

Robust and Sparse Head Pose Estimation without Face Localization

Menti : 김수정, 한경탁

Industry Mentor : Julius Sumantri

Academic Mentor: 이상철 교수님, 홍성은 교수님



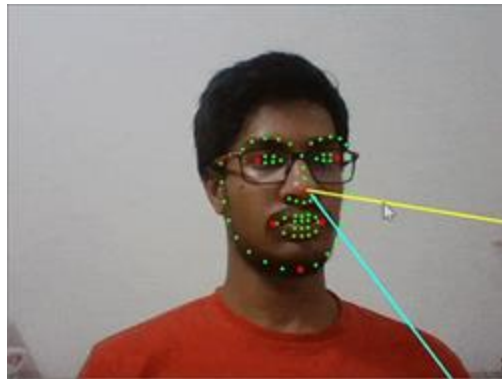
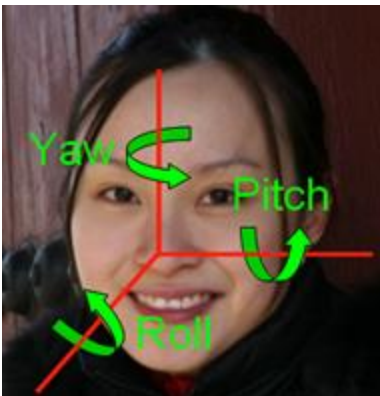
목차

- 프로젝트 개요
- 프로젝트 목표
- 프로젝트 진행 과정
 - Base model, 실험, 데이터
 - 마스크 데이터 생성
 - 다양한 모델 훈련 및 정량/정성적 실험 결과 비교
 - 성능 고도화를 위한 방법 분석 및 구현 – 각 모델 별 성능 비교
- 개선 사항 및 최종 프로젝트 보고

프로젝트 개요

Head Pose Estimation :

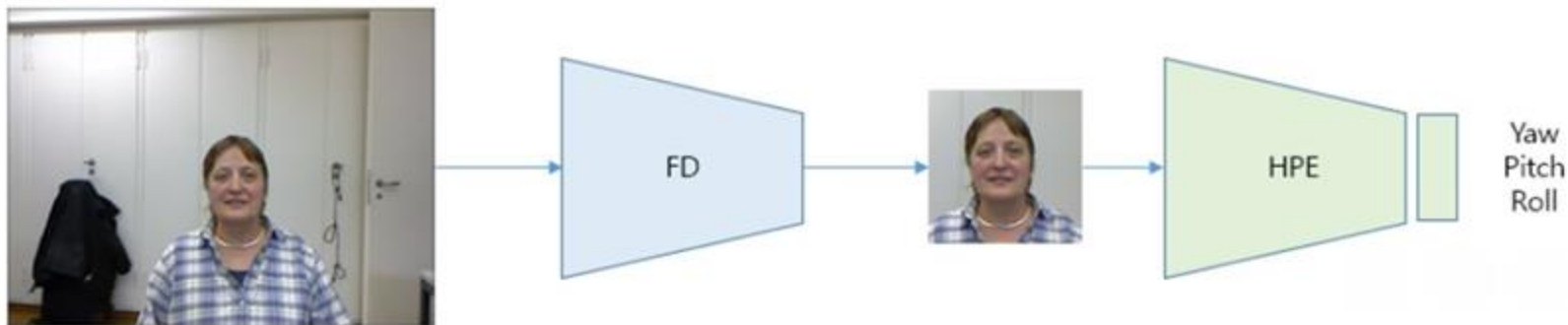
- 사람의 머리 자세를 추정하는 것은 얼굴 분석에 있어 중요한 부분이며 다양한 분야에서 활용할수 있는 응용 분야가 많은 중요한 문제.
- 대표적 응용 분야로는 운전자보조, 가상 현실(augmented reality), HCI 등.
- Yaw, pitch, and roll 예측



프로젝트 개요

연구의 필요성:

