얼굴로 뭐하지?

코드 / 수식 없는 Introduction to Face Deep Learning

발표자 소개

OO대학교 컴퓨터 전공 학부생

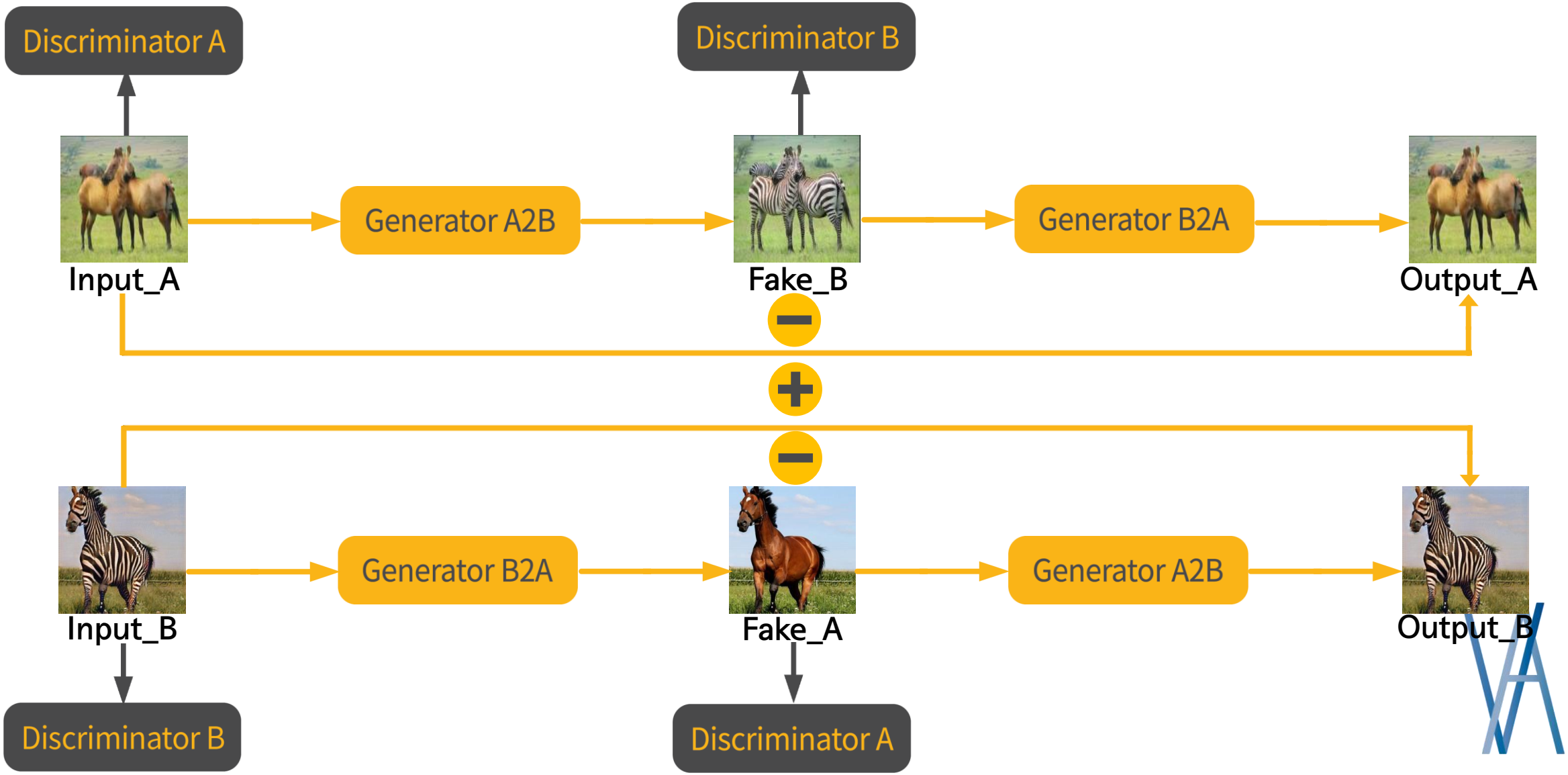
취준(좋은 회사 가고 싶..) or 대학원?

한 분야를 깊게 파기 보다는
AI에 어떤 재미있는 프로젝트가 있을까?

프로젝트 – 동아리 1회, 인턴 2회

What I did... – Character Generation(2017)

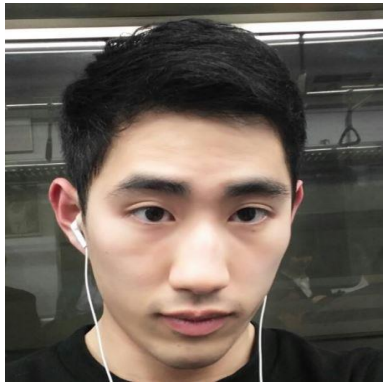
- 사람 얼굴(Input) 을 넣으면 애니메이션 캐릭터(Output)를 만들어주는 인공지능



What I did... – Character Generation(2017)

- 성공(?) 케이스



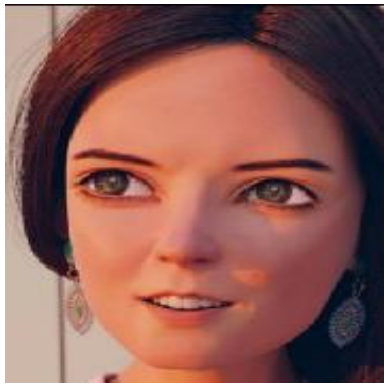





What I did... – Character Generation(2017)

- 실패 케이스








What I did... – Video Synthesis(2018)

What I did... – Video Synthesis(2018)

- 정해인 → 유해진(?)

What I did... – ObamaNet(2019)

- 내가 원하는 말을 하는 오바마 영상을 만들어 줘! → 인공지능 앵커에 적용 가능

제 발표, 이런 분들이 들으면 좋습니다!

제 발표, 이런 분들이 들으면 좋습니다!

코드없는 수식없는 Face Deep learning – 키워드 중심!
난이도 ★☆☆☆☆

☺ 얼굴 데이터에 대해서 전혀 모르시는 분

☺ 잘 모르지만 그냥 한번 돌려보고 싶으신 분

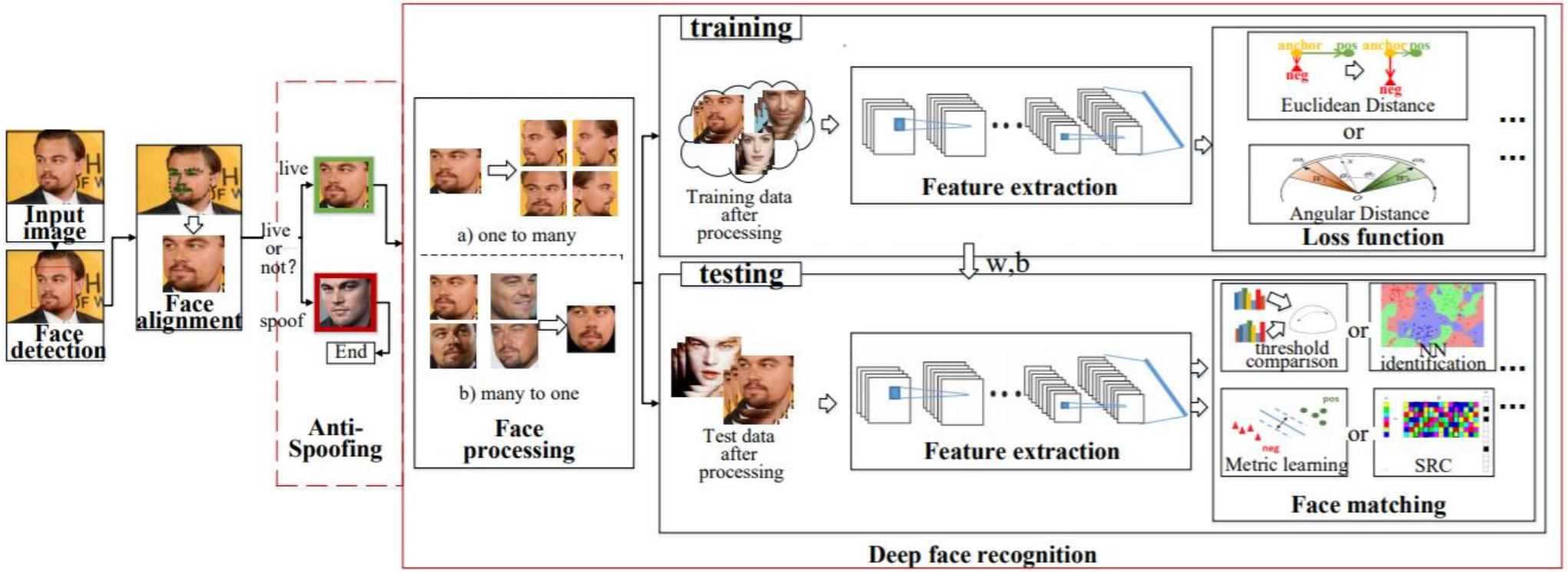
😇; 얼굴 관련 데이터로 프로젝트 하고 계시는 분

😯얼굴 분야 SOTA 논문의 원리를 정확하게 알고 싶으신 분

그럼, 이제부터 진짜 시작합니다!

