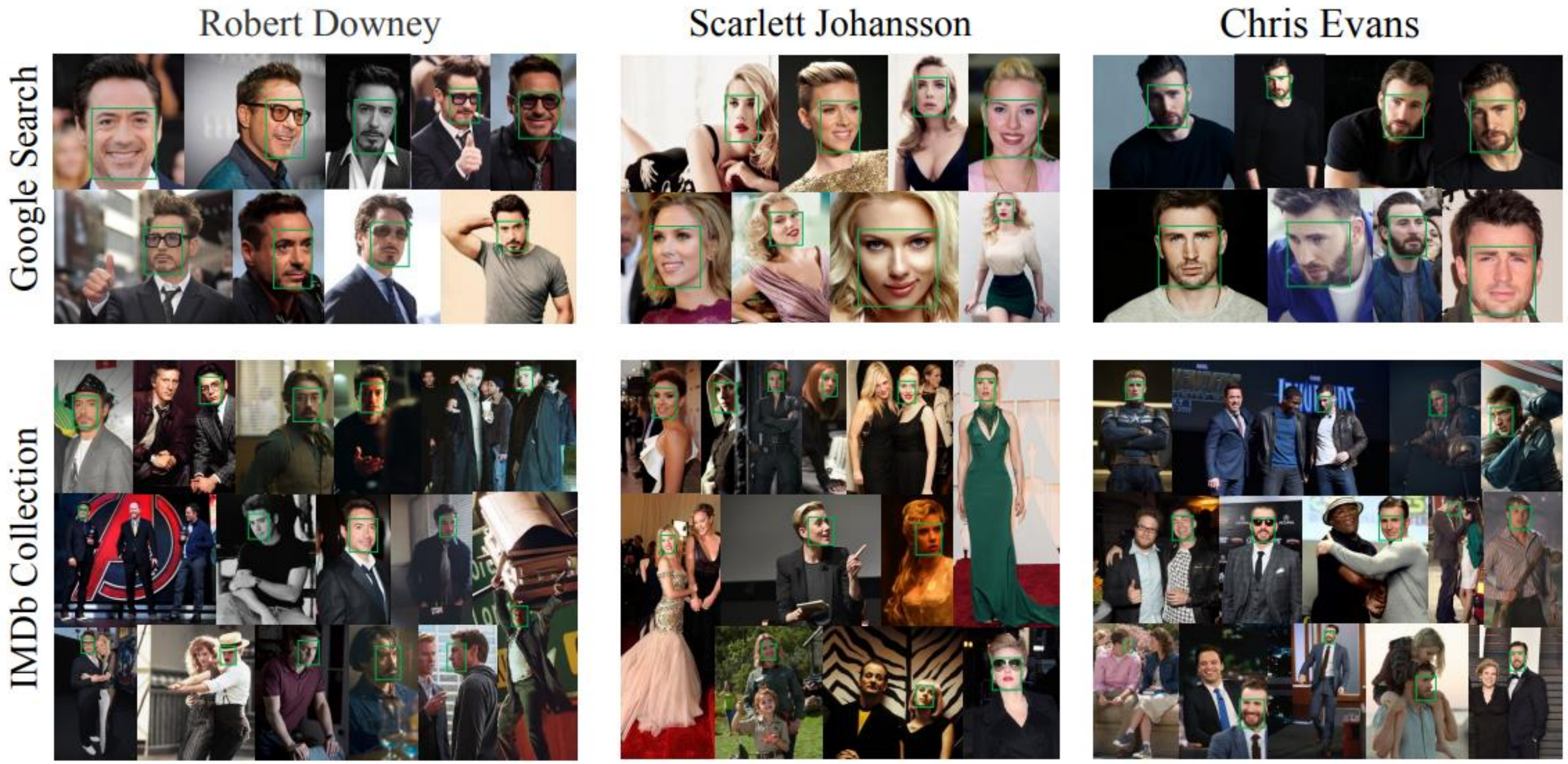
  → 두 agent가 bounding box 조정 및 face landmark regression

Background technique
: RL, CNN

Dataset
: 300-VW

[80] The Devil of Face Recognition is in the Noise

Fei Wang, Liren Chen, Cheng Li, Shiyao Huang, Yanjie Chen, Chen Qian, Chen Change Loy



[80] The Devil of Face Recognition is in the Noise

Fei Wang, Liren Chen, Cheng Li, Shiyao Huang, Yanjie Chen, Chen Qian, Chen Change Loy

Problem
: 기존의 face recognition의 dataset 규모가 커지면서 face recognition을 위한 robust한 CNN 등장하지만
label noise의 원인과 결과에 대한 이해 부족

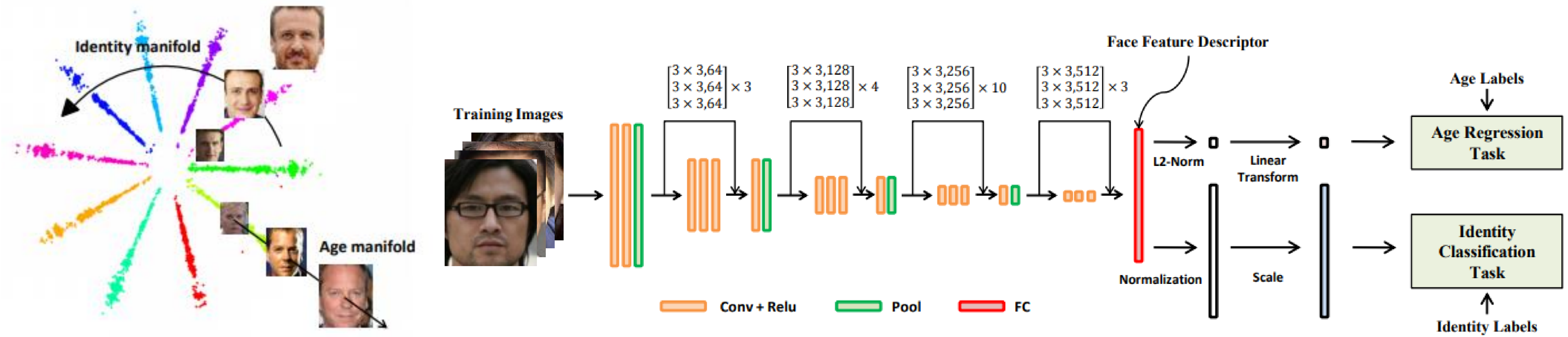
Strategy
- IMDb-Face Dataset 구축

Background technique
: CNN, dataset

Dataset
: MergeFace, MS-Celeb-1M

[81] Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition

Yitong Wang, Dihong Gong, Zheng Zhou, Xing Ji, Hao Wang, Zhifeng Li, Wei Liu, Tong Zhang



201810957 염지현

[81] Orthogonal Deep Features Decomposition for Age-Invariant Face Recognition

Yitong Wang, Dihong Gong, Zheng Zhou, Xing Ji, Hao Wang, Zhifeng Li, Wei Liu, Tong Zhang

Problem
: Age-Invariant Face Recognition

Strategy
- Orthogonal Deep Features Decomposition
    - Deep face features를 두 개의 orthogonal 구성 요소로 분해 → Age, Identity
- CAF라는 dataset 생성

Background technique
: CNN

Dataset
: CAF(Corss-Age Face), MORPH Album 2, CACD-VS, FG-NET)

[82] PyramidBox: A Context-assisted Single Shot Face Detector

Xu Tang, Daniel K. Du, Zeqiang He, Jingtuo Liu

(a) (b) (c)

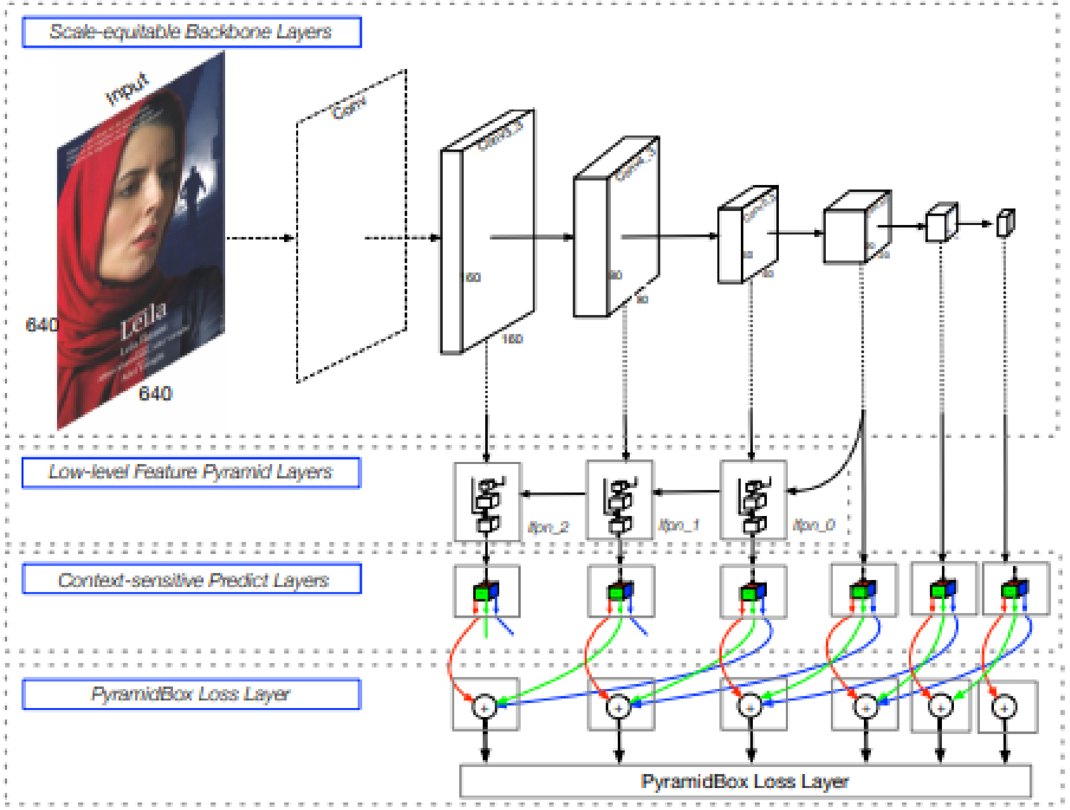


Fig. 2: Architecture of PyramidBox. It consists of **Scale-equitable Backbone Layers**, **Low-level Feature Pyramid Layers (LFPN)**, **Context-sensitive Predict Layers** and **PyramidBox Loss Layer**.

[82] PyramidBox: A Context-assisted Single Shot Face Detector

Xu Tang, Daniel K. Du, Zeqiang He, Jingtuo Liu

Problem

: Face Detection
: Face detection의 문제 중 small, blurred, partially occluded faces in uncontrolled environment

Strategy

- PyramidAnchors
  - Semi-supervised 방법
- PyramideBox
  - Low-level Facial Feature + adequate high-level context semantic feature 결합 → 한 번에 모든 scale의 얼굴 예측 가능
- Data-anchor=sampling
  - 훈련 샘플 확대 → 훈련 데이터의 다양성 증가