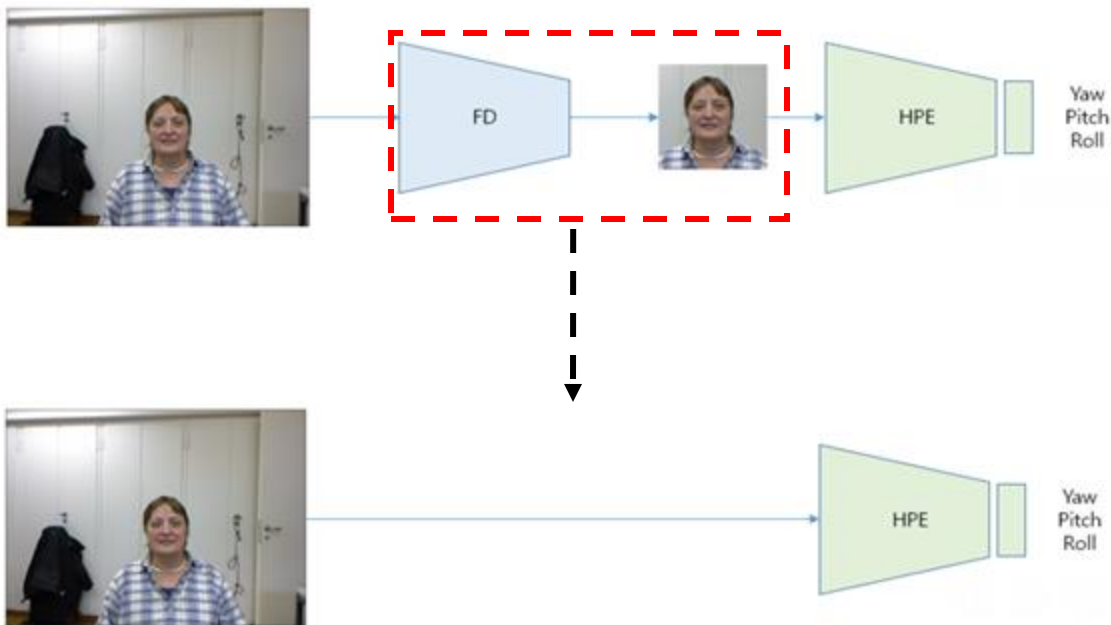
- 최신의 HPE는 sparse input 이미지에서 좋은 성능을 내지 못하며 Face Detection (FD)분야에 많은 dependency가 있다.
- 대부분의 방법론들은 FD 와 Head Pose Estimation (HPE)를 수행하는 2단계로 진행이 된다.
- 마스크 와 같은 face occlusion 발생시 성능 하락이 많이 발생과 함께 필요성 급증.