Q. 얼굴 인식을 하고 싶습니다! 간단하게 돌려볼 수 있는 거 있을까요?

A. 얼굴 인식 하려면 이렇게 많은 것들을 해야 합니다...



https://arxiv.org/pdf/1804.06655.pdf

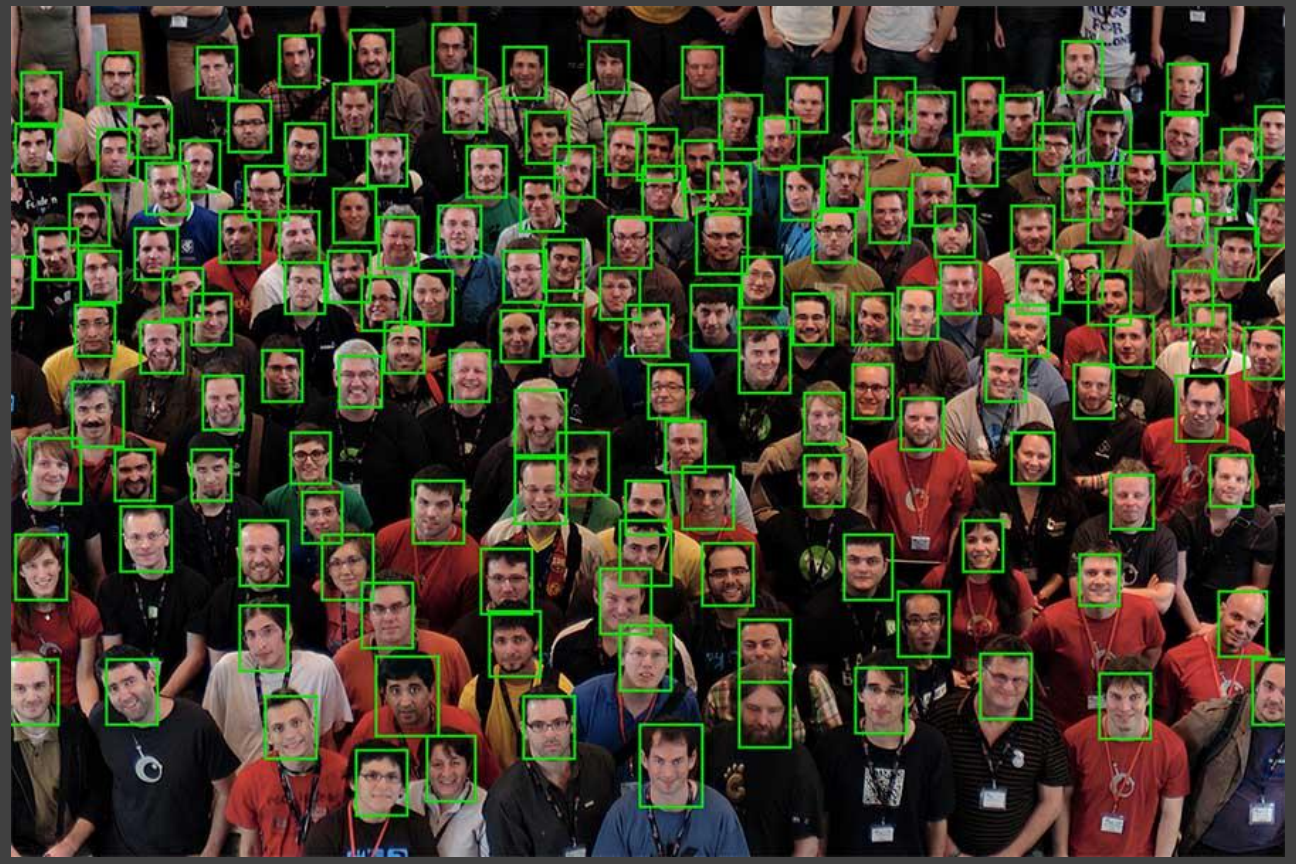
Q. 얼굴 데이터로 어떤 태스크를 해 볼 수 있을까요?(3점)

A. 얼굴 데이터로 이런 것까지!?

- Face Detection(얼굴 탐지)
- Face Alignment(얼굴 정렬)
- Face Recognition(얼굴 인식)
- Face Clustering(얼굴 분류)
- Face Synthesis(얼굴 생성)
- Face Swap(얼굴 변환)
- Face Super-Resolution(얼굴 고해상도 변환)
- 3D Face Reconstruction(3D 얼굴 복원)
- Facial Expression Analysis(감성 분석)
- Age Prediction(나이 예측)
- …(기타 등등)

사람들이 이렇게나 얼굴에 관심이 많다(...) → 연구주제가 엄청 많다!

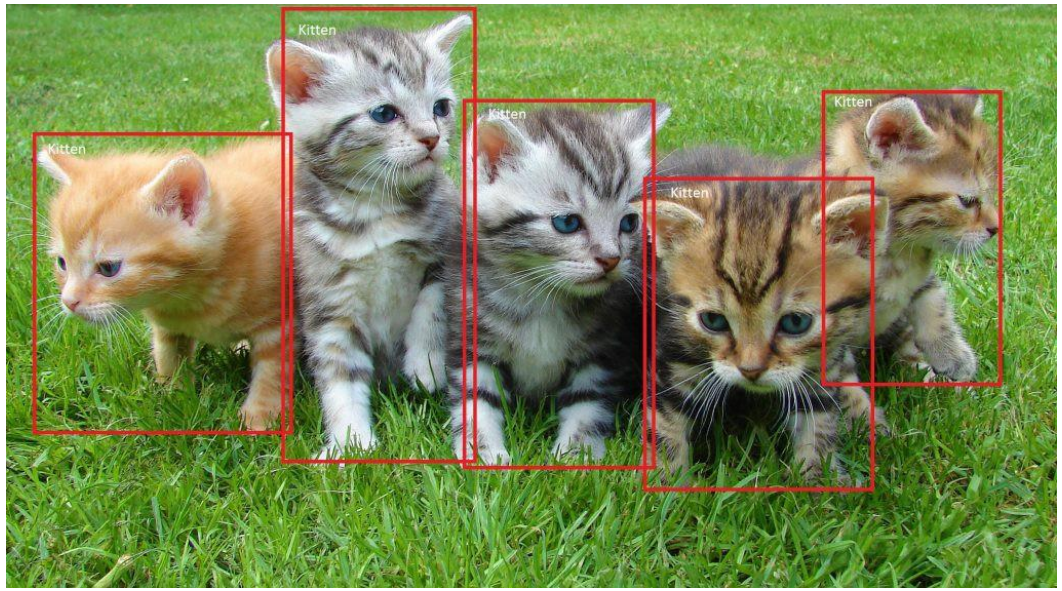
예를 들어, 인공지능 출석체크 어플리케이션을 만든다고 가정해 봅시다.



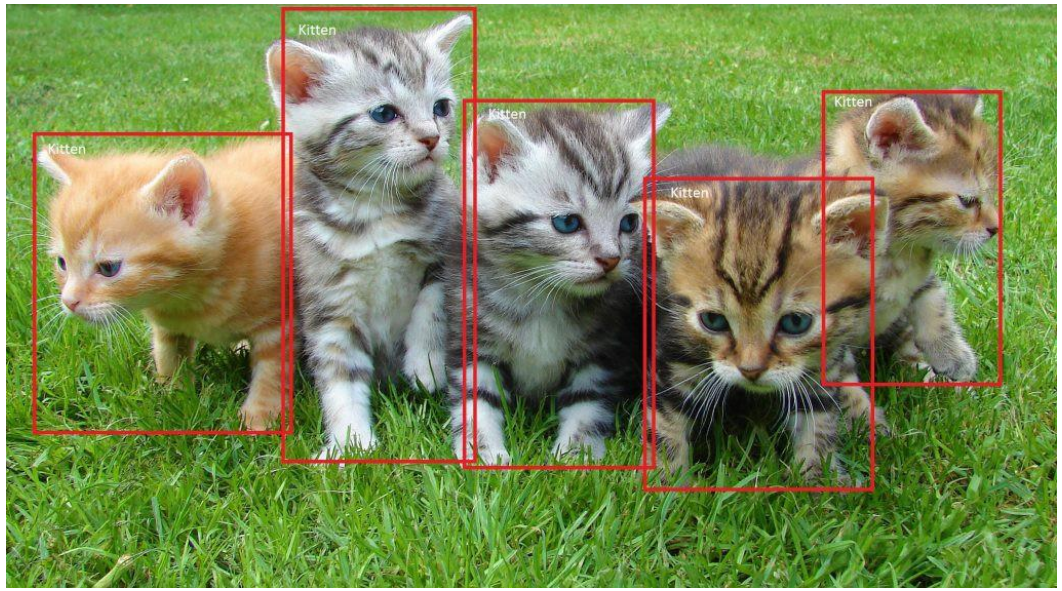
Raw 이미지에서 얼굴부터 찾아야겠다!

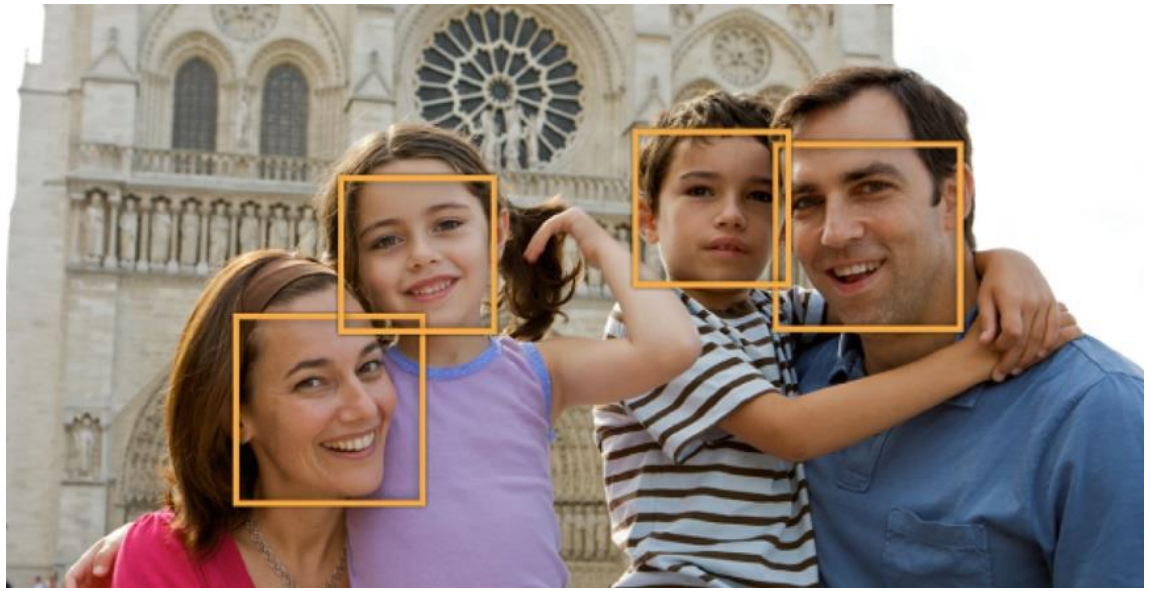
Detection?

Object Detection – Input Image에서 (사전에 정의된 Label) Object 찾아주세요!



Face Detection – Input Image에서 모든 얼굴 찾아주세요!





Face Detection Dataset, 이런 식으로 구성되어 있습니다!

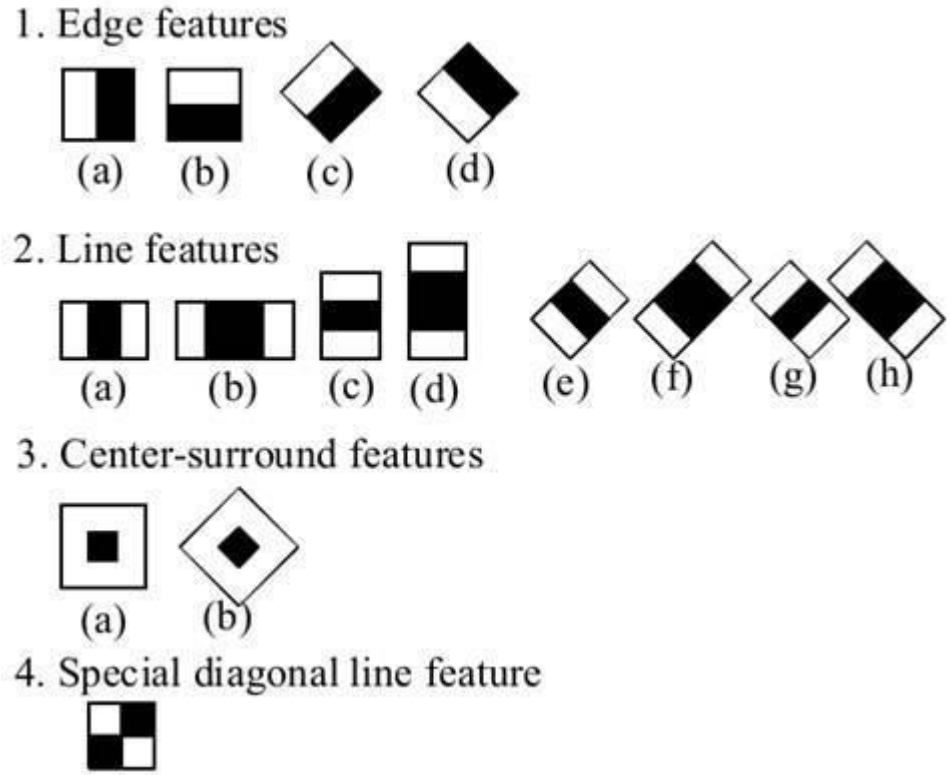
WIDER FACE - x1, y1, w, h, blur, expression, Illumination, invalid, occlusion, pose

http://shuoyang1213.me/WIDERFACE/

딥러닝 이전 Face Detection, 이렇게 했습니다! – Haar Cascade(2001)