Background technique

: CNN

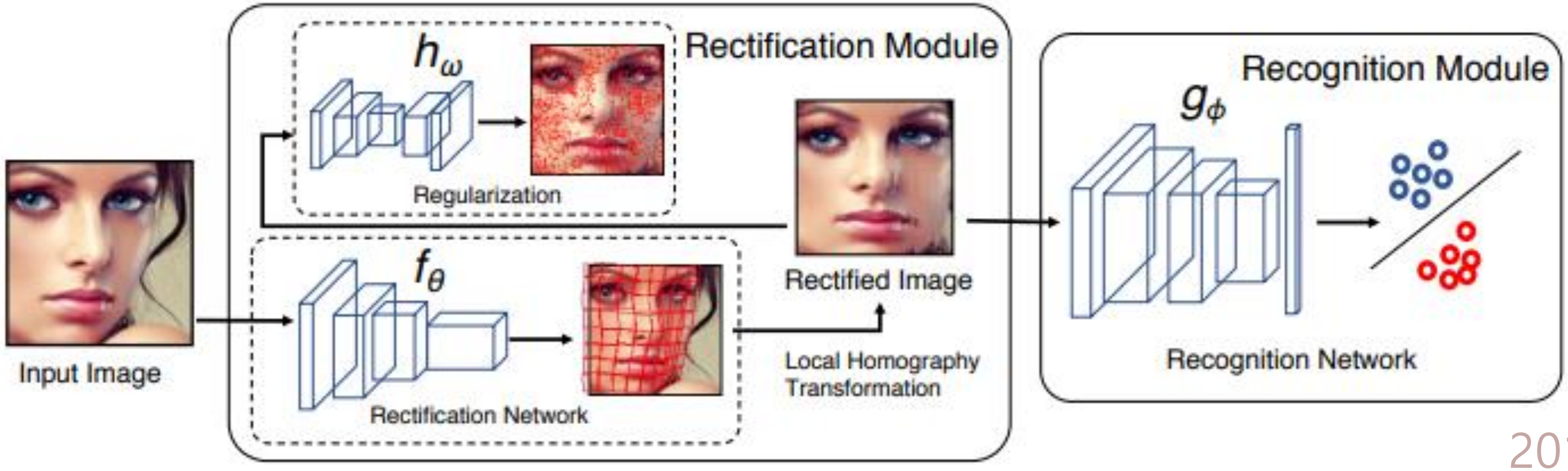
Dataset

: WIDER FACE

201810957 염지현

[83] GridFace: Face Rectification via Learning Local Homography Transformations

Erjin Zhou, Zhimin Cao, and Jian Sun



[83] GridFace: Face Rectification via Learning Local Homography Transformations

Erjin Zhou, Zhimin Cao, and Jian Sun

Problem

: Face Recognition

: Face geometric variation을 줄이고 recognition의 성능 향상이 목표

Strategy

- GridFace

Background technique

: GAN

Dataset

: Multi-PIE, LFW, YTF

201810957 염지현

[84] ELEGANT: Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes

Taihong Xiao, Jiapeng Hong, and Jinwen Ma

(a) removing bangs

(b) adding bangs

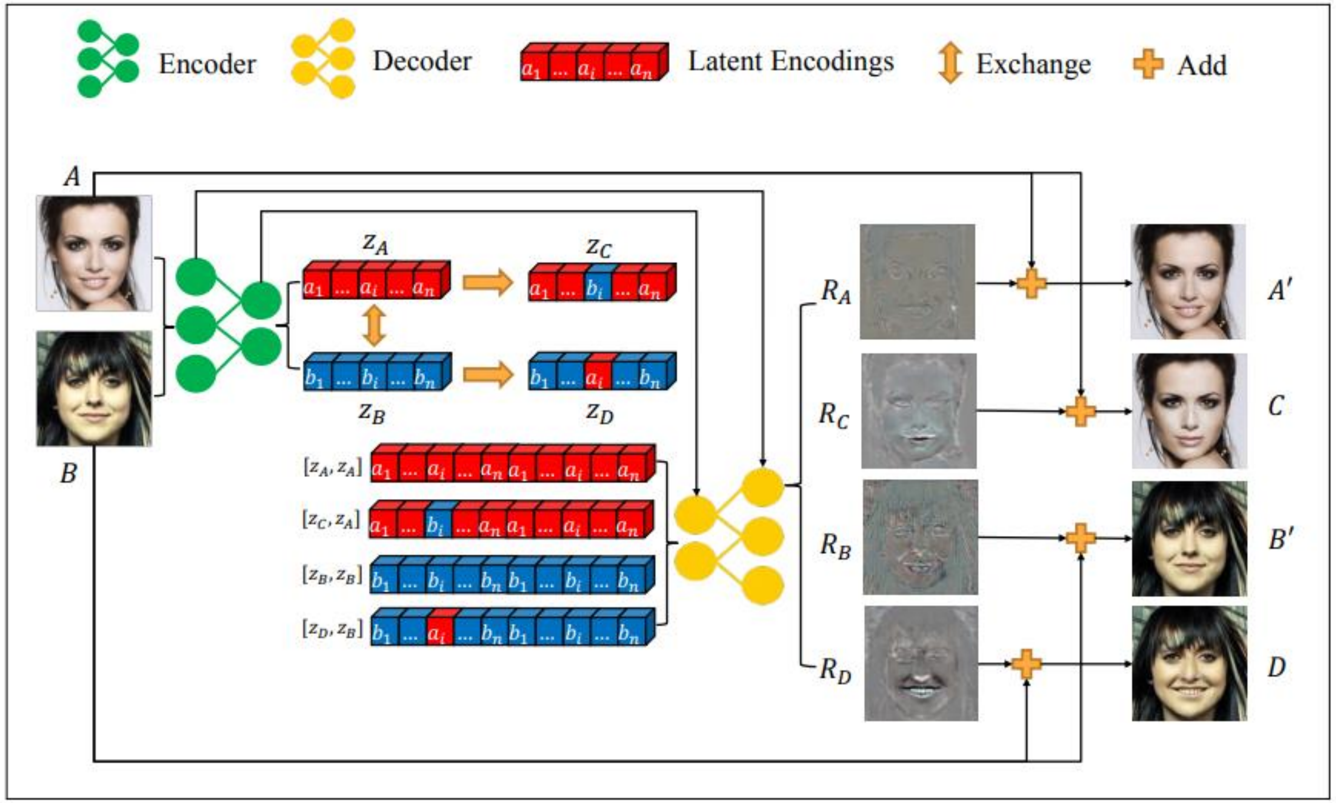


Fig. 6: The ELEGANT model architecture.

[84] ELEGANT: Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes

Taihong Xiao, Jiapeng Hong, and Jinwen Ma

Problem

: Face Attribute Transfer

Strategy

- ELEGANT
  - 입력 : 2장의 이미지
  - 출력 : 서로의 attribute transfer 한 이미지 출력

Background technique

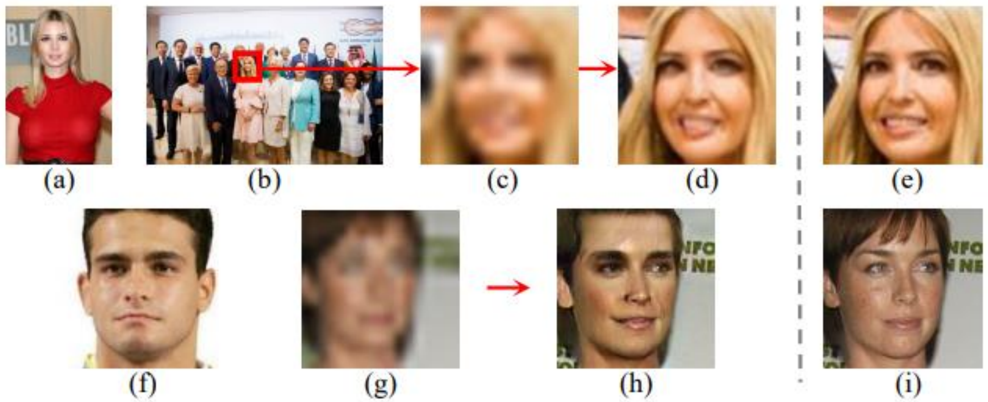
: GAN

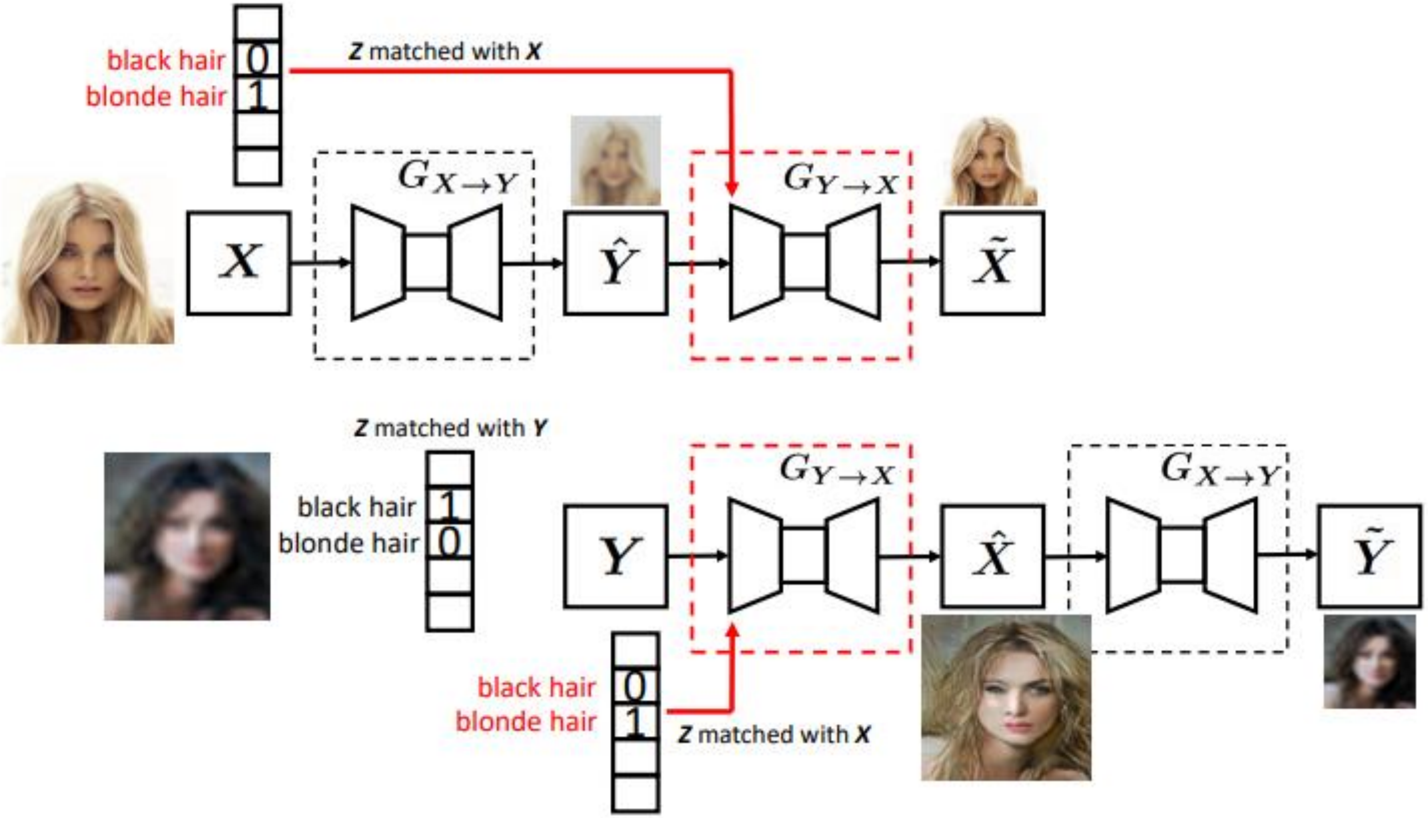
Dataset

: Celeb-A

[85] Attribute-Guided Face Generation Using Conditional CycleGAN

Yongyi Lu 1 , Yu-Wing Tai 2 , and Chi-Keung Tang





201810957 염지현

[85] Attribute-Guided Face Generation Using Conditional CycleGAN

Yongyi Lu 1 , Yu-Wing Tai 2 , and Chi-Keung Tang

Problem
: Face Resolution

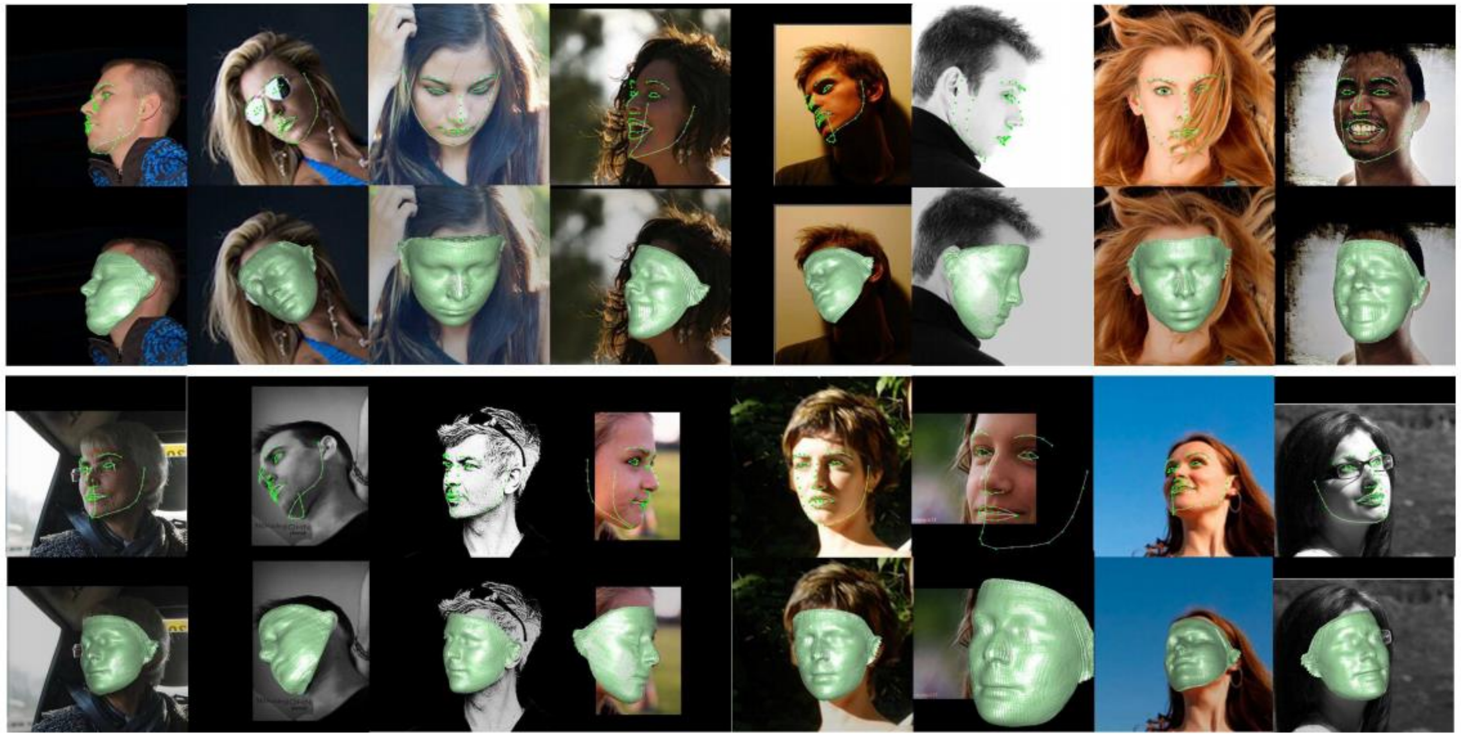
Strategy
- Conditional CycleGAN
  - Non-paired image더라도 resolution 가능
  - Input image의 attribute를 통해 identity preserving

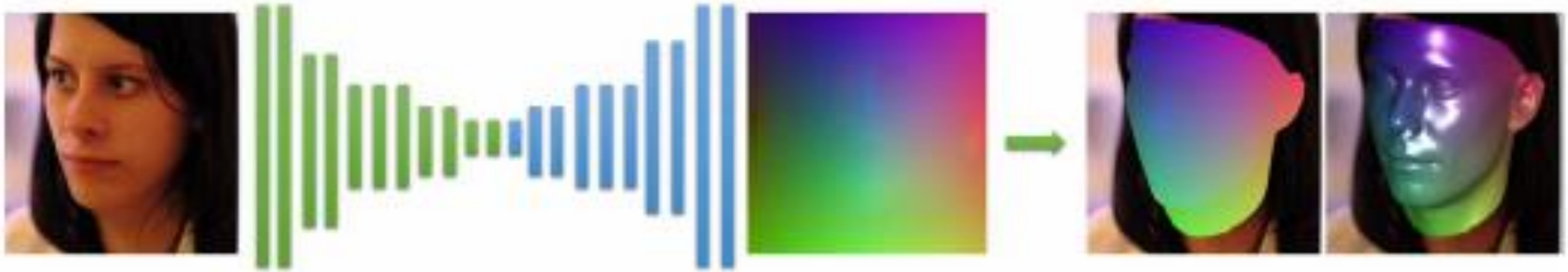
Background technique
: GAN

Dataset
: Celeb-A, MNIST

[86] Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network

Yongyi Lu 1 , Yu-Wing Tai 2 , and Chi-Keung Tang





201810957 염지현

[86] Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network

Yongyi Lu 1 , Yu-Wing Tai 2 , and Chi-Keung Tang

Problem
: 3D Face Reconstruction

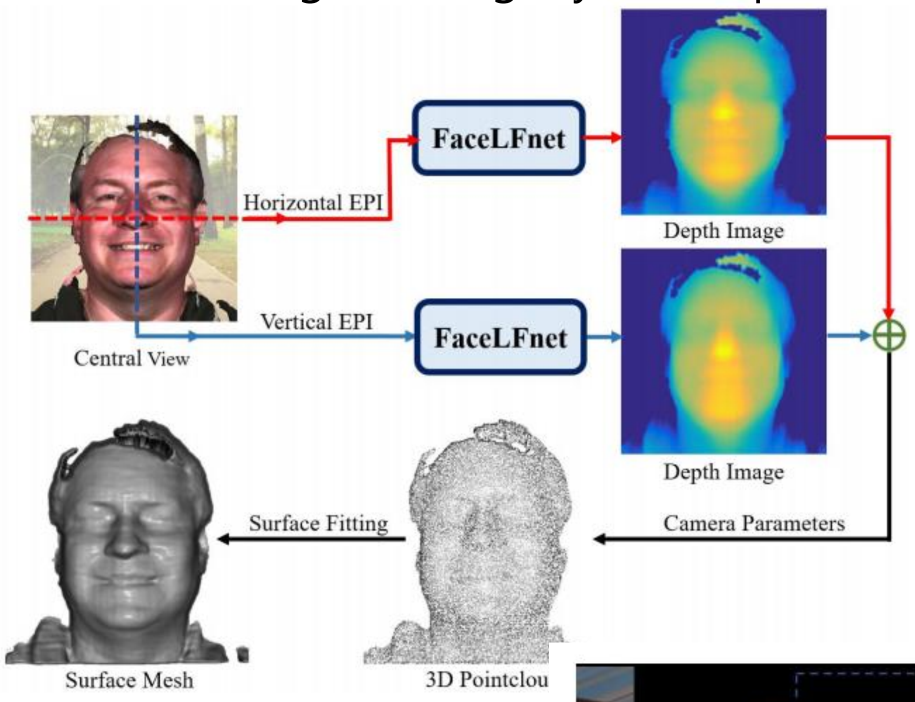
Strategy
- Position Map Regression Network
  - 3D face reconstruction 할 때 UV map이라는 2D image로부터 regression

Background technique
: GAN

Dataset
: 300W-LP

[87] 3D Face Reconstruction from Light Field Images: A Model-free Approach

Mingtao Feng, Syed Zulqarnain Gilani, Yaonan Wang, Ajmal Mian



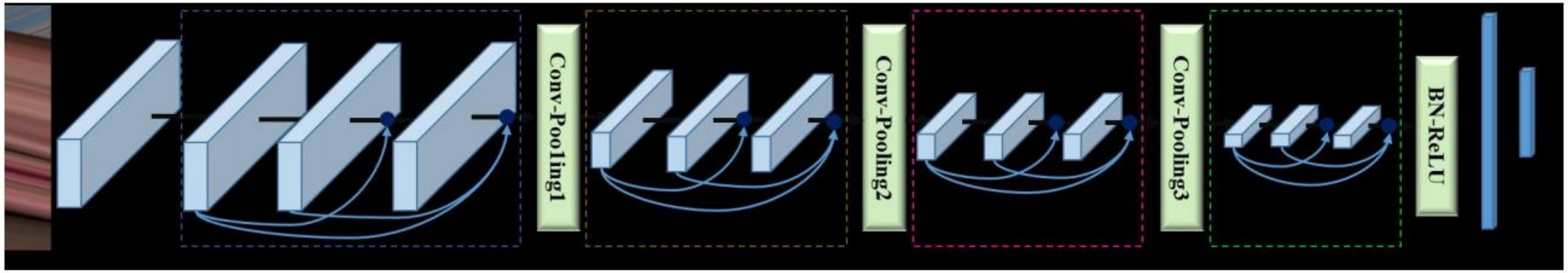


Figure 5. Our proposed FaceLFnet for learning 3D face curves from EPIs. It contains 4 dense blocks, followed by two fully connected layers. The layers between two neighboring blocks are defined as transition layers and change feature map sizes via convolution and pooling [23].

[87] 3D Face Reconstruction from Light Field Images: A Model-free Approach

Mingtao Feng, Syed Zulqarnain Gilani, Yaonan Wang, Ajmal Mian

Problem
: 3D Face Reconstruction

Strategy
- FaceLFnet
  - EPI를 활용 → EPI에서 수평 및 수직 3D Face 곡선 복구

Background technique
: CNN

Dataset
: BU-3DFE, BU-4DFE

201810957 염지현

[88] Remote Photoplethysmography Correspondence Feature for 3D Mask Face Presentation Attack Detection

Si-Qi Liu, Xiangyuan Lan, Pong C. Yuen

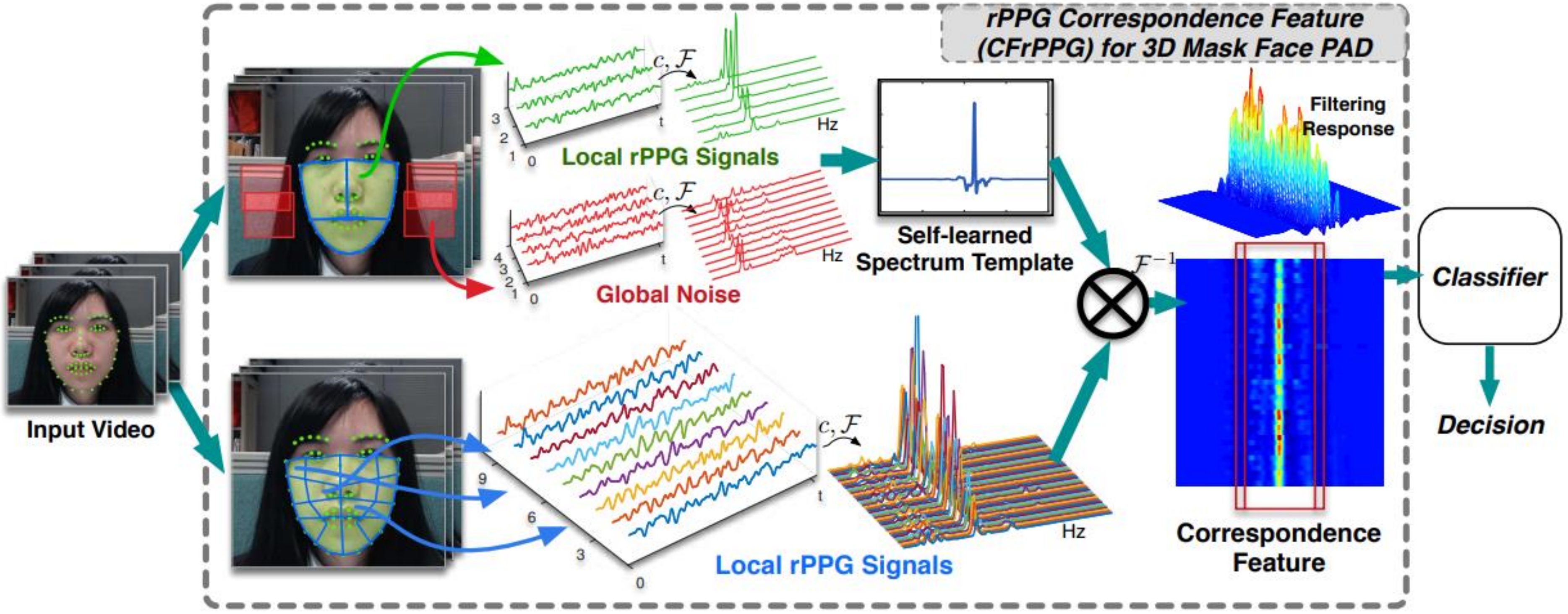


Fig. 1. Block diagram of the proposed CFrPPG feature

201810957 염지현

[88] Remote Photoplethysmography Correspondence Feature for 3D Mask Face Presentation Attack Detection

Si-Qi Liu, Xiangyuan Lan, Pong C. Yuen

Problem

: 3D Face Reconstruction

Strategy

- FaceLFnet
    - EPI를 활용 → EPI에서 수평 및 수직 3D Face 곡선 복구

Background technique

: CNN

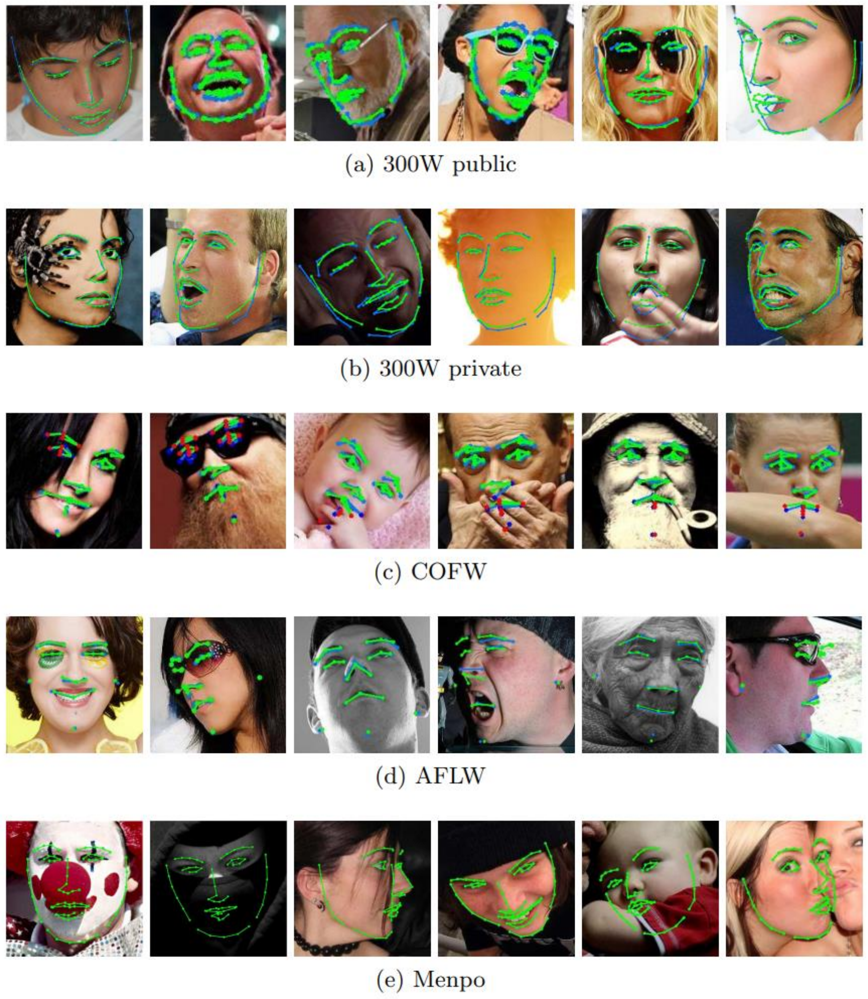
Dataset

: BU-3DFE, BU-4DFE

201810957 염지현

[89] A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment

Roberto Valle, Jose M. Buenaposada, Antonio Valdes, Luis Baumela



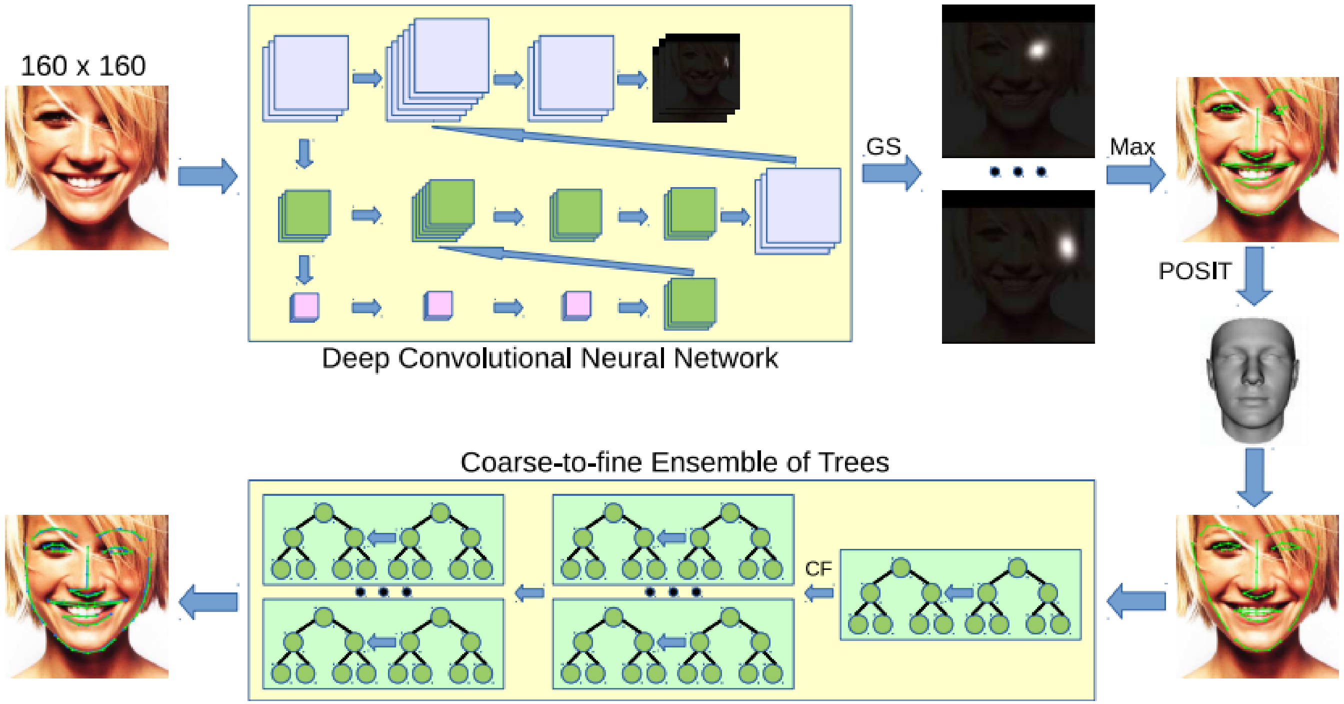


Fig. 1: DCFE framework diagram. GS, Max and POSIT represent the Gaussian smoothing filter, the maximum of each probability map and the 3D pose estimation respectively.

201810957 염지현

[89] A Deeply-initialized Coarse-to-fine Ensemble of Regression Trees for Face Alignment

Roberto Valle, Jose M. Buenaposada, Antonio Valdes, Luis Baumela

Problem
: Face Alignment

Strategy
- DCFE
  - ERT(Ensemble of Regression Tress)를 기반으로 실시간 face landmark regression
  - Landmark 위치의 확률 map 생성
  - 폐색이나 여러 명의 프로필 얼굴 분석 가능

Background technique
: CNN

Dataset
: AFLW, COFW, 300W

201810957 염지현

Head

CVPR

[1] Natural and Effective Obfuscation by Head Inpainting

Qianru Sun, Liqian Ma, Seong Joon Oh, Luc Van Gool, Bernt Schiele, Mario Fritz

Figure 1: Our obfuscation method based on head inpainting generates much more natural patterns than common techniques like blurring, but still results in a more effective identity obfuscation against a recognizer.

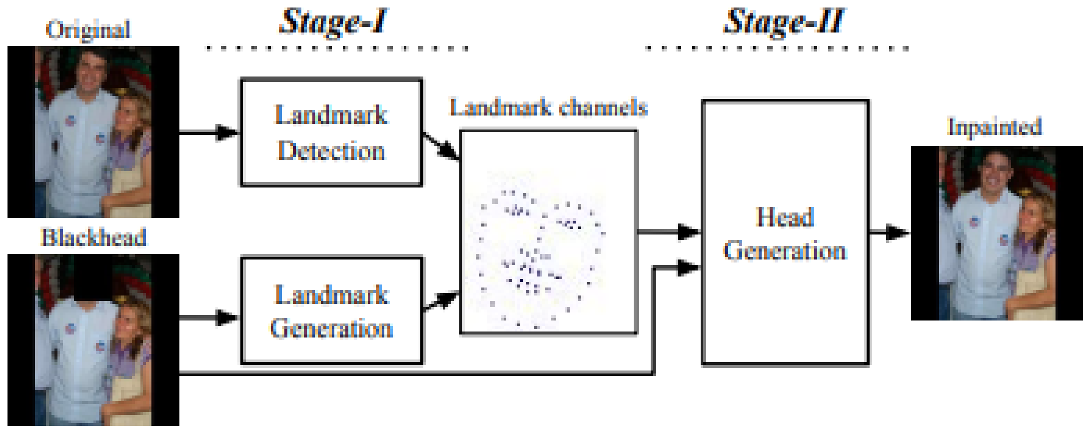


Figure 2: Our two-stage head inpainting framework. The input of stage-I is either the original or the blackhead image. The output is the inpainted image.