딥러닝 이전 Face Detection, 이렇게 했습니다! – Haar Cascade(2001)

Haar-like Feature



딥러닝 이전 Face Detection, 이렇게 했습니다! – Haar Cascade (2001)

영어사전

cascade 미국·영국 [kæ'skeɪd] 영국식 다른 뜻(1건) | 예문보기

1. 작은 폭포 2. 폭포처럼 쏟아지는 물 3. 풍성하게 늘어진 것

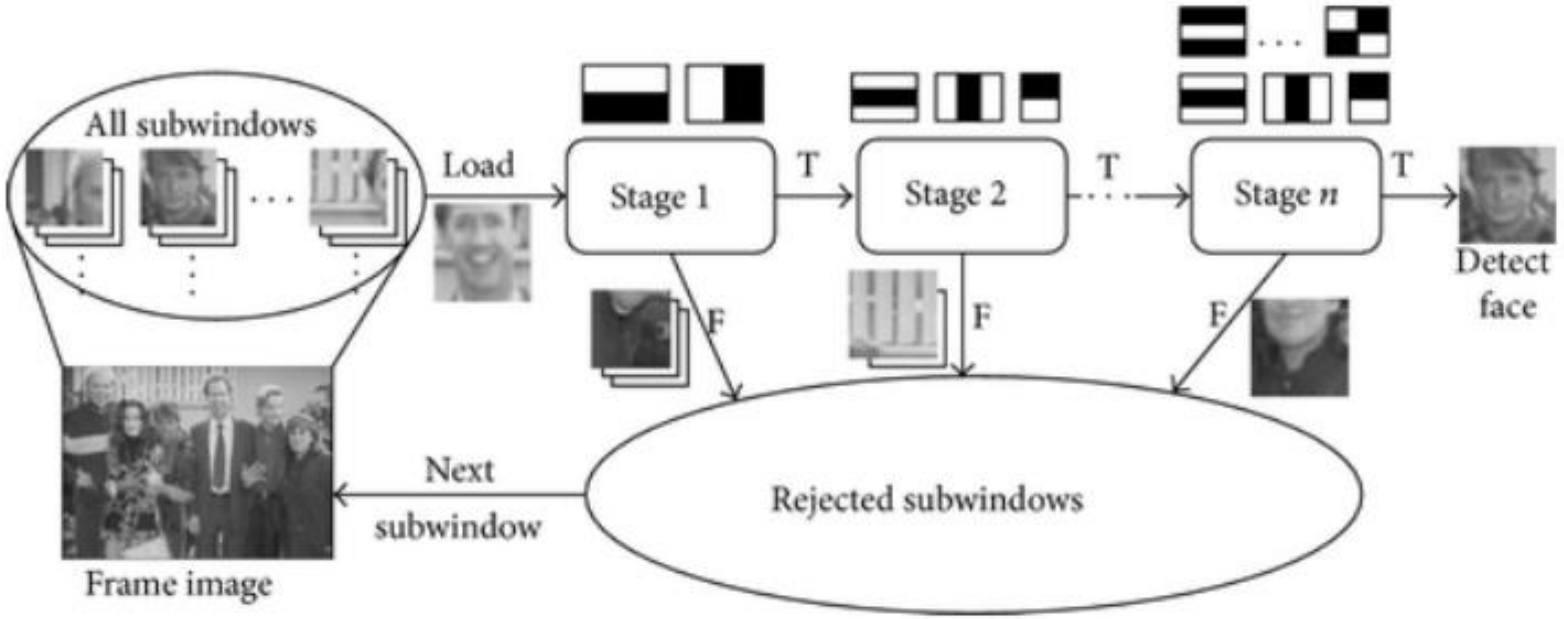
→ Sliding Window를 돌면서 Input Image 에 Haar-like Feature 를 중첩시킨다.

딥러닝 이전 Face Detection, 이렇게 했습니다! – Haar Cascade

영어사전

cascade 미국·영국 [kæ'skeɪd] 영국식 다른 뜻(1건) | 예문보기 → N개의 If – Else (?)

1. 작은 폭포 2. 폭포처럼 쏟아지는 물 3. 풍성하게 늘어진 것



딥러닝 이전 Face Detection, 이렇게 했습니다! – Haar Cascade

https://www.youtube.com/watch?v=hPCTwxF0qf4

Face Detection, Dlib(또는 OpenCV)에 이미 구현되어 있습니다.

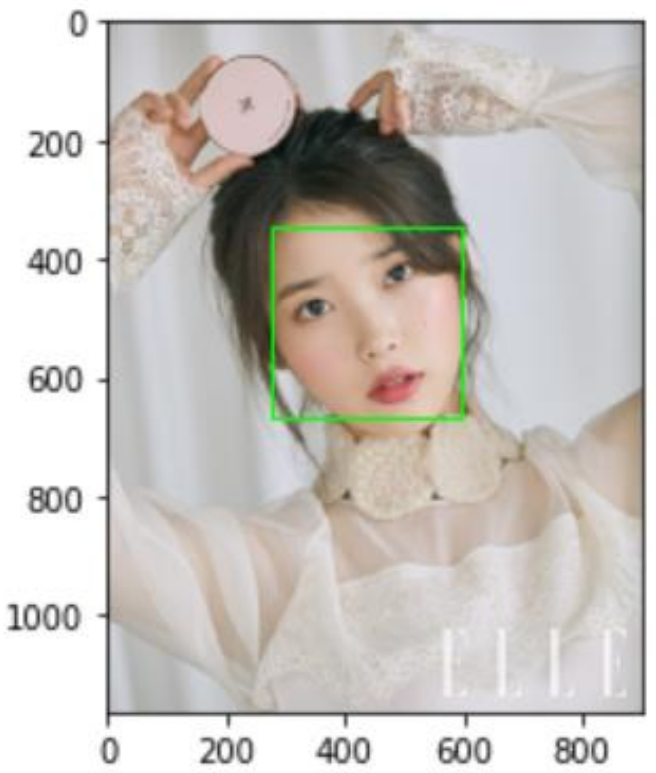
Face Detection, Dlib(또는 OpenCV)에 이미 구현되어 있습니다.

How to Detect Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

```
import dlib
face_detector = dlib.get_frontal_face_detector() # HOG Face Detector
bbox = face_detector(image, 1) # Face Detect
```

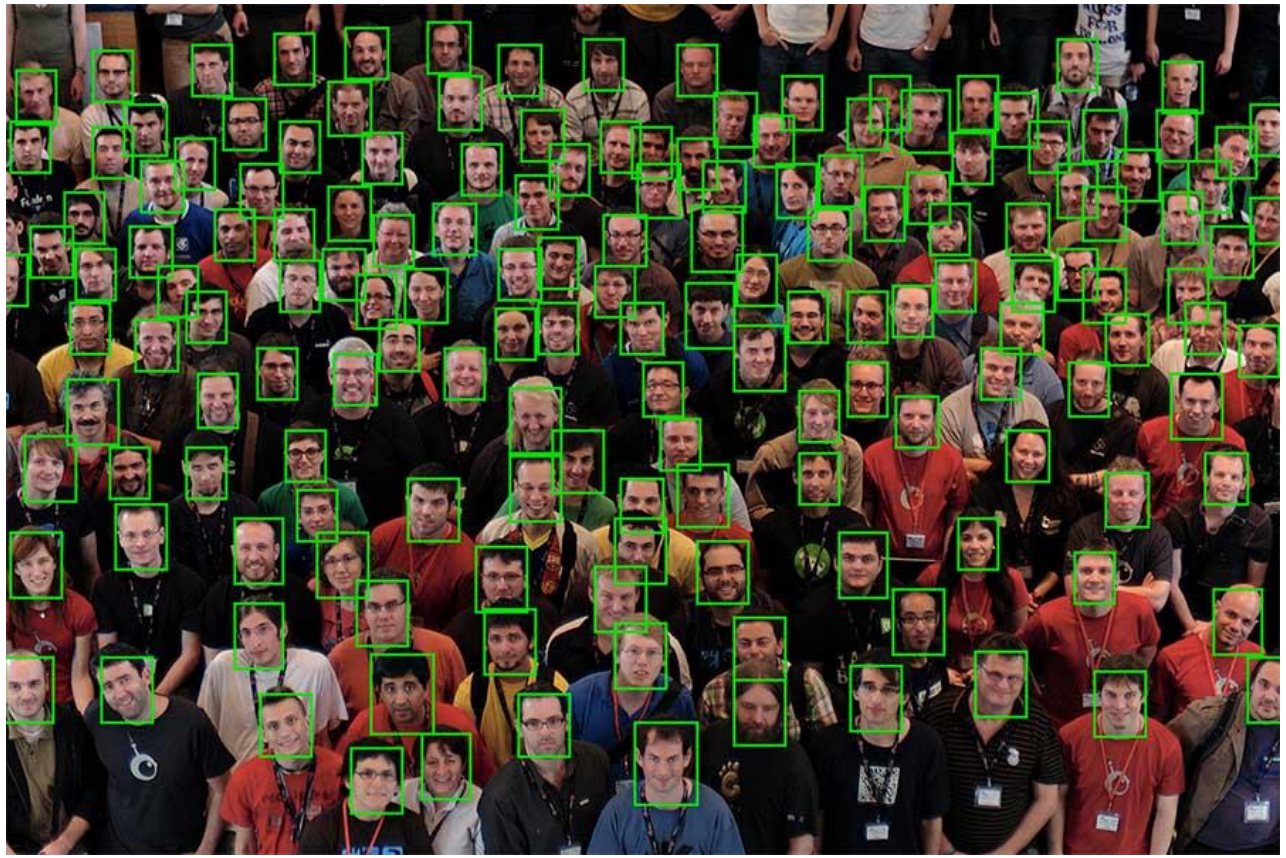
How to Detect Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

```
import dlib
face_detector = dlib.get_frontal_face_detector() # HO
                                                 # Face Detect
```

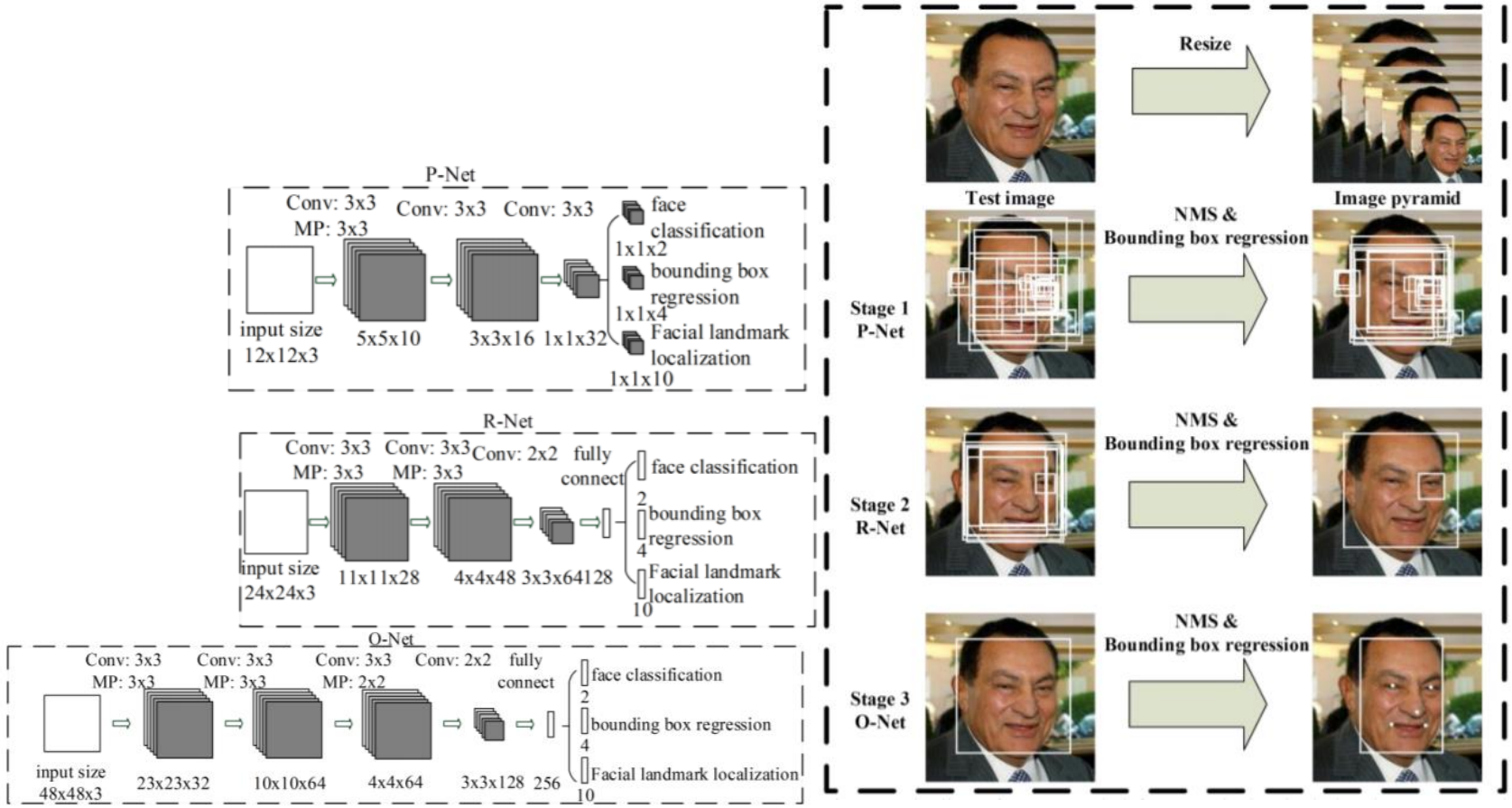


딥러닝 이후 Face Detection, 이렇게 했습니다!

딥러닝 이전 : Sliding Window 돌면서 Filter 규칙에 맞는 Object 찾아줘!(If-Else-_-)
딥러닝 이후 : (엄청 많은) 데이터 넣고 이거랑 비슷한 거 찾아줘!

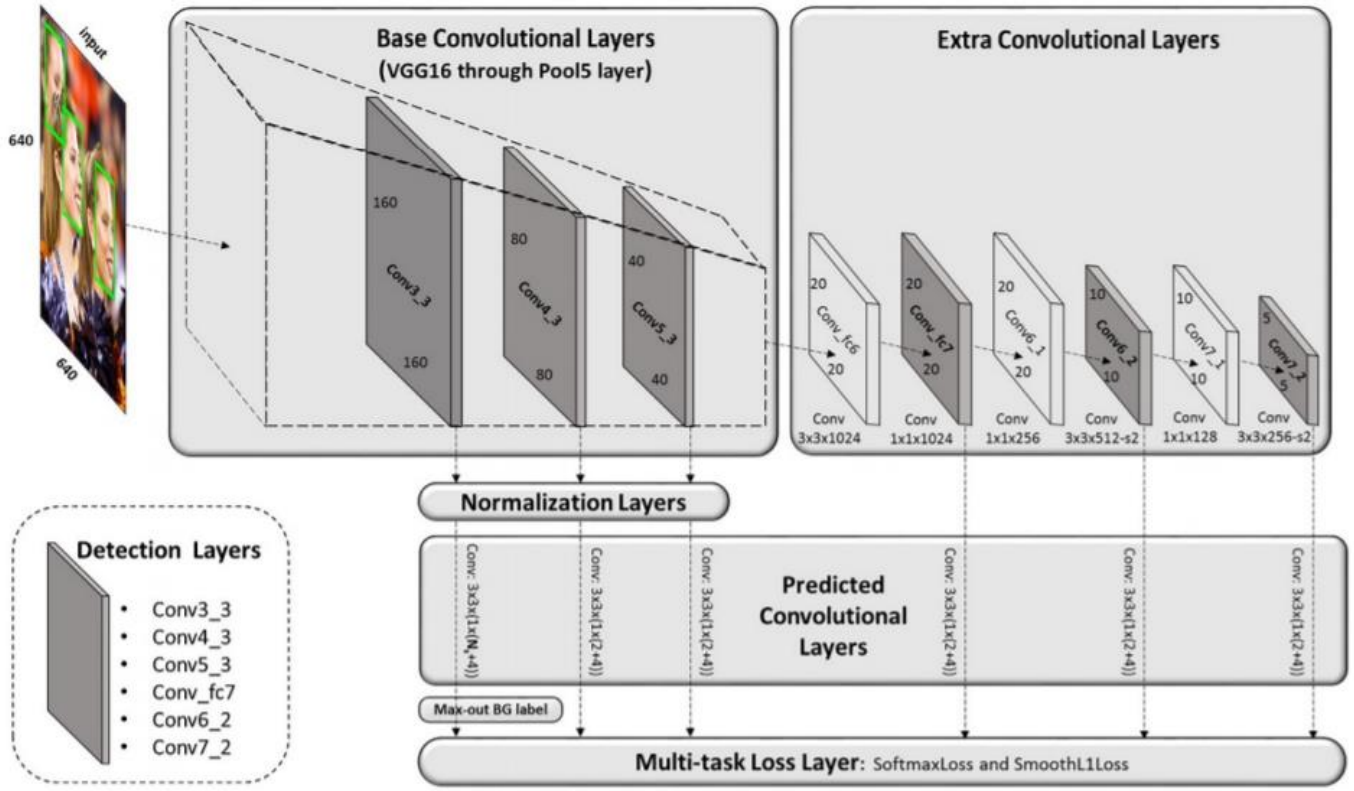


딥러닝 이후 Face Detection, 이렇게 했습니다! – MTCNN(2016)



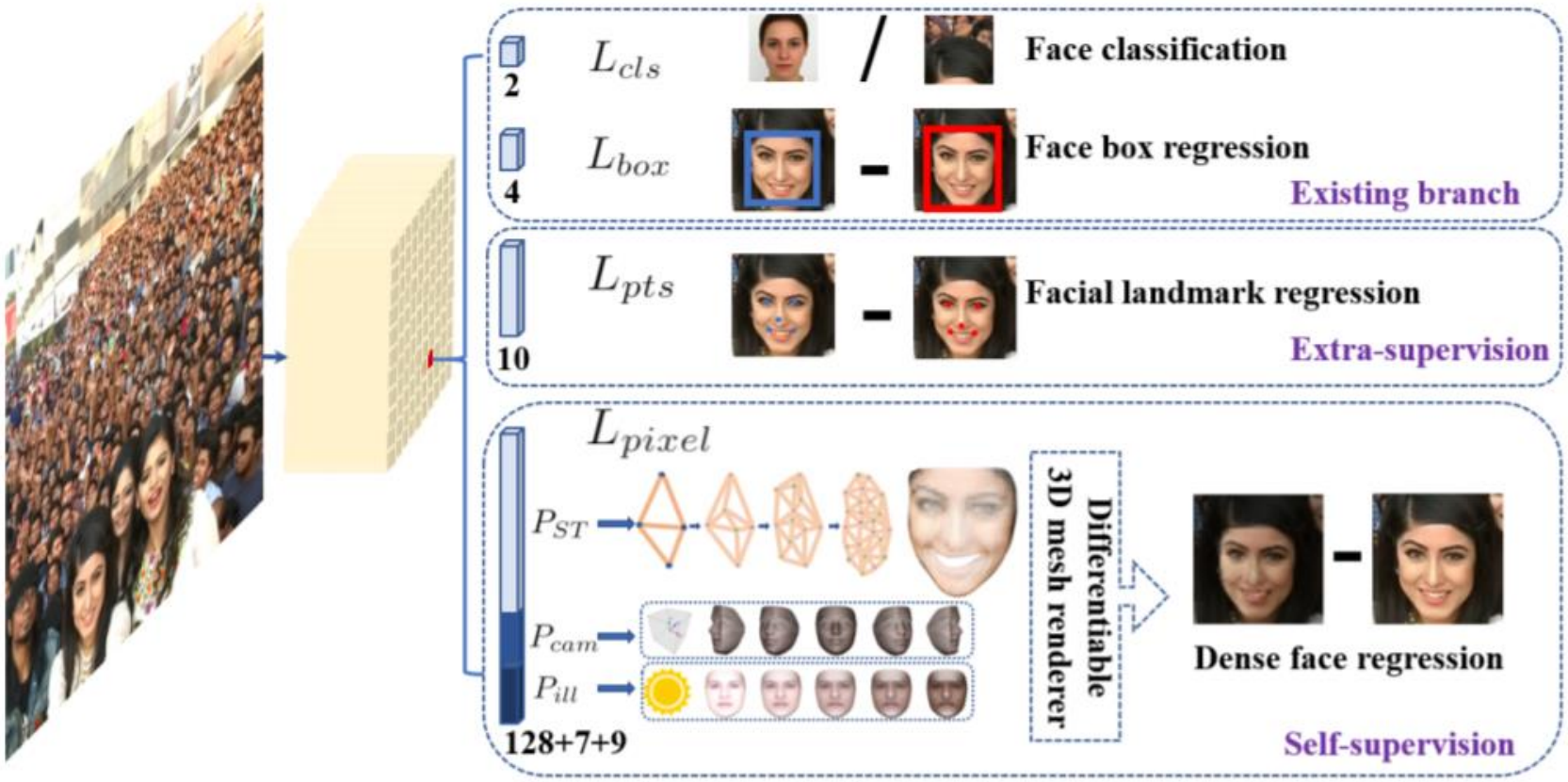
https://arxiv.org/ftp/arxiv/papers/1604/1604.02878.pdf