[1] Natural and Effective Obfuscation by Head Inpainting

Qianru Sun, Liqian Ma, Seong Joon Oh, Luc Van Gool, Bernt Schiele, Mario Fritz

Problem

: Identity Obfuscation → blurred images도 identity가 인식될 수 있음

Strategy

- Head inpaint obfuscation technique
    - Image context에서 face landmark 생성
    - Facial landmark conditional head inpainting

Background technique

: GAN

Dataset

: PIPA

[2] MX-LSTM: Mixing Tracklets and Vislets to Jointly Forecast Trajectories and Head Poses

Irtiza Hasan, Francesco Setti, Theodore Tsesmelis, Alessio Del Bue, Fabio Galasso, Marco Cristani

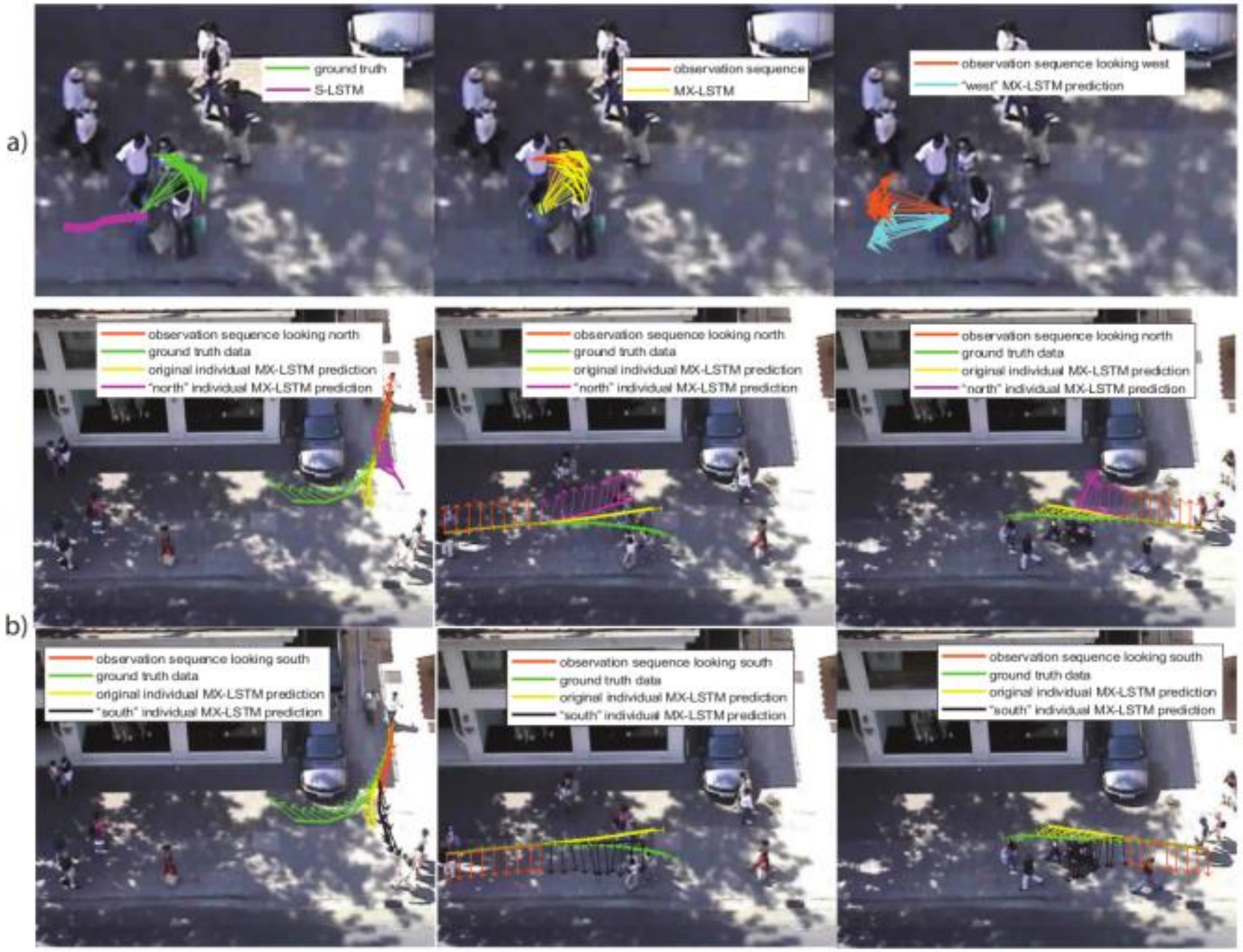


Figure 3. Qualitative results: a) MX-LSTM b) Ablation qualitative study on Individual MX-LSTM (better in color).

201810957 염지현

[2] MX-LSTM: Mixing Tracklets and Vislets to Jointly Forecast Trajectories and Head Poses

Irtiza Hasan, Francesco Setti, Theodore Tsesmelis, Alessio Del Bue, Fabio Galasso, Marco Cristani

Problem
: Predict the future positions of pedestrians

Strategy
- MX-LSTM = tracklet + vislet ➔ 궤적 예측

Background technique
: LSTM

Dataset
: Zara01, Zara02, UCY, TownCentre

Portrait

CVPR

201810957 염지현

[1] Facelet-Bank for Fast Portrait Manipulation

Ying-Cong Chen, Huaijia Lin, Michelle Shu, Ruiyu Li, Xin Tao, Xiaoyong Shen, Yangang Ye, Jiaya Jia

(a) Original Image (b) Get younger (c) Add facial hair

Figure 1. Illustration of face manipulation using our model.

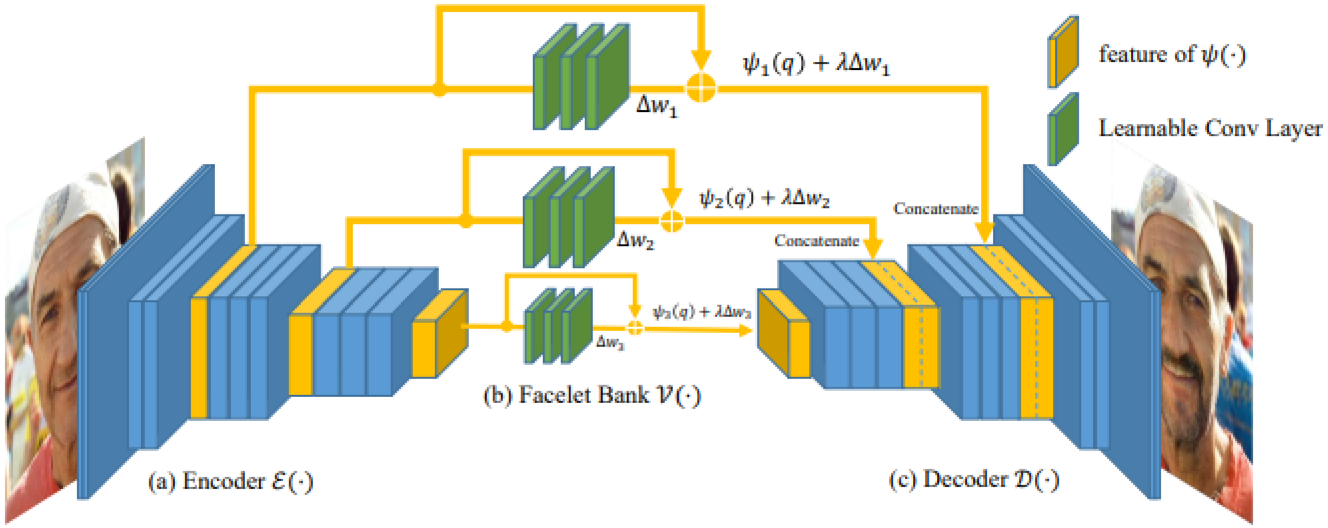


Figure 2. Illustration of our framework. (a) is the encoder $\mathcal{E}(\cdot)$; (b) are convolutional layers of our facelet bank $\mathcal{V}(\cdot)$; (c) is the decoder $\mathcal{D}(\cdot)$. The structure of our facelet bank is Conv-ReLU-Conv-ReLU-Conv, where all Convs are with 3×3 kernels. Also, all Convs of the facelet bank do not change the height, width and number of channels given the previous input.

[1] Facelet-Bank for Fast Portrait Manipulation

Ying-Cong Chen, Huaijia Lin, Michelle Shu, Ruiyu Li, Xin Tao, Xiaoyong Shen, Yangang Ye, Jiaya Jia

Problem
: Portrait Manipulation

Strategy
- End-to-end convolutional network – unpaired image set
  → 다양한 style 적용 가능

Background technique
: CNN

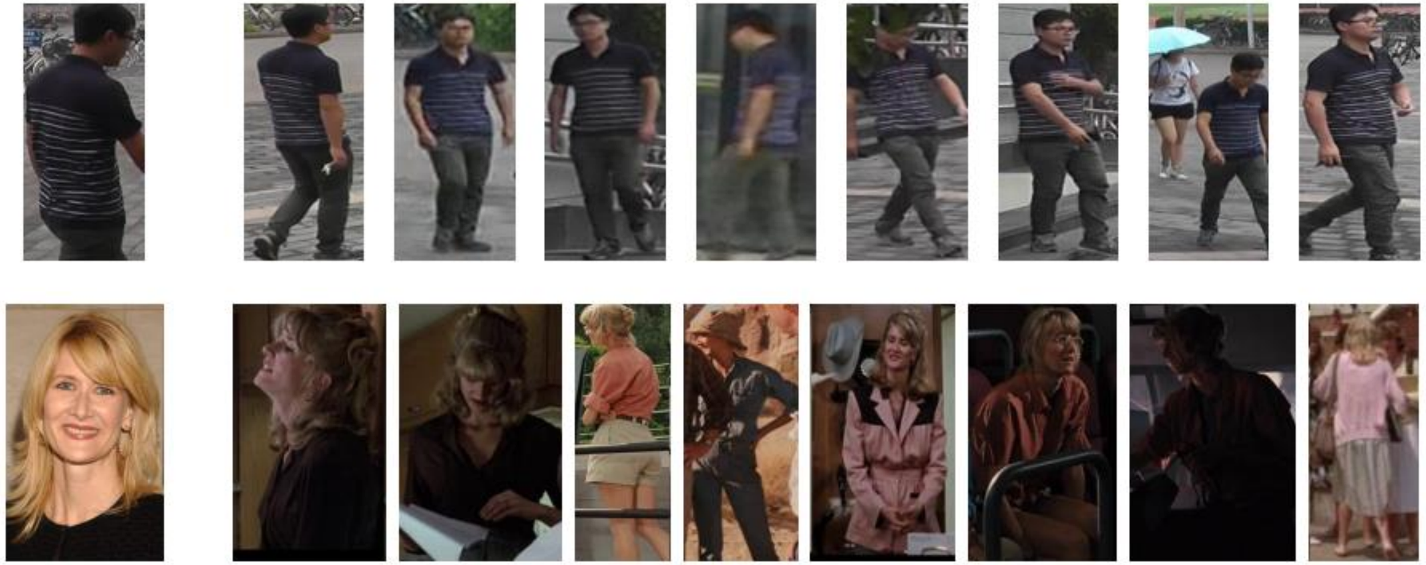
Dataset
: CelebA

Portrait

ECCV

[2] Person Search in Videos with One Portrait Through Visual and Temporal Links

Qingqiu Huang, Wentao Liu, Dahua Lin



201810957 염지현

[2] Person Search in Videos with One Portrait Through Visual and Temporal Links

Qingqiu Huang, Wentao Liu, Dahua Lin

Problem
: Person Search

Strategy
- Tracklet : Identity invariant 고려
- Visual and the temporal link : Person Identity 전파

Background technique
: CNN

Dataset
: CSM, MARS

Style
&
Stylization

CVPR

201810957 염지현

[1] PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup

Huiwen Chang, Jingwan Lu, Fisher Yu, Adam Finkelstein

Figure 1: Three source photos (top row) are each modified to match makeup styles in three reference photos (left column) to produce nine different outputs (3×3 lower right).

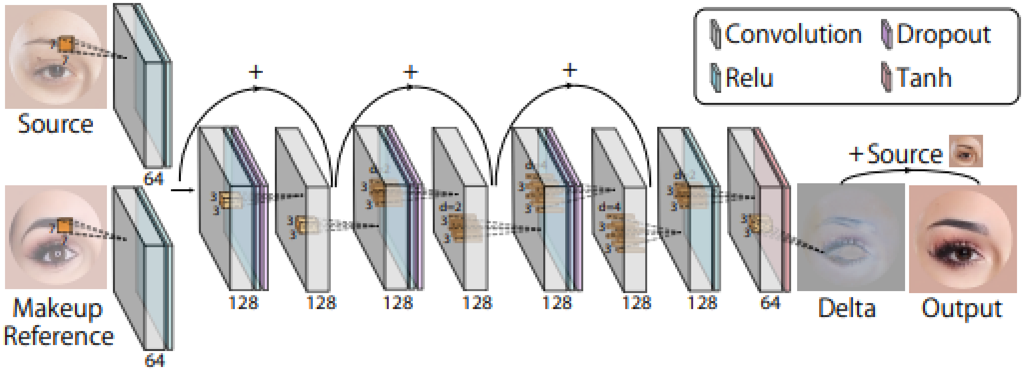


Figure 5: Generator Architecture. For generator G, we use DRN [25] with 3 dilated residual blocks to retain small spatial information (such as eye makeup). d indicates the dilation factor. We use two degridding layers (a 2-dilated 3×3 convolution and a 3×3 convolution) at the end of the network to avoid grid artifacts. The architecture of F is similar. The only difference is that it takes one image as input and therefore does not need concatenation.

201810957 염지현

[1] PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup

Huiwen Chang, Jingwan Lu, Fisher Yu, Adam Finkelstein

Problem
: Applying and Removing Makeup

Strategy
- PairedCycleGan
  - Network1 : 기존 face에 style을 입히는 network
  - Network2 : style을 입힌 face에서 다시 remove style → cycle Consistencyloss

Background technique
: GAN

Dataset
: Yutube makeup tutorial video

[2] Disentangling Structure and Aesthetics for Style-Aware Image Completion

Andrew Gilbert, John Collomosse, Hailin Jin, Brian Price

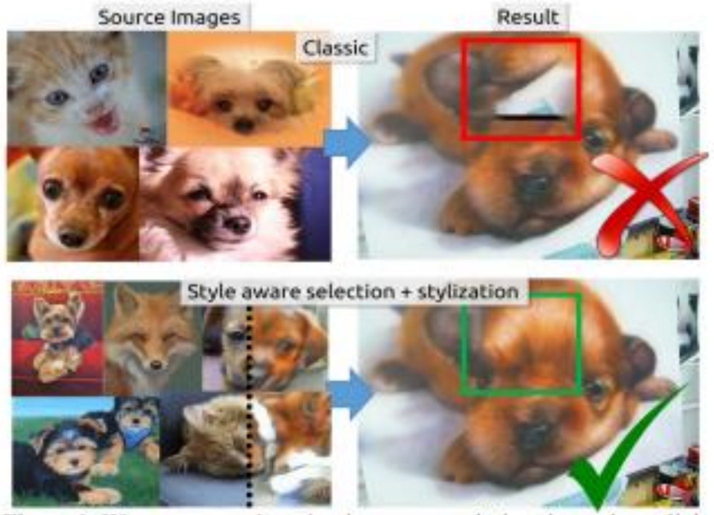


Figure 1. We propose enhancing image completion through explicit consideration of visual aesthetics (style) alongside structure and semantics. A deep convnet is used to disentangle patch structure and style, driving 1) style-aware patch search; 2) a style-aware optimization for patch selection; 3) stylization of patch content to enable seamless image completion with coherent visual aesthetic.