딥러닝 이후 Face Detection, 이렇게 했습니다! – S^3FD(2017)



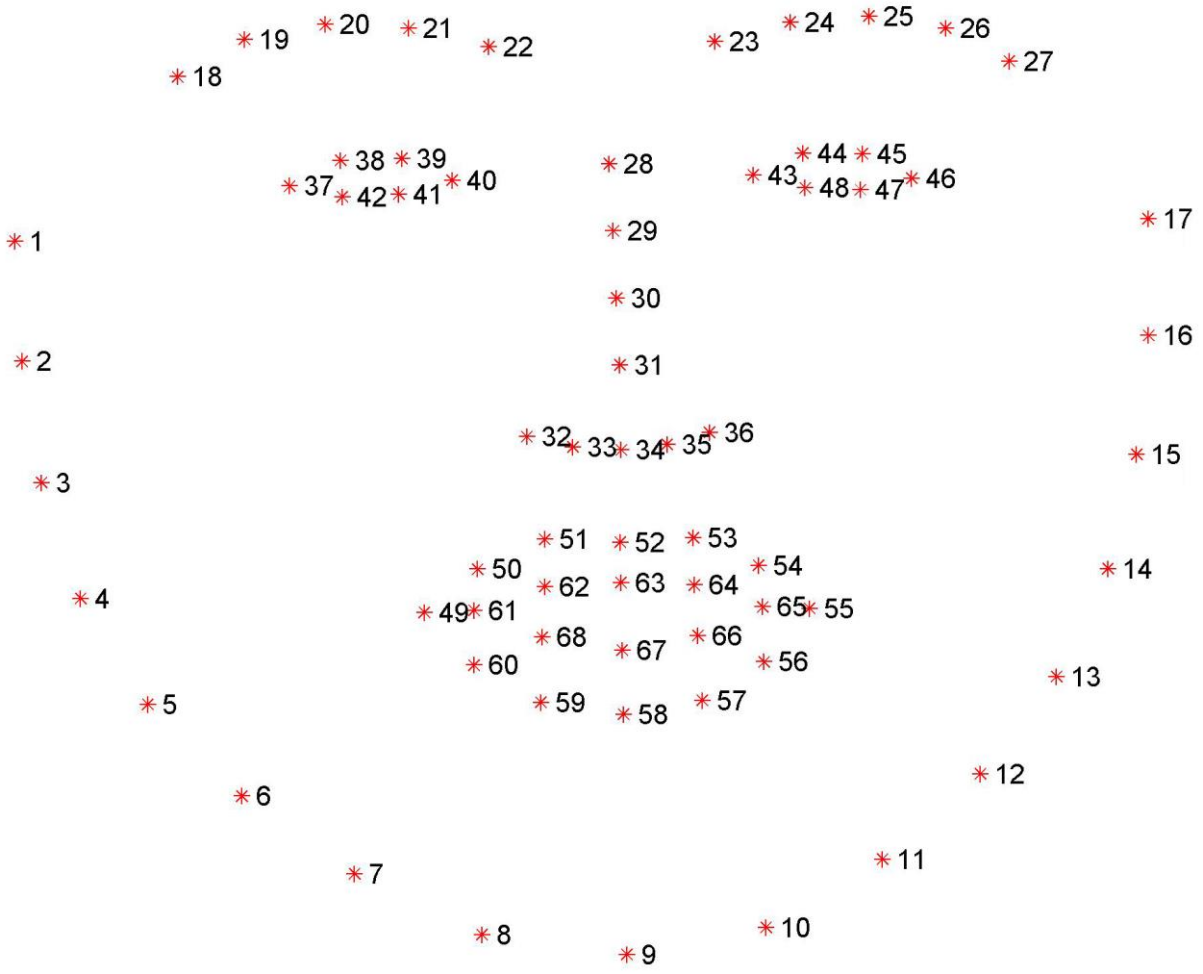
https://arxiv.org/pdf/1708.05237.pdf

딥러닝 이후 Face Detection, 이렇게 했습니다! – RetinaFace(2019)

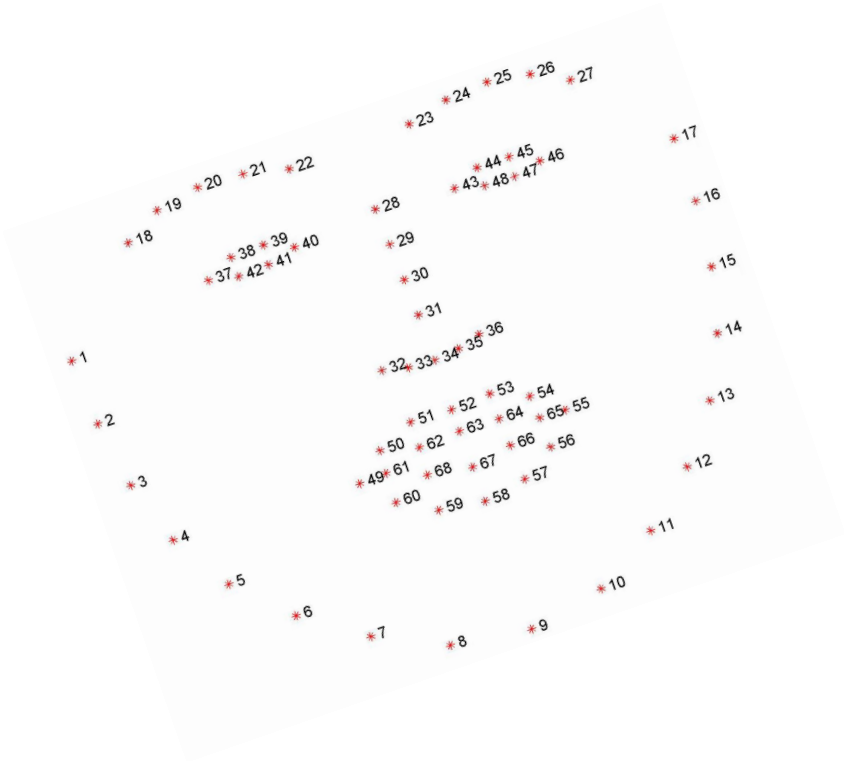


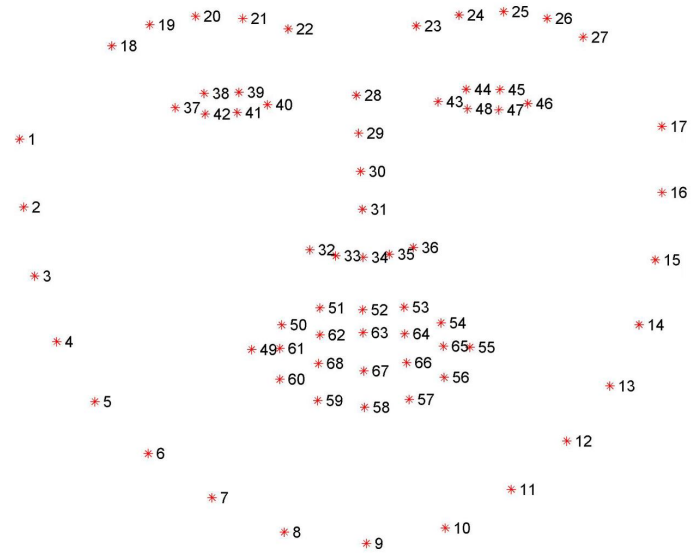
날 것의(Raw) 얼굴을 Align(일종의 전처리?)

Facial Landmark – 얼굴 특징점!



How to Align Face – 얼굴 특징점을 바탕으로 돌리면 됩니다!





How to Align Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

How to Align Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

```
import dlib
face_detector = dlib.get_frontal_face_detector() # HOG Face Detector
bbox = face_detector(image, 1) # Face Detect
landmark_detector = dlib.shape_predictor('./shape_predictor_68_face_landmarks.dat')
face = landmark_detector(image, bbox[0]) # Facial Landmark Detection
aligned_face = dlib.get_face_chip(image, face,size = 256, padding = 0.1)
```

How to Align Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

```
import dlib
face_detector = dli
bbox = face_detecto
landmark_detector =
face = landmark_det
aligned_face = dlib
```

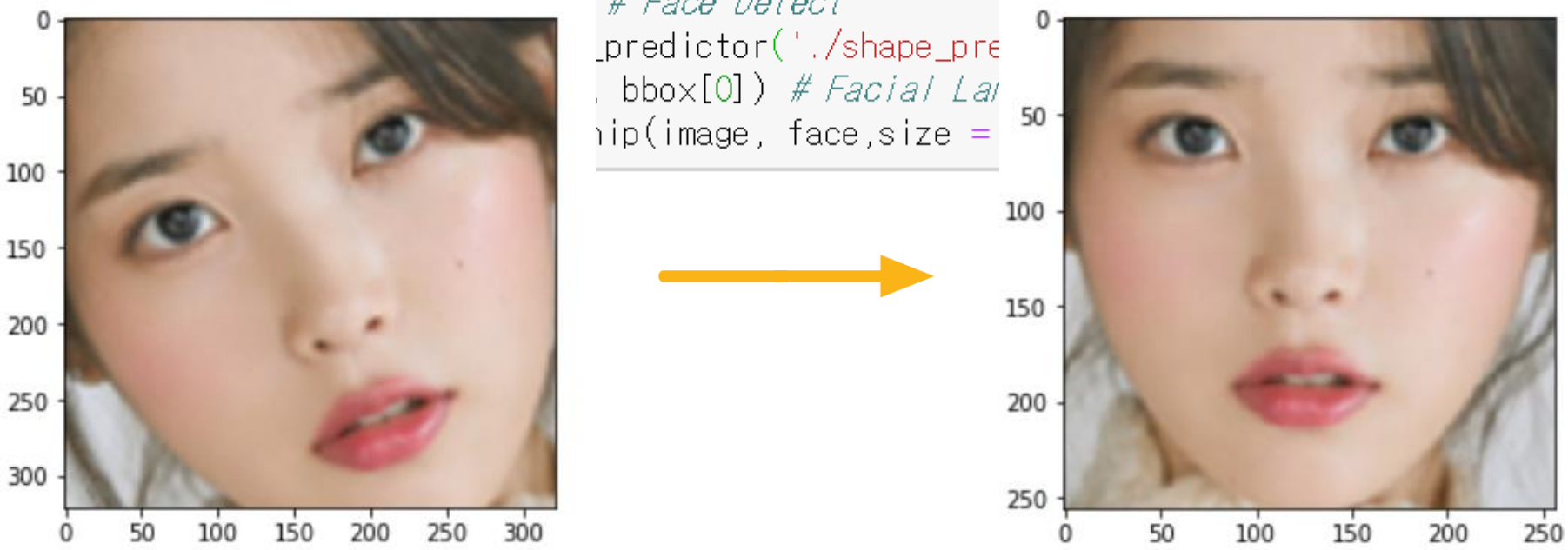
```
ector

face_landmarks.dat')
ction
ng = 0.1)
```

How to Align Face – Dlib이 묻지도 따지지도 않고 알아서 다 해줍니다.

```
import dlib
face_detector = dlib.get_frontal_face_detector() # HOG Face Detector
```

```
# Face Detect
_predictor('./shape_pre
 bbox[0]) # Facial La
ip(image, face,size =
```



Face Alignment Dataset, 이런 식으로 구성되어 있습니다!