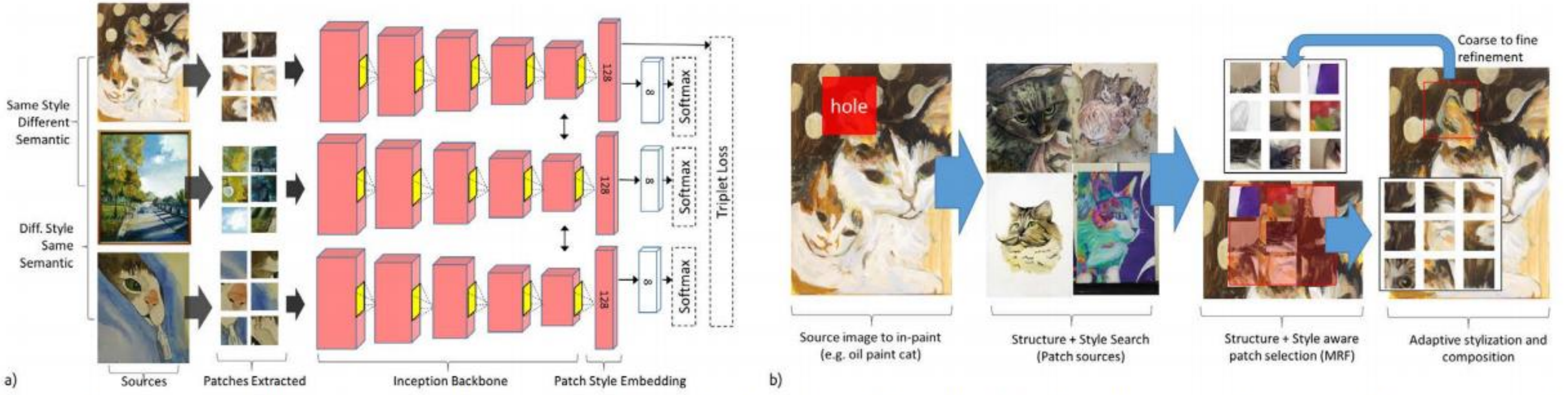


Figure 2. Overview of the proposed method. (a) Learning a local patch model to disentangle aesthetics and structure. (b) our proposed 3-stage algorithm for style aware image in-painting: 1) source image used as a query for a style-aware web-scale search for in-filling candidates; 2) patch aggregation and selection via MRF balancing structure and style; 3) adaptive stylization of patches using MRF weights.

[2] Disentangling Structure and Aesthetics for Style-Aware Image Completion

Andrew Gilbert, John Collomosse, Hailin Jin, Brian Price

Problem
: Content-aware image completion and in-painting

Strategy
- 구조 & style의 일관성을 모두 적용한 non-parametric in-painting algorithm
  - Patch search and selection : image 구조와 style 분리
  - Selected patches를 target 이미지에 맞추기 위해 adaptive stylization of patch 수행

Background technique
: GAN

Dataset
: Places2, BAM

[3] Real-Time Monocular Depth Estimation Using Synthetic Data With Domain Adaptation via Image Style Transfer

Amir Atapour-Abarghouei, Toby P. Breckon

Figure 1: Our monocular depth estimation (KITTI [55]).

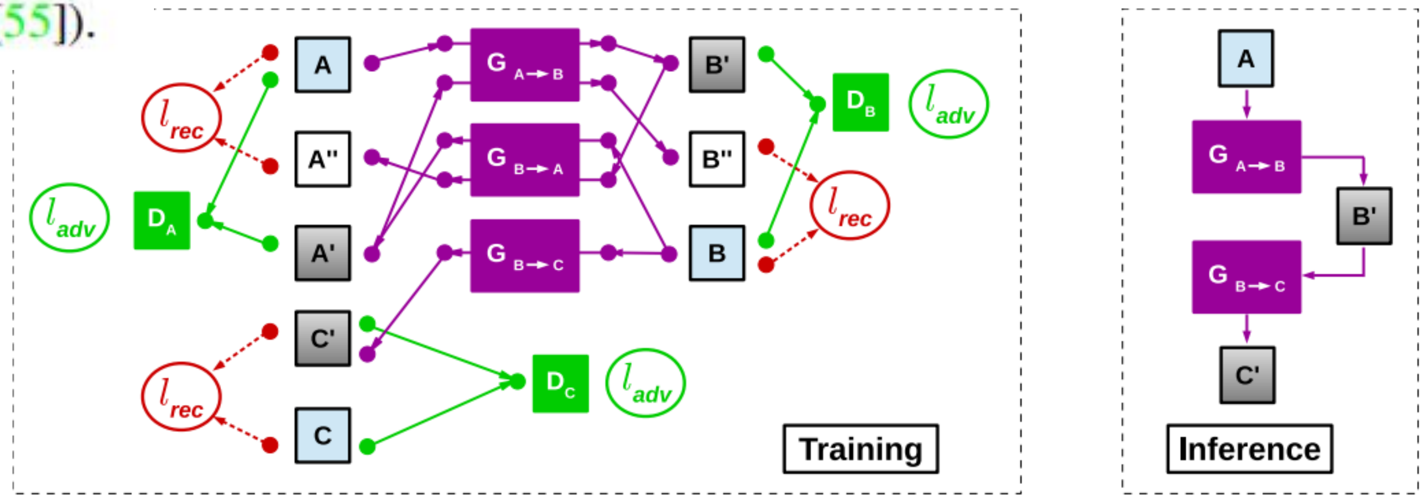


Figure 2: Our approach using [90]. Domain A (real-world RGB) is transformed into B (synthetic RGB) and then to C (pixel-perfect depth). A, B, C denote ground truth, A', B', C' generated images, and A'', B'' cyclically regenerated images.

201810957 염지현

[3] Real-Time Monocular Depth Estimation Using Synthetic Data With Domain Adaptation via Image Style Transfer

Amir Atapour-Abarghouei, Toby P. Breckon

Problem

: Real-Time Monocular Depth Estimation(rely on large quantities of ground truth)

Strategy

- Image style transfer와 domain adaptation 연결
- Style transfer + adversarial training
  → monocular 실제 color 이미지에서 Depth Estimation 측정

Background technique

: GAN

Dataset

: KITTI

[4] Camera Style Adaptation for Person Re-Identification

Zhun Zhong, Liang Zheng, Zhedong Zheng, Shaozi Li, Yi Yang

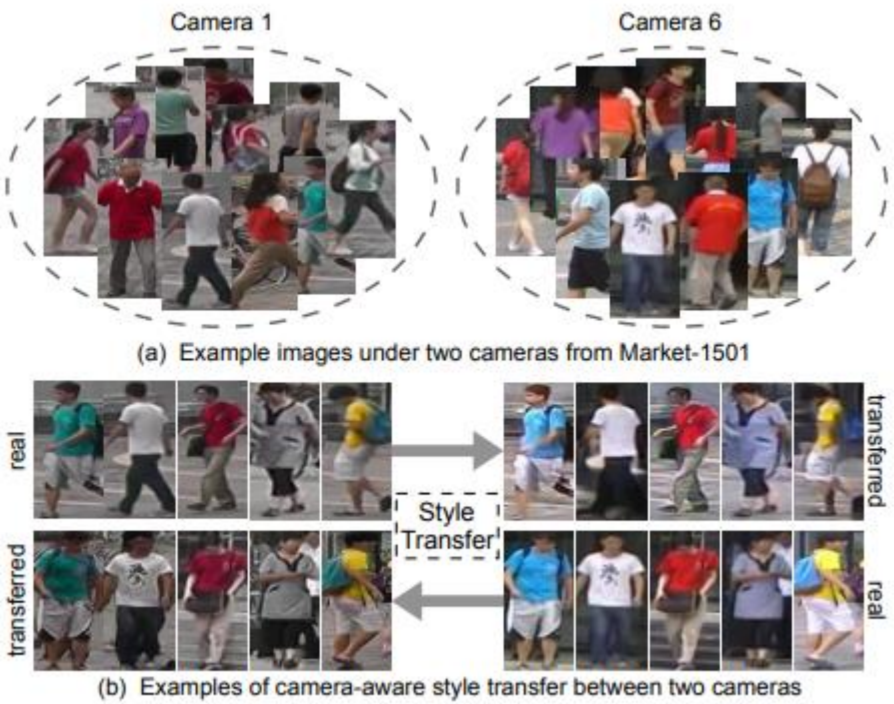


Figure 1. (a) Example images from Market-1501 [42]. (b) Examples of camera-aware style transfer between two cameras using our method. Images in the same column represent the same person.

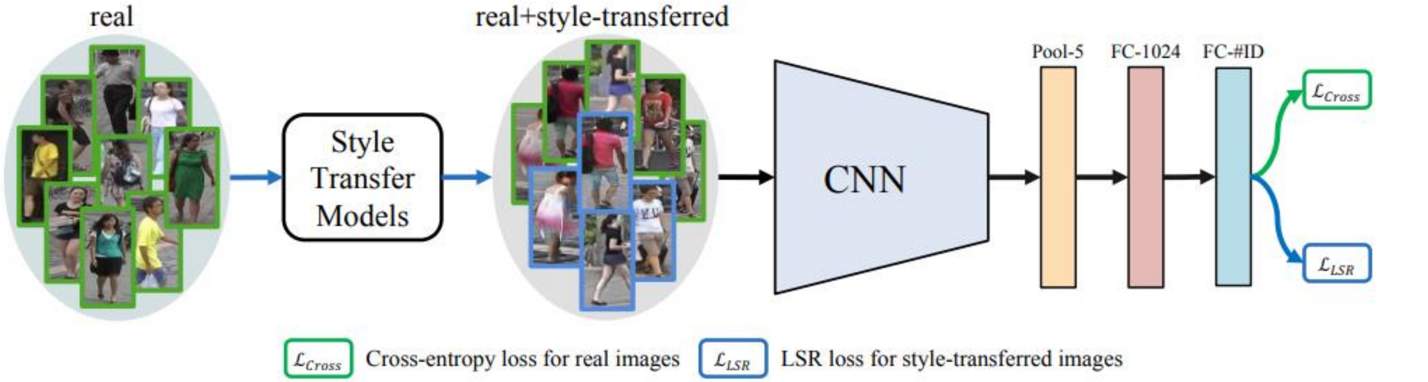


Figure 3. The pipeline of our method. The camera-aware style transfer models are learned from the real training data between different cameras. For each real image, we can utilize the trained transfer model to generate images which fit the style of target cameras. Subsequently, real images (green boxes) and style-transferred images (blue boxes) are combined to train the re-ID CNN. The cross-entropy loss and the label smooth regularization (LSR) loss are applied to real images and style-transferred images, respectively.

[4] Camera Style Adaptation for Person Re-Identification

Zhun Zhong, Liang Zheng, Zhedong Zheng, Shaozi Li, Yi Yang

Problem

: Person Re-identification

: Person Re-identification 할 대 다른 카메라로 촬영해서 같은 사람임에도 불구하고 다른 사람으로 인식하는
 경우 발생

Strategy

- Camera Style Adaptation
  - CycleGAN : Data 다양성 증가but 노이즈 발생
  - LSR(Label Smooth Regularization) : 노이즈 완화

Background technique

: CNN, GAN

Dataset

: Market-4501, DuckMTMC-reID

201810957 염지현

[5] Stereoscopic Neural Style Transfer

Dongdong Chen, Lu Yuan, Jing Liao, Nenghai Yu, Gang H

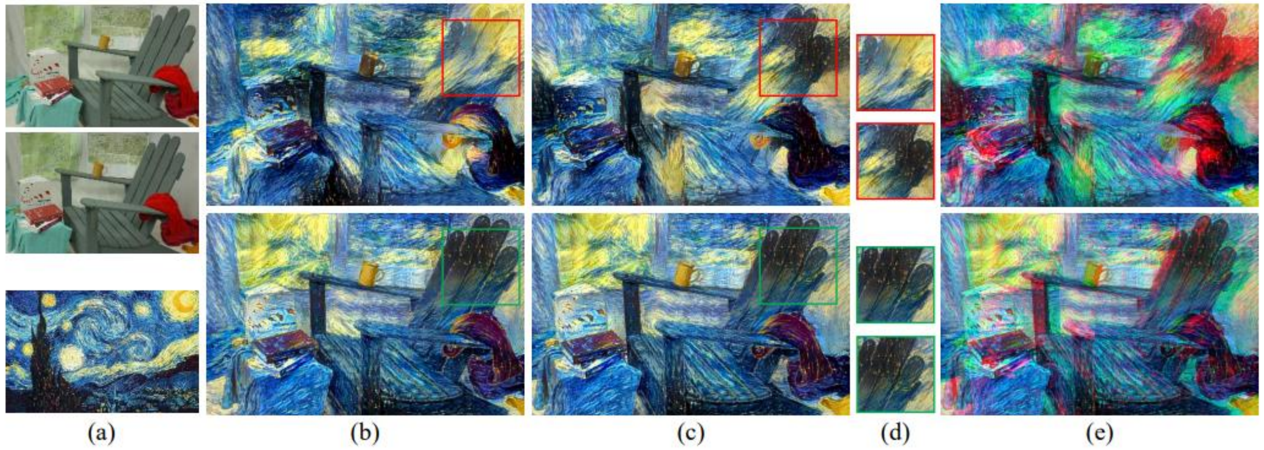


Figure 1. (a) Given a stereoscopic image pair and a style image, when the left and right view are stylized separately (first row), the left stylization result (b) will be inconsistent with that of the right view (c) in spatially corresponding areas (d).This will lead to undesirable vertical disparities and incorrect horizontal disparities, subsequently causing 3D visual fatigue in anaglyph images (e) . In contrast, by introducing a new disparity consistency constraint, our method (second row) can produce consistent stylization results for the two views.

<STEREOSCOPIC
>

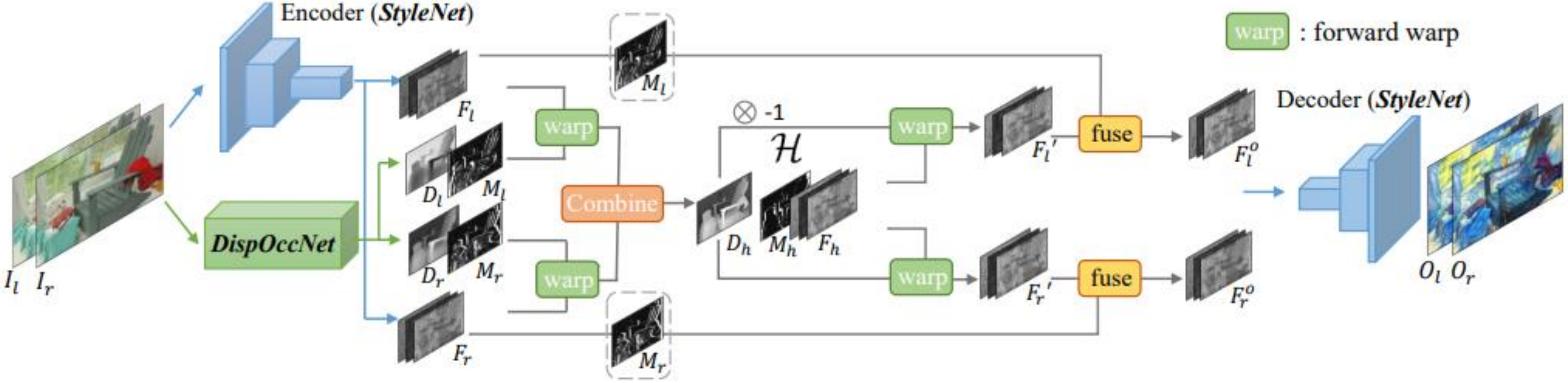


Figure 3. The overall network structure for fast stereoscopic image style transfer. It consists of two sub-networks: *StyleNet* and *DispOccNet*, which are integrated in the feature level middle domain $\mathcal{H}$.

[5] Stereoscopic Neural Style Transfer

Dongdong Chen, Lu Yuan, Jing Liao, Nenghai Yu, Gang Hua

Problem

: Stereoscopic Neural Style Transfer
: 3D 영화 또는 AR/VR에 대한 새로운 수요에 대응하는 stereoscopic neural style transfer
: 3D 영상을 통한 눈의 피로감

Strategy

- Non-occluded regions에서 bidirectional disparity constraint 시행
  → new disparity loss + adapted style loss
- Stylization sub-network & disparity sub-network

Background technique

: GAN

Dataset

: Microsoft COCO

[6] Multi-Content GAN for Few-Shot Font Style Transfer

Samaneh Azadi, Matthew Fisher, Vladimir G. Kim, Zhaowen Wang, Eli Shechtman, Trevor Darrell

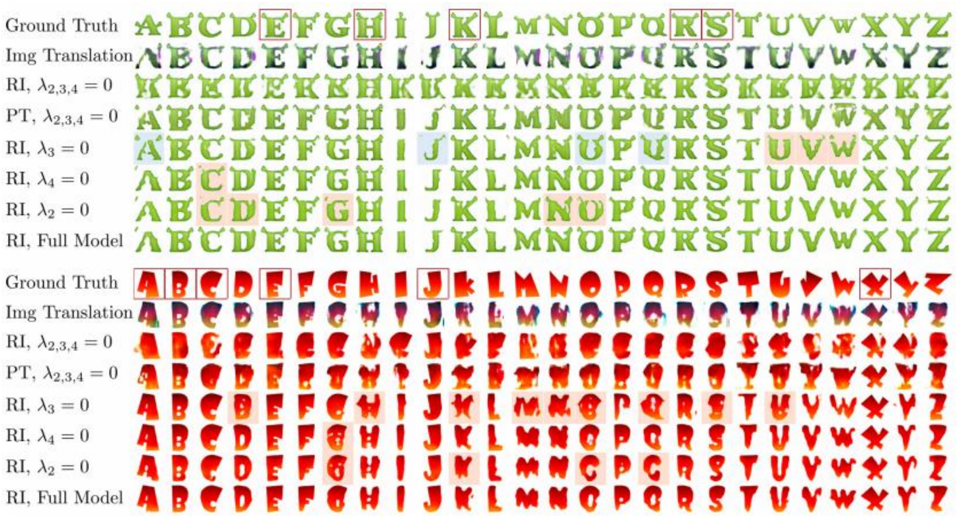


Figure 5: Ablation study on our MC-GAN model components: For each exemplar font, we show ground truth (**1st row**), observed letters (red squares in the 1st row), predictions of a baseline image translation network (**2nd row**), predictions of our end-to-end model with randomly initialized (RI) OrnaNet and $\lambda_2 = \lambda_3 = \lambda_4 = 0$ (**3rd row**), with pre-trained (PT) OrnaNet weights and $\lambda_2 = \lambda_3 = \lambda_4 = 0$ (**4th row**), selectively disabled loss terms (**rows 5-7**), and the full end-to-end MC-GAN model (**bottom row**). Style transfer improvements by λ_3 are highlighted in blue and degradation in the predictions by omitting each individual regularizer is highlighted in red.

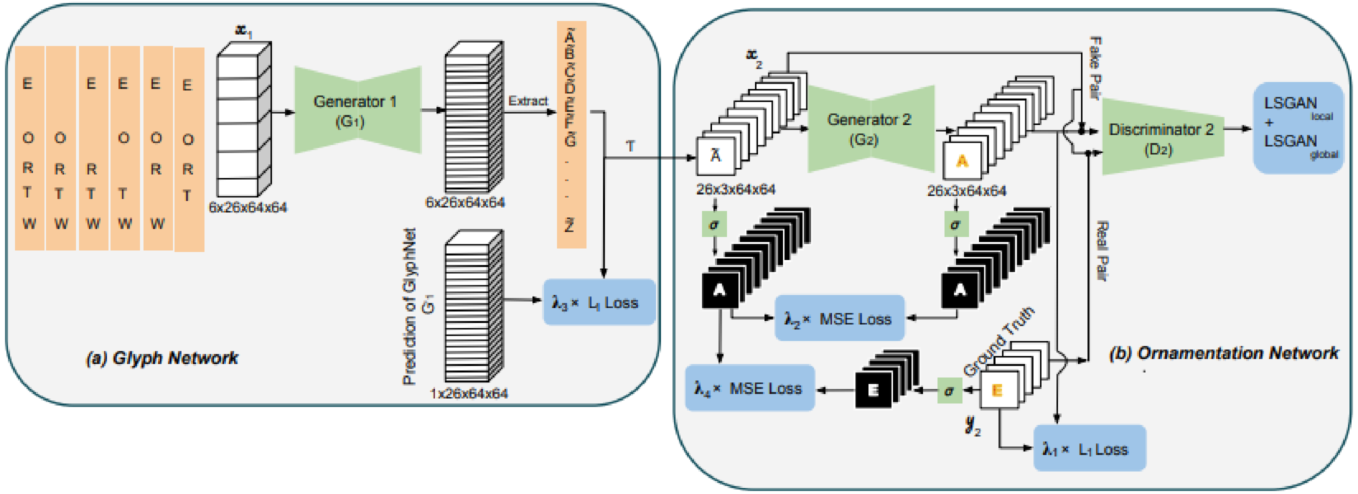


Figure 2: Schematic of our end-to-end MC-GAN model including **(a)** GlyphNet and **(b)** OrnaNet. Inputs and Outputs are illustrated in white, network layers in green, and loss functions are shown in blue. We use a leave-one-out approach among all observed letters of a word like *TOWER* (in orange) to construct a batch of input image stacks to be fed into G_1: For each input stack in the batch, we extract the left out generated glyph. In addition, the remaining 21 glyphs will be generated by feeding in all observed letters together. After a reshape and gray-scale channel repetition, $\mathcal{T}$, these extracted generated glyphs, $\tilde{A}, \tilde{B}, \cdots, \tilde{Z}$ will be fed into OrnaNet.

[6] Multi-Content GAN for Few-Shot Font Style Transfer

Samaneh Azadi, Matthew Fisher, Vladimir G. Kim, Zhaowen Wang, Eli Shechtman, Trevor Darrell

Problem
: Font Style Transfer

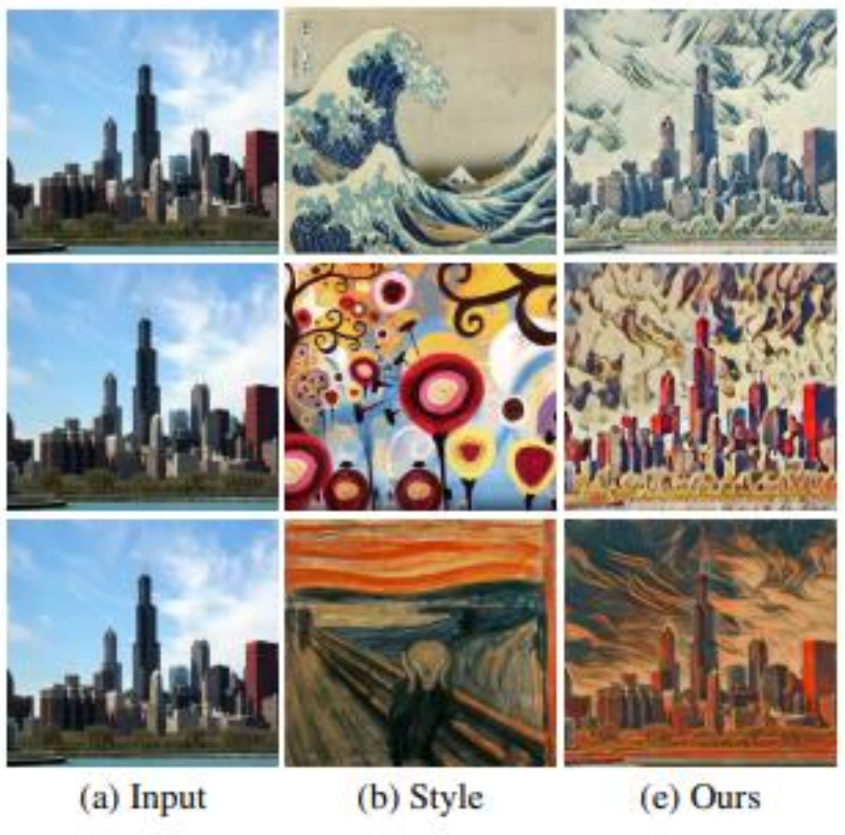
Strategy
- Multi-Content GAN
  - 주어진 글리프의 STYLE을 보이지 않는 콘텐츠로 이전하여 stylization 된 글꼴 캡쳐

Background technique
: GAN

Dataset
: 언급 X

[7] Neural Style Transfer via Meta Networks

Falong Shen, Shuicheng Yan, Gang Zeng



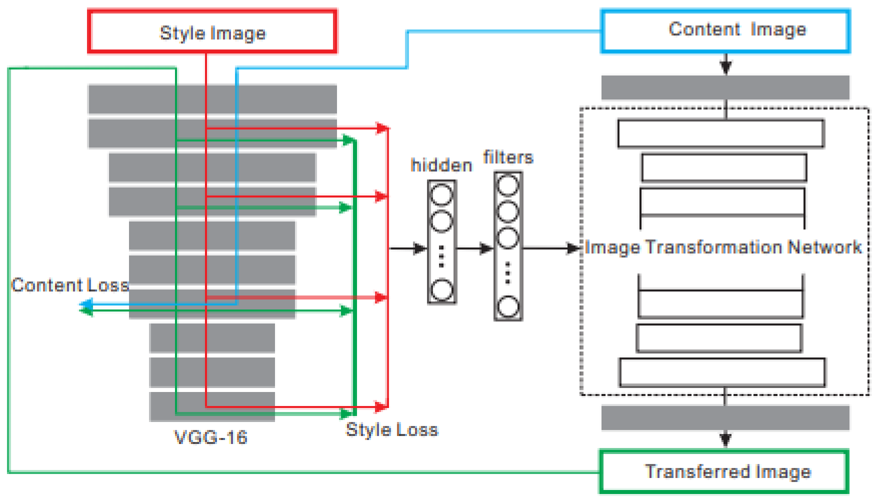


Figure 2: Model architecture. The style image is fed into the fixed VGG-16 to get the style feature, which goes through two fully connected layers to construct the filters for each `conv` layer in the corresponding image transformation network. We fixed the scale and bias in the instance `batchnorm` layer to 1 and 0. The dimension of `hidden` vector is 1792 without specification. The dimension of `filters` vector is in the range from 1×10^5 to 2×10^6, depending on the size of image transformation networks. The hidden features are connected with the filters of each `conv` layer of the network in a group manner to decrease the parameter size, which means a 128 dimensional hidden vector for each `conv` layer. Then we compute the style loss and the content loss for the transferred image with the style image and the content image respectively through the fixed VGG-16.

[7] Neural Style Transfer via Meta Networks

Falong Shen, Shuicheng Yan, Gang Zeng

Problem

: Style Transfer
: 일반적으로 새로운 스타일에 대해 이미지 변환 네트워크를 훈련시키는데 이때 스타일에 대한 generalization이 부족할 경우, SGD(Stochastic Gradient Descent)를 엄청 반복해야 함