- 각 점의 Index와 좌표(+ Pose?)

(c) 300W-test

(d) Menpo2D-test

https://arxiv.org/pdf/1812.01936.pdf

정제된 얼굴 데이터를 본격적으로 분석!

Face Recognition Dataset, 엄청 많다...

- LFW, CASIA , MegaFace, Youtube Faces, MS-Celeb, VGGFace2...

TABLE VI
THE COMMONLY USED FR DATASETS FOR TRAINING

Datasets	Publish Time	#photos	#subjects	# of photos per subject [1]	Key Features
MS-Celeb-1M (Challenge 1)[69]	2016	10M 3.8M(clean)	100,000 85K(clean)	100	breadth; central part of long tail; celebrity; knowledge base
MS-Celeb-1M (Challenge 2)[69]	2016	1.5M(base set) 1K(novel set)	20K(base set) 1K(novel set)	1/-/100	low-shot learning; tailed data; celebrity
MS-Celeb-1M (Challenge 3) [2]	2018	4M(MSv1c) 2.8M(Asian-Celeb)	80K(MSv1c) 100K(Asian-Celeb)	-	breadth;central part of long tail; celebrity
MegaFace [105], [145]	2016	4.7M	672,057	3/7/2469	breadth; the whole long tail;commonalty
VGGFace2 [22]	2017	3.31M	9,131	87/362.6/843	depth; head part of long tail; cross pose, age and ethnicity; celebrity
CASIA WebFace [243]	2014	494,414	10,575	2/46.8/804	celebrity
UMDFaces-Videos [10]	2017	22,075	3,107	–	video
VGGFace [149]	2015	2.6M	2,622	1,000	depth; celebrity; annotation with bounding boxes and coarse pose
CelebFaces+ [187]	2014	202,599	10,177	19.9	private
Google [176]	2015	>500M	>10M	50	private
Facebook [195]	2014	4.4M	4K	800/1100/1200	private

[1] The min/average/max numbers of photos or frames per subject

Face Recognition Dataset, 이런 식으로 구성되어 있습니다!

- 1번째 사람 N장, 2번째 사람 N장 ... M번째 사람 N 장 → Number of Identity, Image per Identity

(c) John Wesley Shipp

(d) Leymah Gbowee

(e) Princess Haya Bint Al Hussein

(f) Julio César Chávez Jr.

(g) Roy Jones Jr.

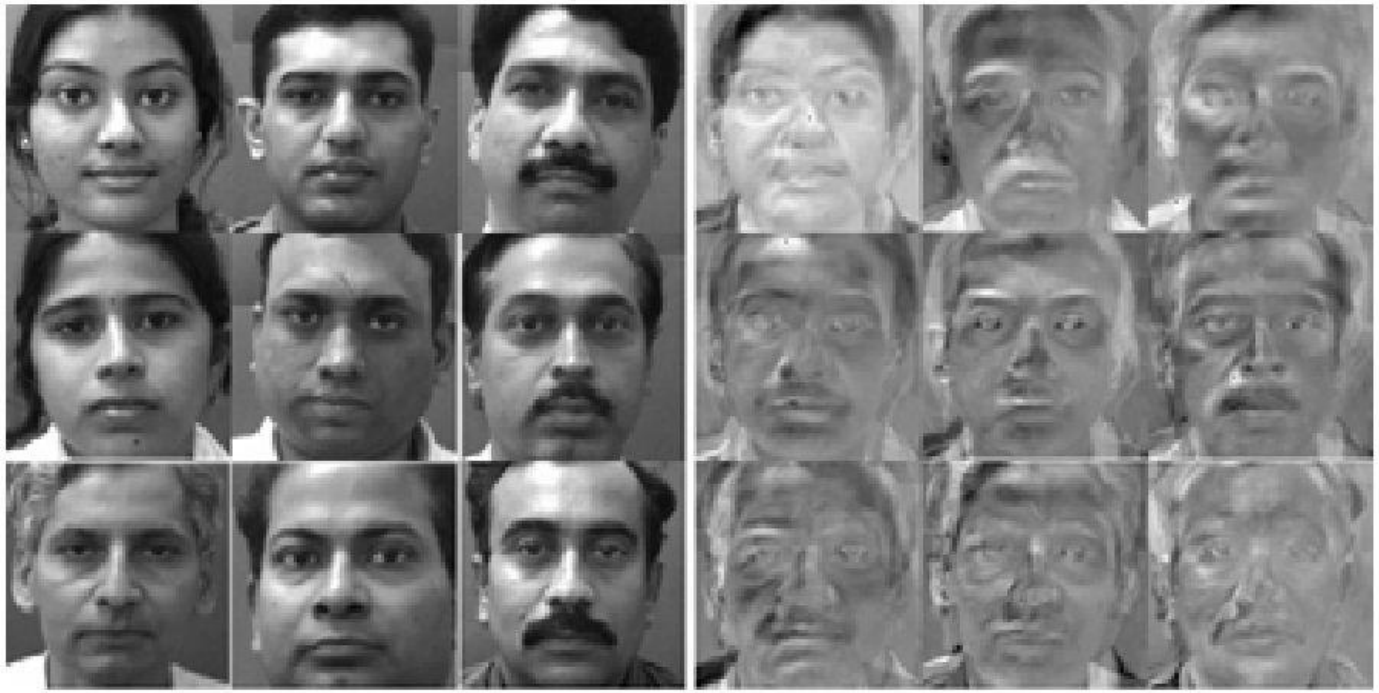
(h) Ruby Lin

(i) Additi Gupta

(j) Lee Joon-gi

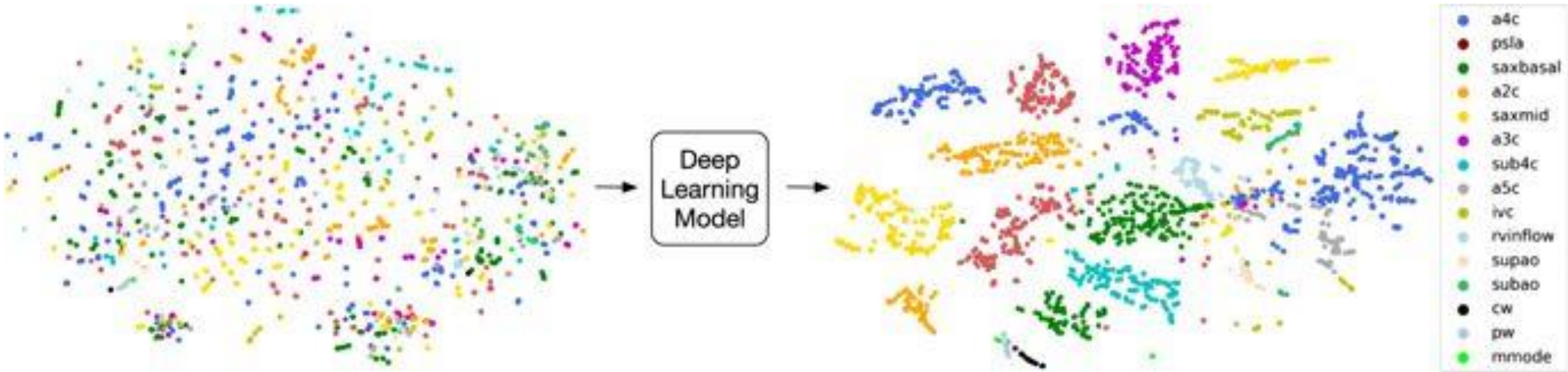
https://arxiv.org/pdf/1710.08092.pdf

딥러닝 이전 Face Recognition, 이렇게 했습니다! – Eigenface(1991)

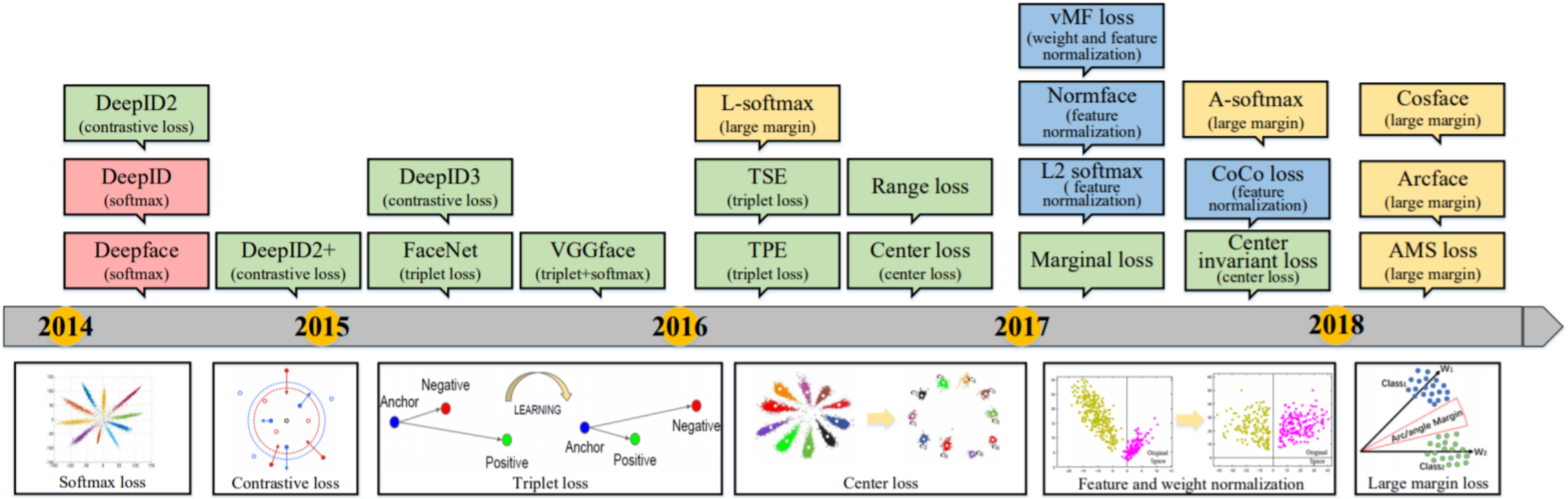


딥러닝 이후 Face Recognition ≒ Classification

비슷한 얼굴끼리는 가깝게, 다른 얼굴끼리는 최대한 멀게!



딥러닝 이후 Face Recognition, 너무 많다...



https://arxiv.org/pdf/1804.06655.pdf

딥러닝 이후 Face Recognition, 이렇게 했습니다! – FaceNet(2015)

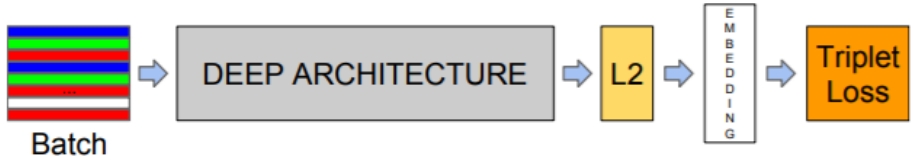


Figure 2. **Model structure.** Our network consists of a batch input layer and a deep CNN followed by L_2 normalization, which results in the face embedding. This is followed by the triplet loss during training.