Strategy

- Meta Network
    - Style image를 가져와 이미지 변환 네트워크를 직접 생성

Background technique

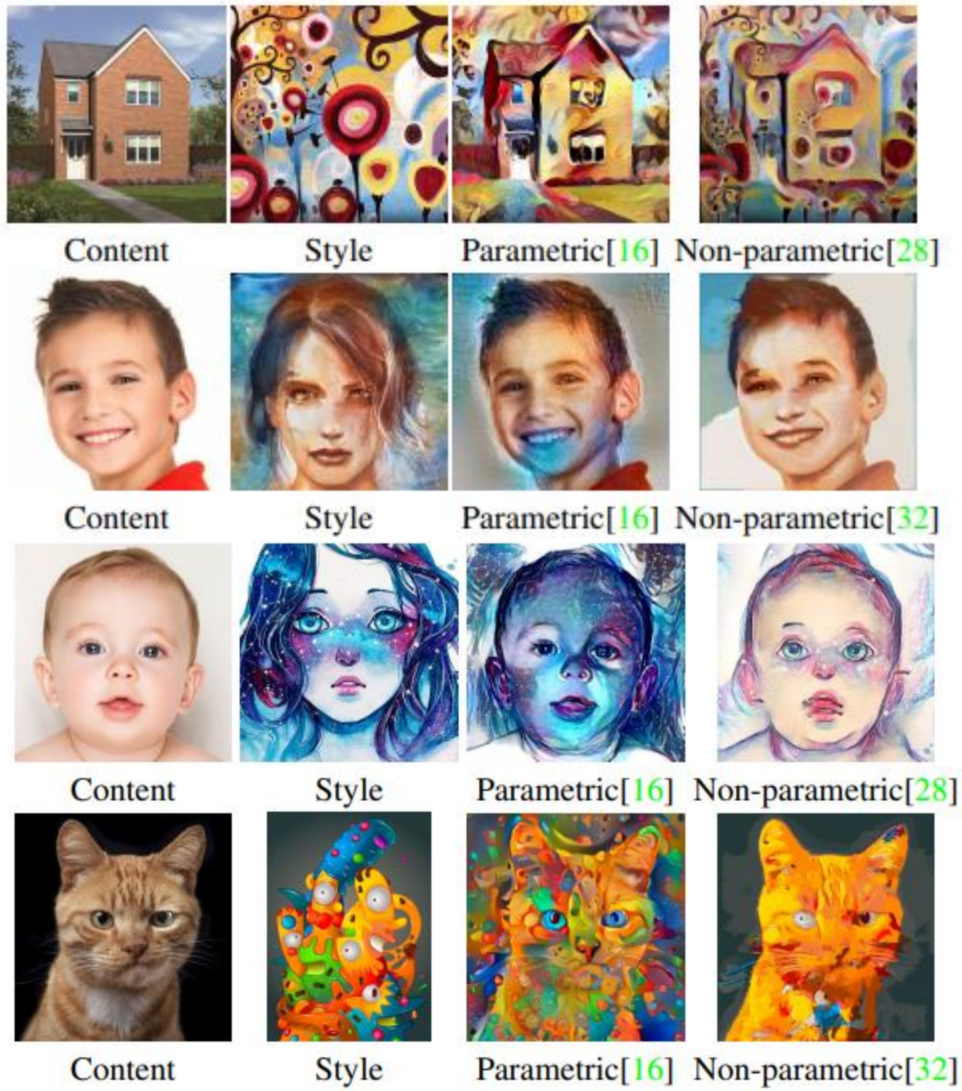
: GAN

Dataset

: WikiArt, MS-COCO

[8] Arbitrary Style Transfer with Deep Feature Reshuffle

Shuyang Gu, Congliang Chen, Jing Liao, Lu Yuan



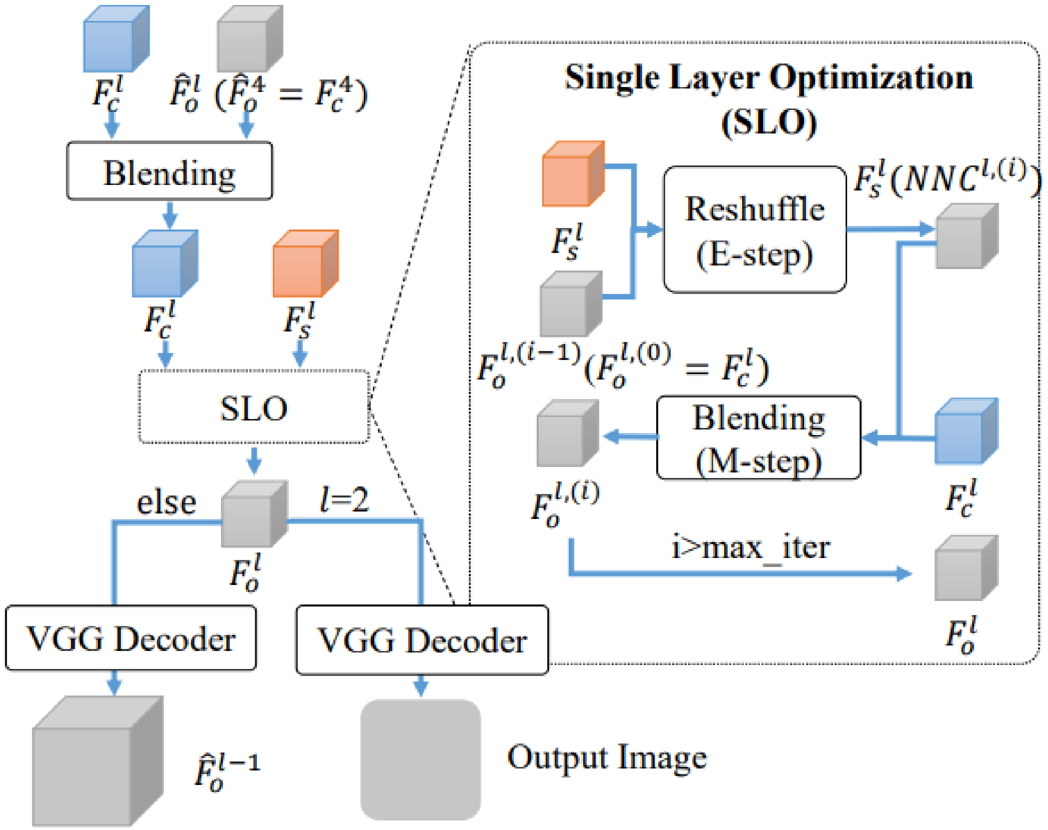


Figure 4. System Pipeline

[8] Arbitrary Style Transfer with Deep Feature Reshuffle

Shuyang Gu, Congliang Chen, Jing Liao, Lu Yuan

Problem
: Arbitrary Style Transfer

Strategy
- Reshuffling deep features of the style image – reshuffle을 기반으로 하는 새로운 style loss는 global/local style loss를 connect
  - 왜곡 방지, semantic-level의 transfer 가능
  - Artifact 방지

Background technique
: CNN

Dataset
: ImageNet

[9] Avatar-Net: Multi-scale Zero-shot Style Transfer by Feature Decoration

Lu Sheng, Ziyi Lin, Jing Shao, Xiaogang Wang

Figure 1. Exemplar stylized results by the proposed Avatar-Net, which faithfully transfers the *Lena* image by arbitrary style.

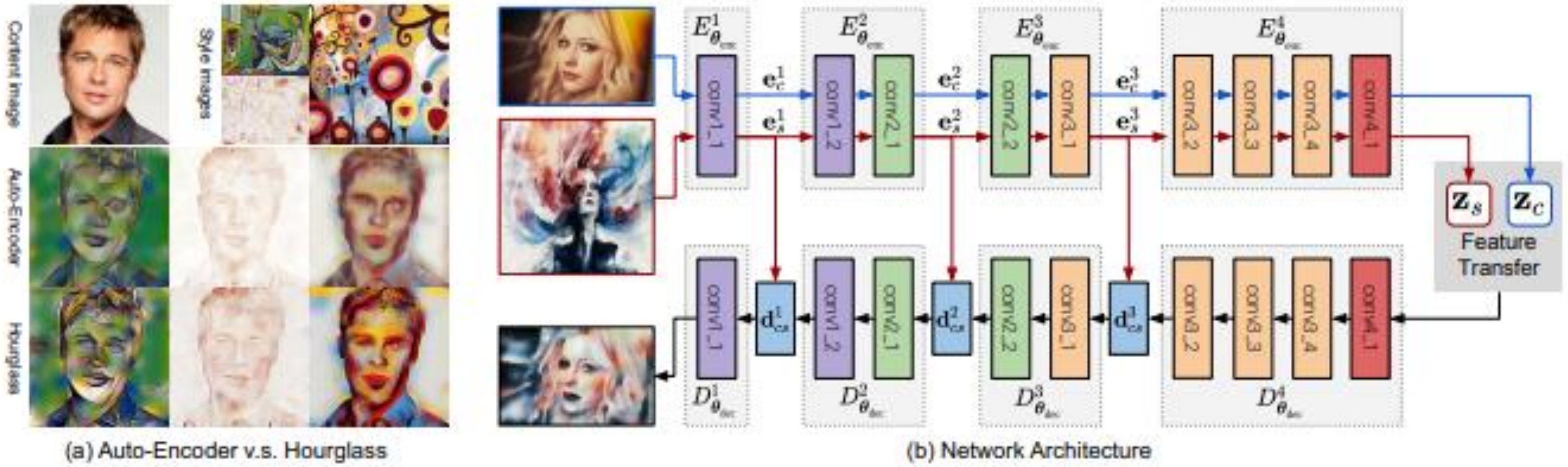


Figure 6. (a) Stylization comparison by auto-encoder and style-augmented hourglass networks. Style decorator is applied as the feature transfer module. The auto-encoder and the hourglass networks share the same main branch. (b) The network architecture of the proposed style-augmented hourglass network. Detailed implementation is depicted in Sec. 5.1.

201810957 염지현

[9] Avatar-Net: Multi-scale Zero-shot Style Transfer by Feature Decoration

Lu Sheng, Ziyi Lin, Jing Shao, Xiaogang Wang

Problem

: Zero-shot Style Transfer
: generalization과 efficiency 사이의 균형 → Zero-shot style transfer 방해

Strategy

- Avatar-Net
  - 임의의 style image에서 semantically aligned style features로 content features 구성
    → 전체적 feature 분포 일치 & 세부 style pattern 보호

Background technique

: CNN, GAN

Dataset

: MS-COCO, MPI Sintel

[10] Separating Style and Content for Generalized Style Transfer

Yexun Zhang, Ya Zhang, Wenbin Cai

	L1 loss	RMSE	PDAR
Pix2pix	0.0105	0.0202	0.17
AEGN	0.0112	0.0202	0.3001
Zitozi	0.0091	**0.0184**	0.1659
C-GAN	0.0112	0.02	0.3685
EMD	**0.0087**	**0.0184**	**0.1332**

Figure 9. Comparison of image generation for known styles and novel contents. Equal number of image pairs with source and target styles are used to train the baselines.

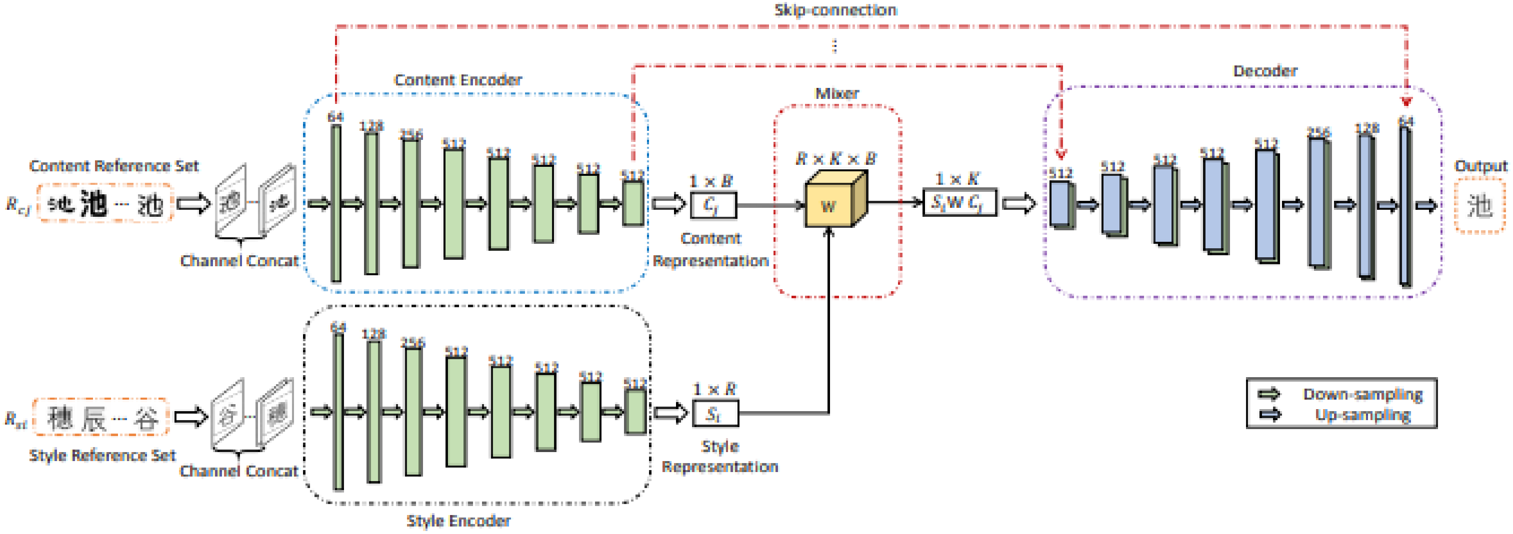


Figure 2. The detailed architecture of the proposed generalized *EMD* model for style transfer.

201810957 염지현

[10] Separating Style and Content for Generalized Style Transfer

Yexun Zhang, Ya Zhang, Wenbin Cai

Problem

: Generalized Style Transfer
: 기존 - source image → target image = 1개의 style로만 transfer 가능

Strategy

- Generalized Style Transfer Net
  - Style과 contents에 대한 표현 분리
  - Style encode, content encoder, mixer, decoder

Background technique

: CNN, GAN

Dataset

: 언급 x

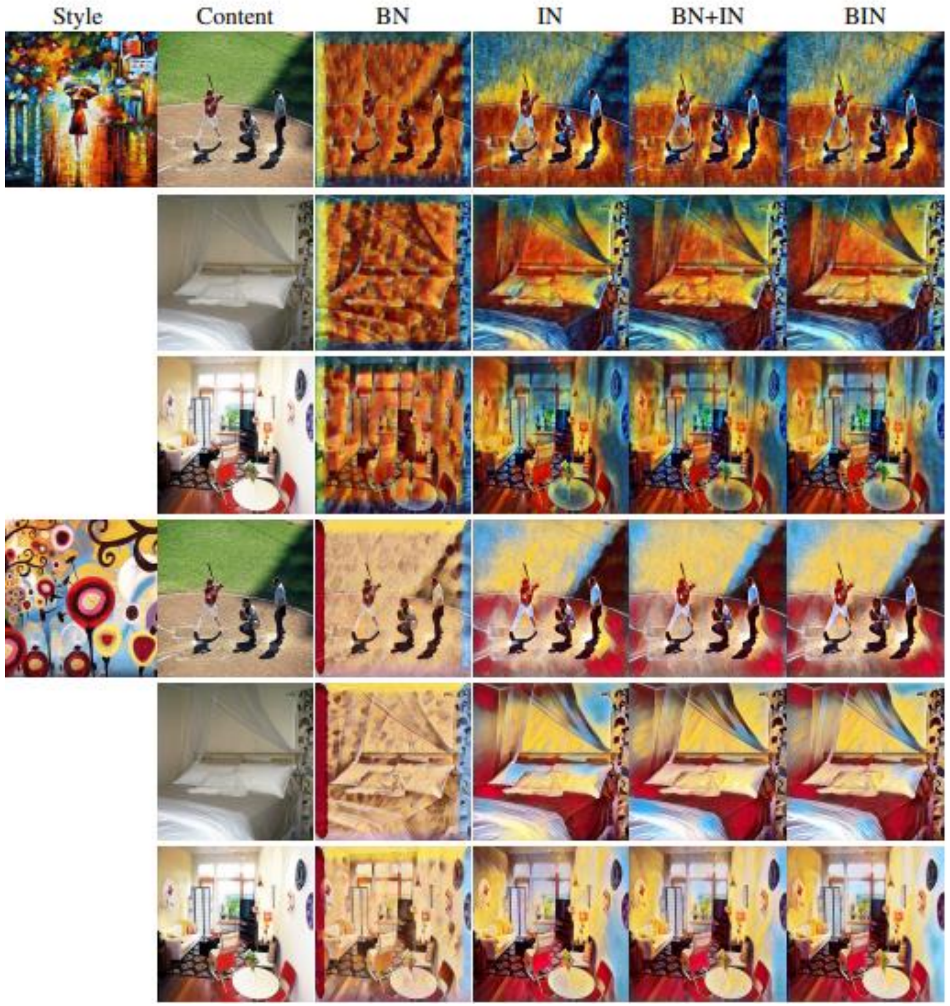
Style
&
Stylization

AAAI

201810957 염지현

[11] Batch-Instance Normalization for Adaptively Style-Invariant Neural Networks

Hyeonseob Nam, Hyo-Eun Kim



[11] Batch-Instance Normalization for Adaptively Style-Invariant Neural Networks

Hyeonseob Nam, Hyo-Eun Kim

Problem
: real-world image recognition은 textures, lighting conditions 등으로 인해 어려움
: 해결을 위해 암묵적으로 data & layer 수의 증가