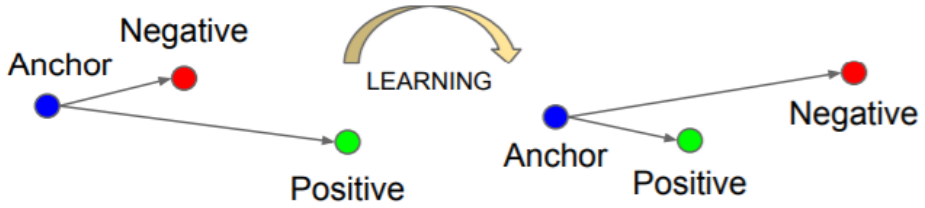


Figure 3. The **Triplet Loss** minimizes the distance between an *anchor* and a *positive*, both of which have the same identity, and maximizes the distance between the *anchor* and a *negative* of a different identity.

```python
def triplet_loss(anchor, positive, negative, alpha):
    """Calculate the triplet loss according to the FaceNet paper

    Args:
      anchor: the embeddings for the anchor images.
      positive: the embeddings for the positive images.
      negative: the embeddings for the negative images.

    Returns:
      the triplet loss according to the FaceNet paper as a float tensor.
    """
    with tf.variable_scope('triplet_loss'):
        pos_dist = tf.reduce_sum(tf.square(tf.subtract(anchor, positive)), 1)
        neg_dist = tf.reduce_sum(tf.square(tf.subtract(anchor, negative)), 1)

        basic_loss = tf.add(tf.subtract(pos_dist,neg_dist), alpha)
        loss = tf.reduce_mean(tf.maximum(basic_loss, 0.0), 0)

    return loss
```

https://arxiv.org/pdf/1503.03832.pdf

딥러닝 이후 Face Recognition, 이렇게 했습니다! – ArcFace(2018)

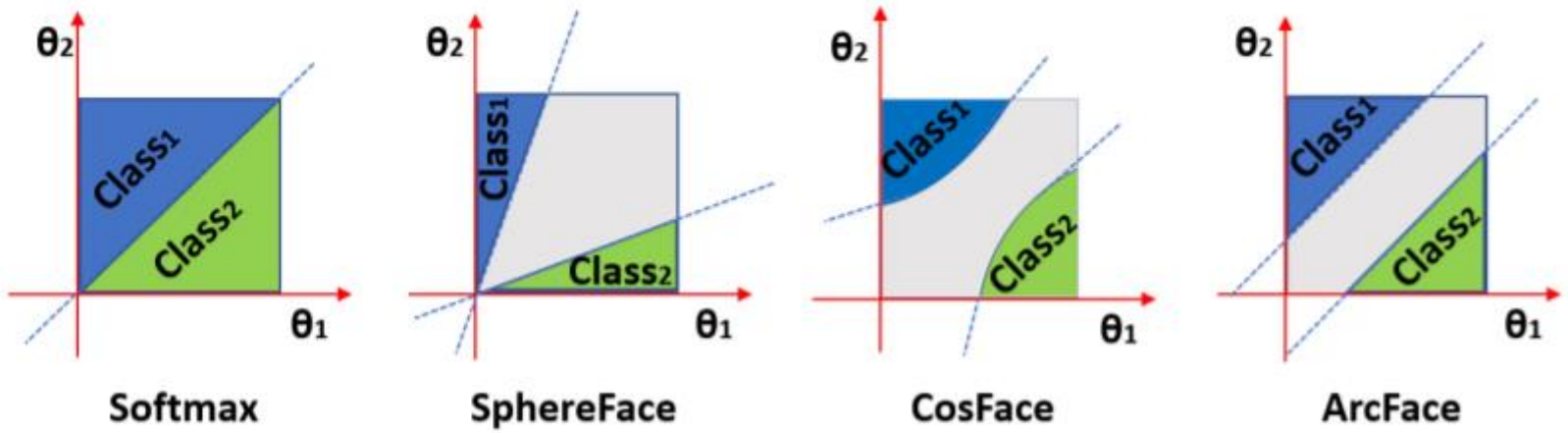


Figure 5. Decision margins of different loss functions under binary classification case. The dashed line represents the decision boundary, and the grey areas are the decision margins.

https://arxiv.org/pdf/1801.07698.pdf

딥러닝 이후 Face Recognition, 벌써 이렇게 잘 된다고?

사실은 공개 데이터셋에만 잘 되는 것일수도 있습니다.
→ 실제로 적용하실 때는 잘 되지 않을 수도 있습니다.

Method	Public. Time	Loss	Architecture	Number of Networks	Training Set	Accuracy±Std(%)
DeepFace [195]	2014	softmax	Alexnet	3	Facebook (4.4M,4K)	97.35±0.25
DeepID2 [187]	2014	contrastive loss	Alexnet	25	CelebFaces+ (0.2M,10K)	99.15±0.13
DeepID3 [188]	2015	contrastive loss	VGGNet-10	50	CelebFaces+ (0.2M,10K)	99.53±0.10
FaceNet [176]	2015	triplet loss	GoogleNet-24	1	Google (500M,10M)	99.63±0.09
Baidu [124]	2015	triplet loss	CNN-9	10	Baidu (1.2M,18K)	99.77
VGGface [149]	2015	triplet loss	VGGNet-16	1	VGGface (2.6M,2.6K)	98.95
light-CNN [225]	2015	softmax	light CNN	1	MS-Celeb-1M (8.4M,100K)	98.8
Center Loss [218]	2016	center loss	Lenet+-7	1	CASIA-WebFace, CACD2000, Celebrity+ (0.7M,17K)	99.28
L-softmax [126]	2016	L-softmax	VGGNet-18	1	CASIA-WebFace (0.49M,10K)	98.71
Range Loss [261]	2016	range loss	VGGNet-16	1	MS-Celeb-1M, CASIA-WebFace (5M,100K)	99.52
L2-softmax [157]	2017	L2-softmax	ResNet-101	1	MS-Celeb-1M (3.7M,58K)	99.78
Normface [206]	2017	contrastive loss	ResNet-28	1	CASIA-WebFace (0.49M,10K)	99.19
CoCo loss [130]	2017	CoCo loss	-	1	MS-Celeb-1M (3M,80K)	99.86
vMF loss [75]	2017	vMF loss	ResNet-27	1	MS-Celeb-1M (4.6M,60K)	99.58
Marginal Loss [43]	2017	marginal loss	ResNet-27	1	MS-Celeb-1M (4M,80K)	99.48
SphereFace [125]	2017	A-softmax	ResNet-64	1	CASIA-WebFace (0.49M,10K)	99.42
CCL [155]	2018	center invariant loss	ResNet-27	1	CASIA-WebFace (0.49M,10K)	99.12
AMS loss [205]	2018	AMS loss	ResNet-20	1	CASIA-WebFace (0.49M,10K)	99.12
Cosface [207]	2018	cosface	ResNet-64	1	CASIA-WebFace (0.49M,10K)	99.33
Arcface [42]	2018	arcface	ResNet-100	1	MS-Celeb-1M (3.8M,85K)	99.83
Ring loss [272]	2018	Ring loss	ResNet-64	1	MS-Celeb-1M (3.5M,31K)	99.50

https://arxiv.org/pdf/1804.06655.pdf

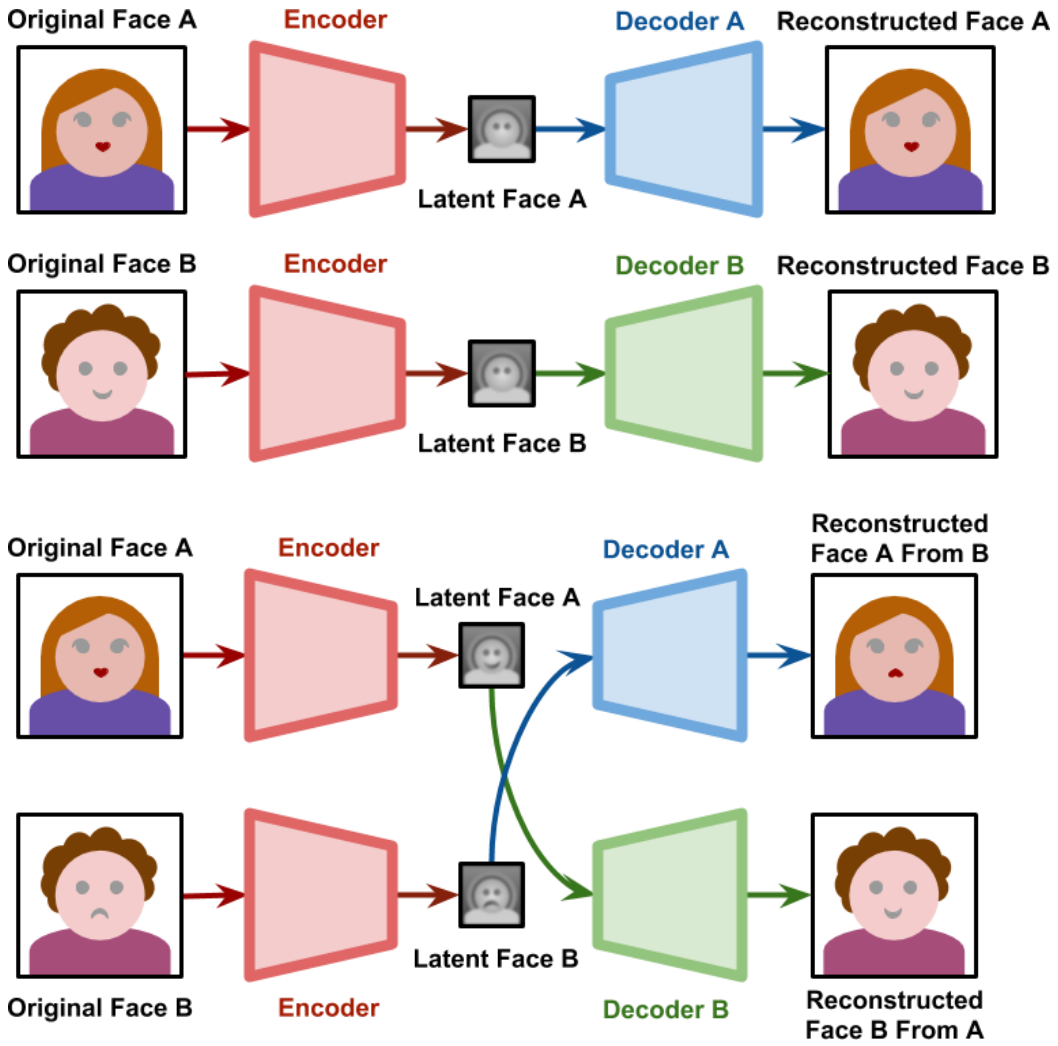
Face Recognition, 입맛에 맞는 Repo를 골라서 적용할 수 있다!

face_recognition - https://github.com/ageitgey/face_recognition
Vggface2 - https://github.com/ox-vgg/vgg_face2
Facenet - https://github.com/davidsandberg/facenet
OpenFace - https://cmusatyalab.github.io/openface/
Sphereface - https://github.com/wy1iu/sphereface
Arcface - https://github.com/deepinsight/insightface

너무 많다...

Bonus : Face Swap - Deepfakes

Bonus : Face Swap - Deepfakes

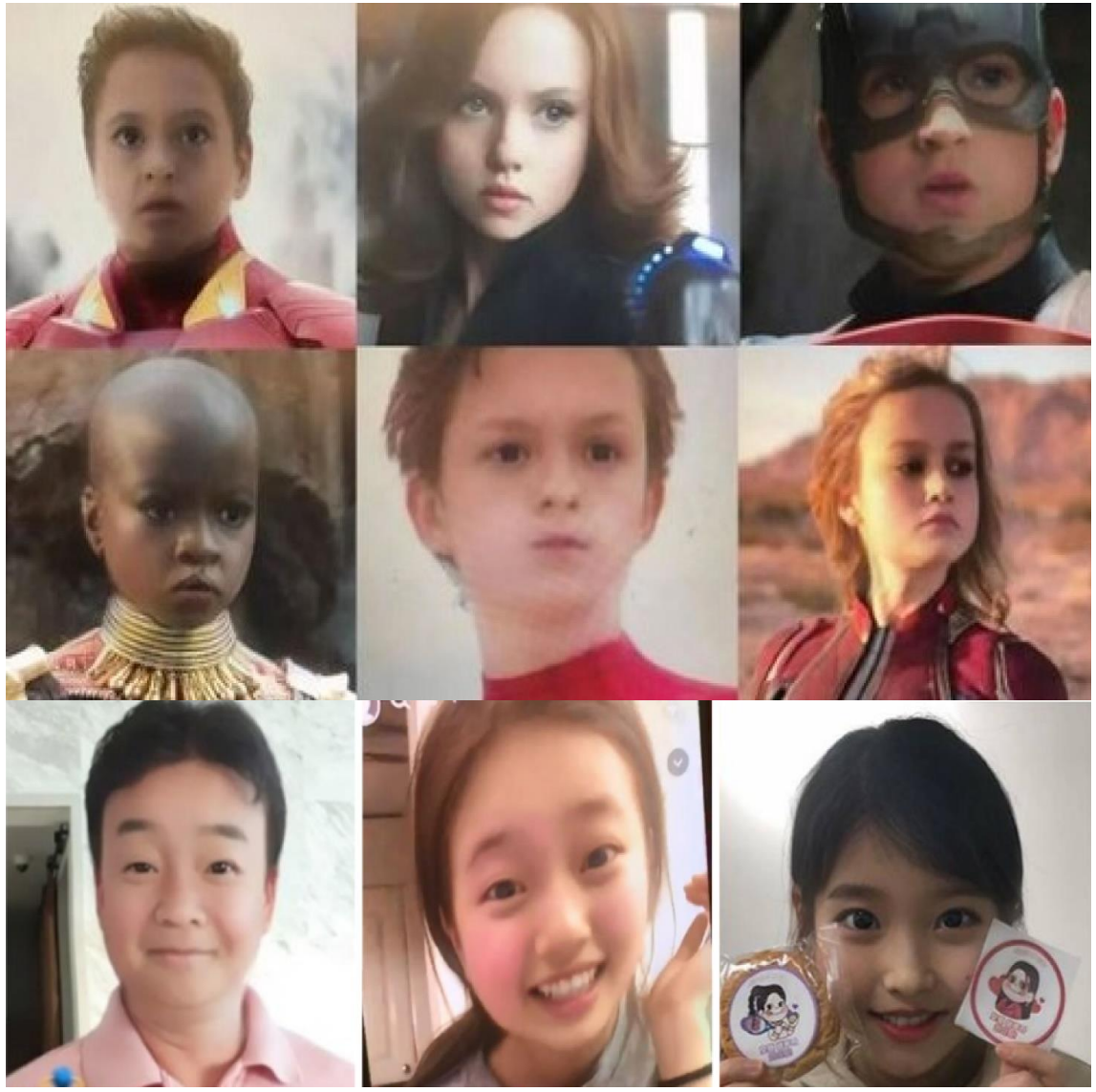


https://www.alanzucconi.com/2018/03/14/understanding-the-technology-behind-deepfakes/

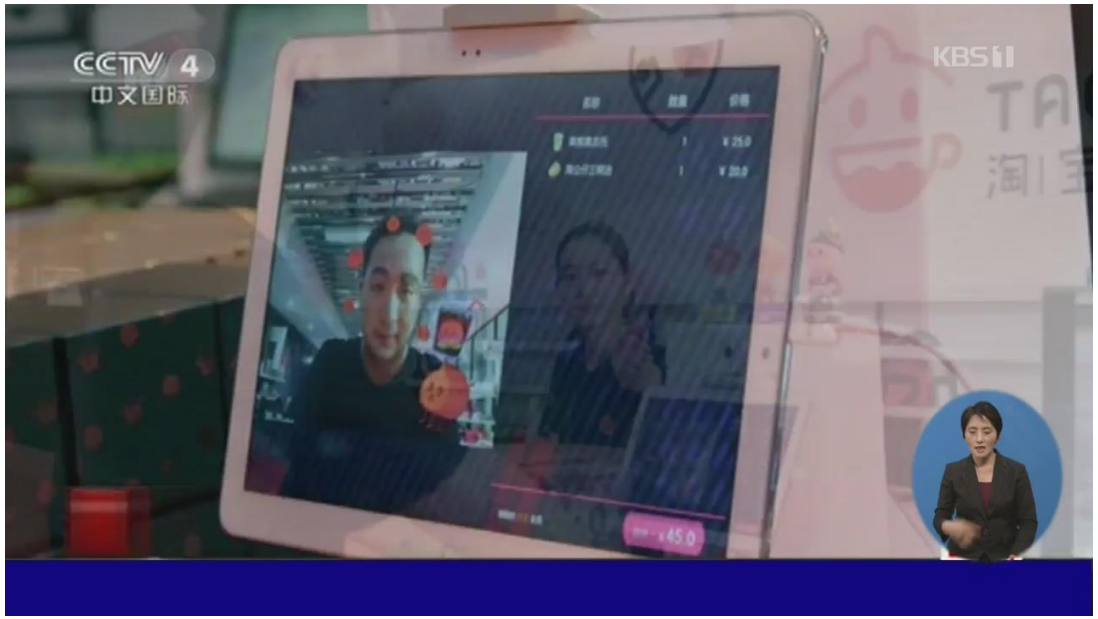
그 밖에 얼굴로 할 수 있는 것들?

요즘에는 얼굴로 뭐하지? – Snow(Face Swap)

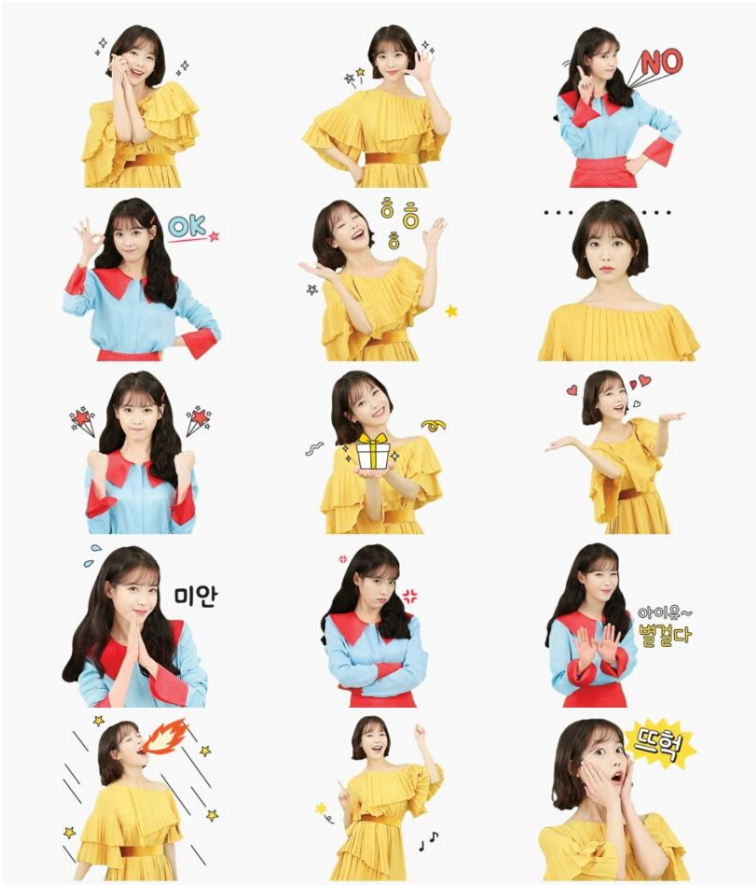
요즘에는 얼굴로 뭐하지? – Snapchat 아기 필터(Style Transfer?)



요즘에는 얼굴로 뭐하지? – 중국 얼굴인식 결제..(Face Recognition)



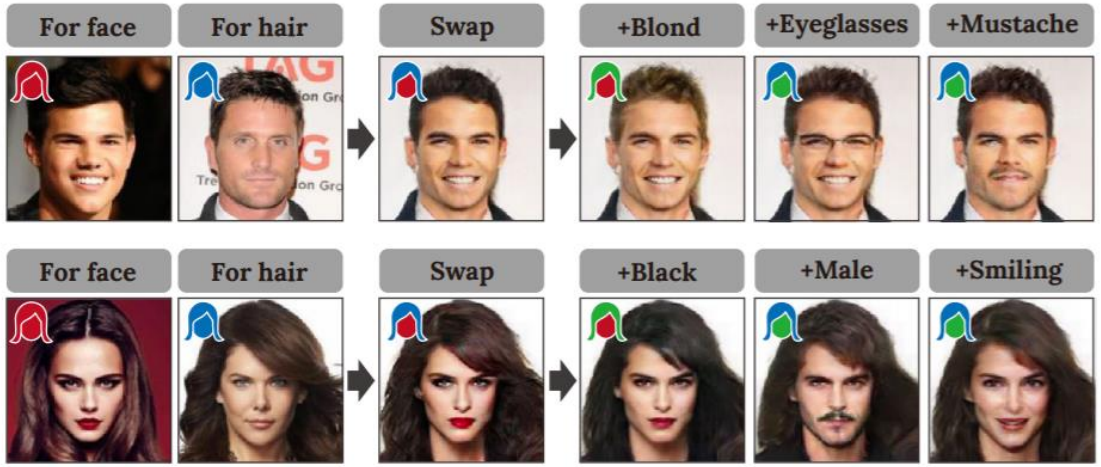
미래에 얼굴로 뭐하지? – Custom 캐릭터, 이모티콘



Further Issues?

TODO

- Few Shot Learning(학습 데이터가 적을 때에도 잘 되게 하고 싶다...)
- Tiny Face(검출하려는 얼굴 크기가 너무 작다?)
- Real-time(or Lightweight) Face Image Processing
  (느리면 답답... or 역시 모바일 디바이스에서 돌아가려면 가벼워야지!)
- Profile Face Detection(옆모습)
- 머리카락 Segmentation(일렉트릭 정화?)
- 기타 등등...



감사합니다.

wogurlee@naver.com