Strategy
- 명시적으로 style 정보를 조작하자!
- BIN(Batch-Instance Normalization)
  - 필요한 style 유지
  - 방해되는 style 정규화

Background technique
: GAN, IN, BN

Dataset
: CIFAR-10/100, ImageNet

Style
&
Stylization

ECCV

201810957 염지현

[12] Game of Sketches: Deep Recurrent Models of Pictionary-Style Word Guessing

Ravi Kiran Sarvadevabhatla, Shiv Surya, Trisha Mittal, R. Venkatesh Babu

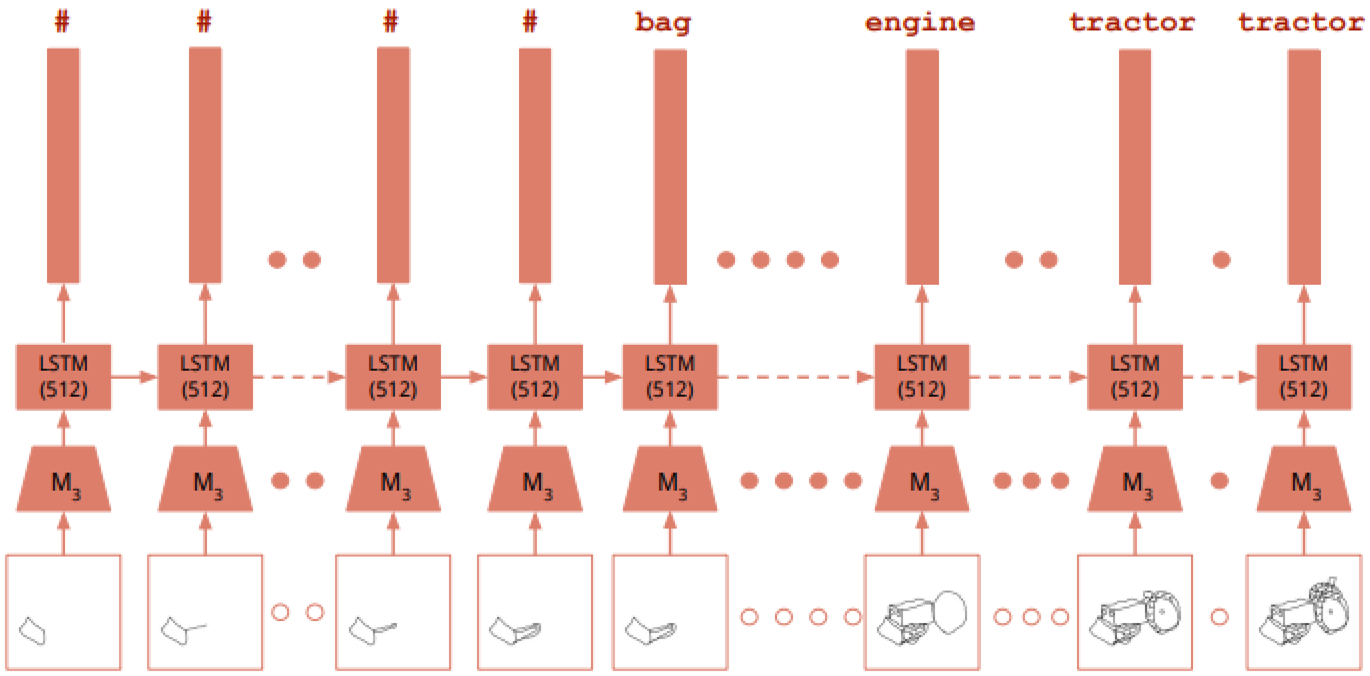


Fig. 9: The architecture for our deep neural model of word guessing. The rectangular bars correspond to guess-word embeddings. M_3 corresponds to the CNN regressor whose penultimate layer's outputs are used as input features to the LSTM model. "#" reflects our choice of modelling 'no guess' as a pre-defined non-word embedding. See Section 6 for details.

[12] Game of Sketches: Deep Recurrent Models of Pictionary-Style Word Guessing

Ravi Kiran Sarvadevabhatla, Shiv Surya, Trisha Mittal, R. Venkatesh Babu

Problem
: 그림 보고 어떤 개체인지 맞추기

Strategy
- 스케치의 시간적 변화 → 단어 추측
- Sketch-QA → 인간과 같은 실수를 일부로 넣으면서 인간 모방 요인 충족

Background technique
: CNN, LSTM

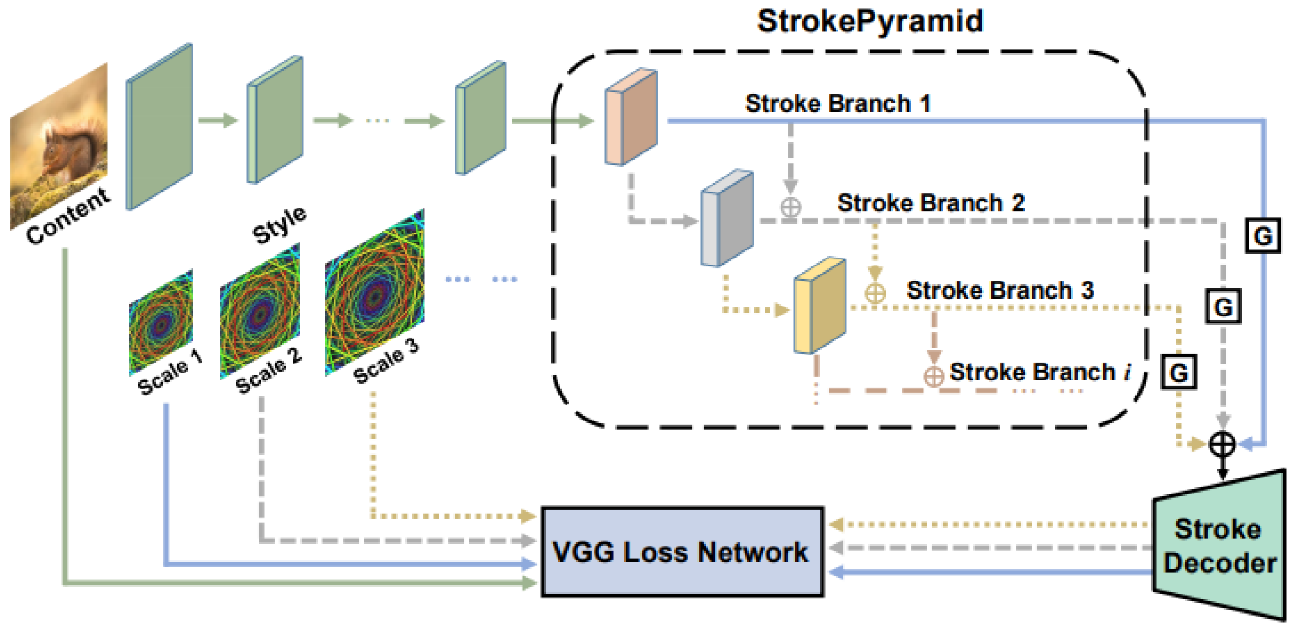
Dataset
: WordGUESS-160, TV-Berlin

[13] Stroke Controllable Fast Style Transfer with Adaptive Receptive Fields?

Yongcheng Jing, Yang Liu, Yezhou Yang, Zunlei Feng, Yizhou Yu, Dacheng Tao, Mingli Song

(a) Content & Style

(d) Our proposed approach



[13] Stroke Controllable Fast Style Transfer with Adaptive Receptive Fields?

Yongcheng Jing, Yang Liu, Yezhou Yang, Zunlei Feng, Yizhou Yu, Dacheng Tao, Mingli Song

Problem
: Stroke Controllable Fast Style Transfer

Strategy
- Stroke controllable style transfer network
  - stackPyramid : adaptive receptive field
  - Achieve faster convergence & 훈련된 model에서 new stroke size 증가

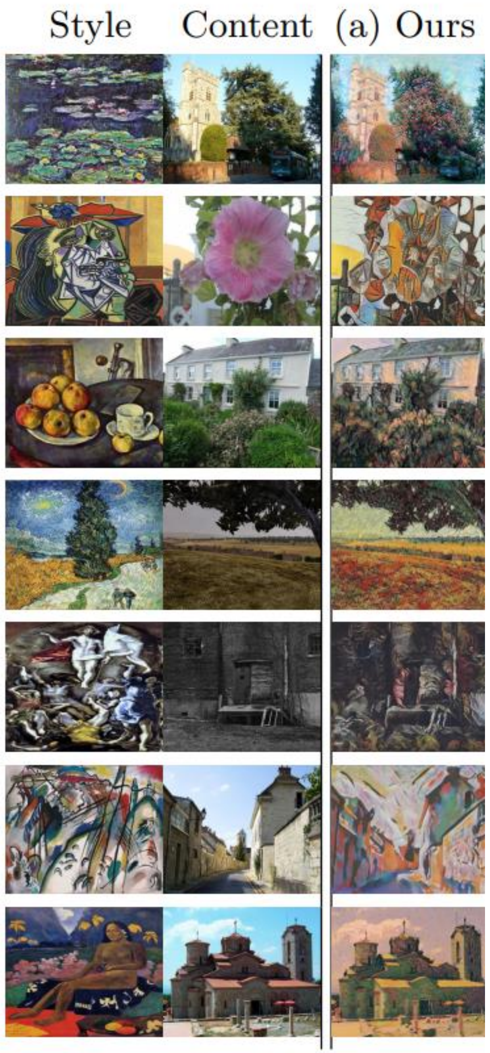
Background technique
: CNN

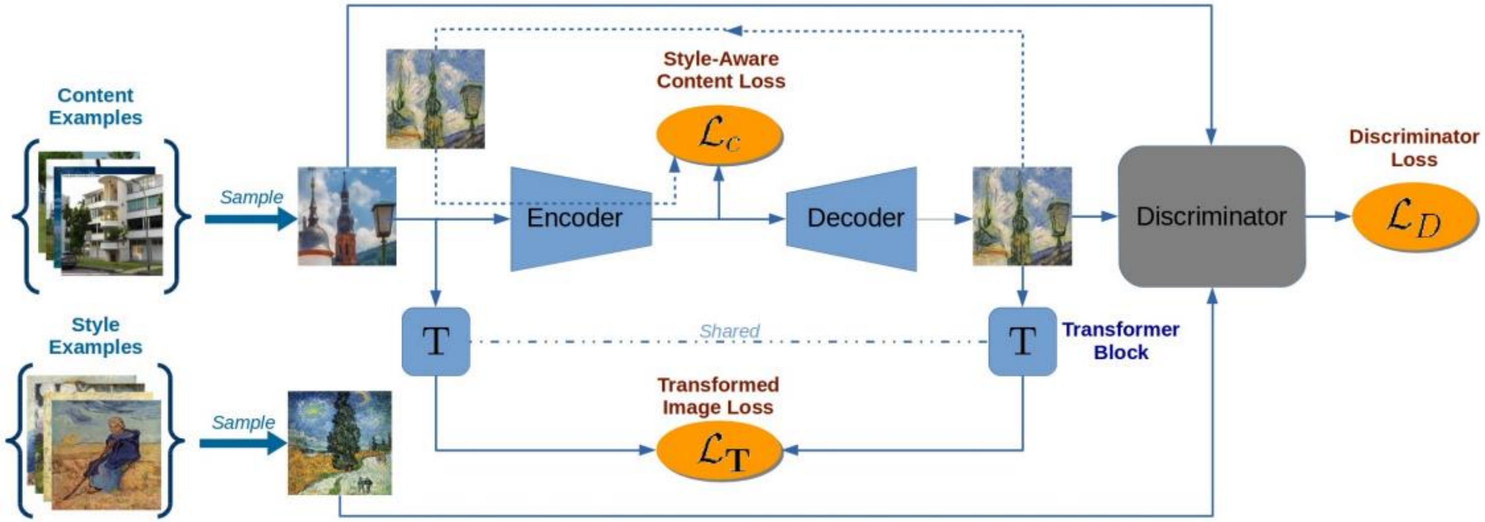
Dataset
: MS-COCO

201810957 염지현

[14] A Style-Aware Content Loss for Real-time HD Style Transfer

Artsiom Sanakoyeu, Dmytro Kotovenko, Sabine Lang, Bjorn Ommer





201810957 염지현

[14] A Style-Aware Content Loss for Real-time HD Style Transfer

Artsiom Sanakoyeu, Dmytro Kotovenko, Sabine Lang, Bjorn Ommer

Problem

: Style transfer – 속도 높이기
: 이전 작업들은 RGB 영역 또는 ImageNet에서 사전 훈련된 CNN에서 comparison art에 의존
→ Label이 지정된 object bounding box 필요
→ Bias 발생 위험

Strategy

- Deep encoder-decoder network for real-time
- High-resolution
- Style-aware content loss
- 미술사가 직접 품질 평가

Background technique

: CNN

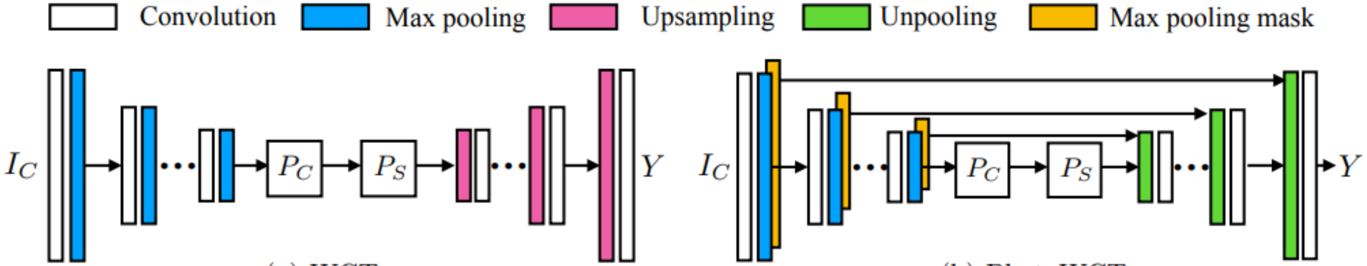
Dataset

: Wikiart

[15] A Closed-form Solution to Photorealistic Image Stylization

Artsiom Sanakoyeu, Dmytro Kotovenko, Sabine Lang, Bjorn Ommer

(a) Style (b) Content (e) Ours



(a) WCT (b) PhotoWCT

Fig. 3: The PhotoWCT and WCT share the same encoder architecture and projection steps. In the PhotoWCT, we replace the upsampling layers (pink) with unpooling layers (green). Note that the unpooling layer is used together with the pooling mask (yellow) which records *where* carries the *maximum* over each max pooling region in the corresponding pooling layer [35].

[15] A Closed-form Solution to Photorealistic Image Stylization

Artsiom Sanakoyeu, Dmytro Kotovenko, Sabine Lang, Bjorn Ommer

Problem

: Photorealistic Image Stylization → artifact 존재, 일관성(X)

Strategy

- Stylization step: reference image style로 content image stylization
- Smoothing step : 공간적으로 일관된 stylization

Background technique

: CNN

Dataset

: MS-COCO

감사합니
다