프로젝트 목표

1단계 목표: Sparse 이미지에서 FD없이 HPE 가능한 end-to-end 모델 개발

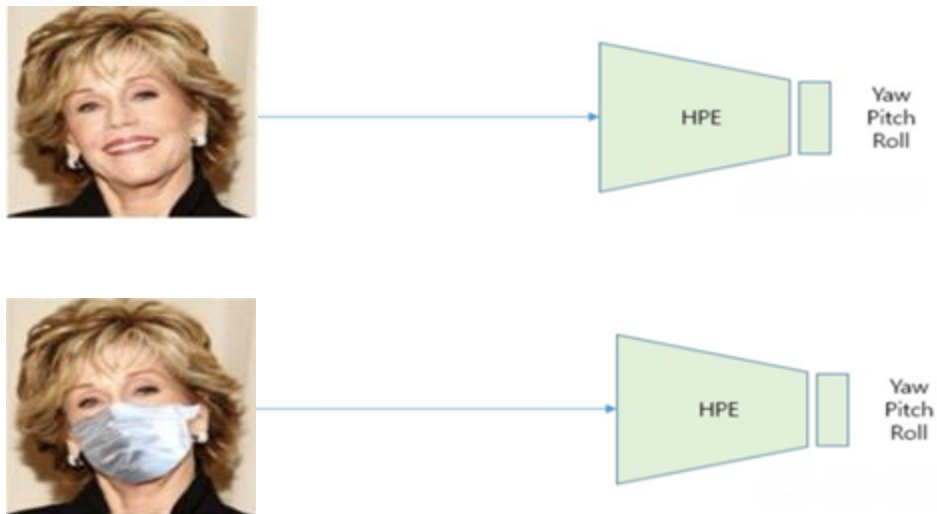
- FD 포함한 네트워크와 비슷한 성능 유지



프로젝트 목표

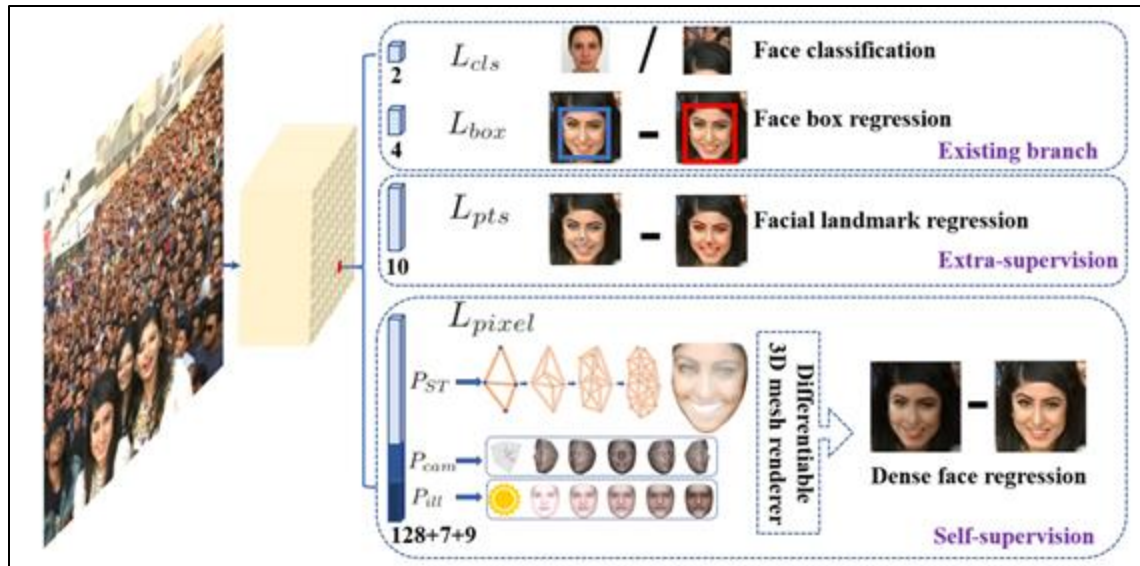
2단계 목표: Face occlusion (mask)에 강인한 HPE 개발.

- Masked HPE 성능 ~ Non-masked HPE 성능



프로젝트 진행 과정 : Based model

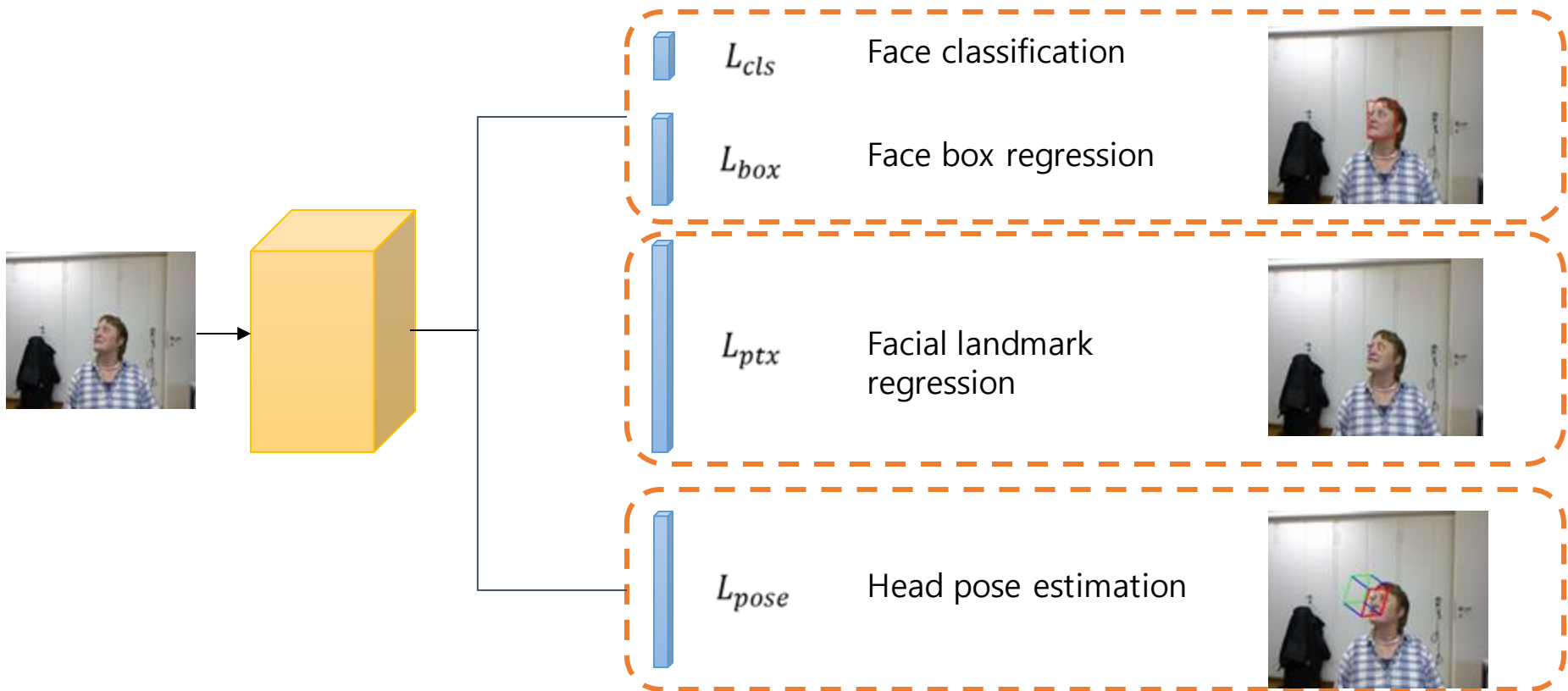
Based model: **Retina-face**



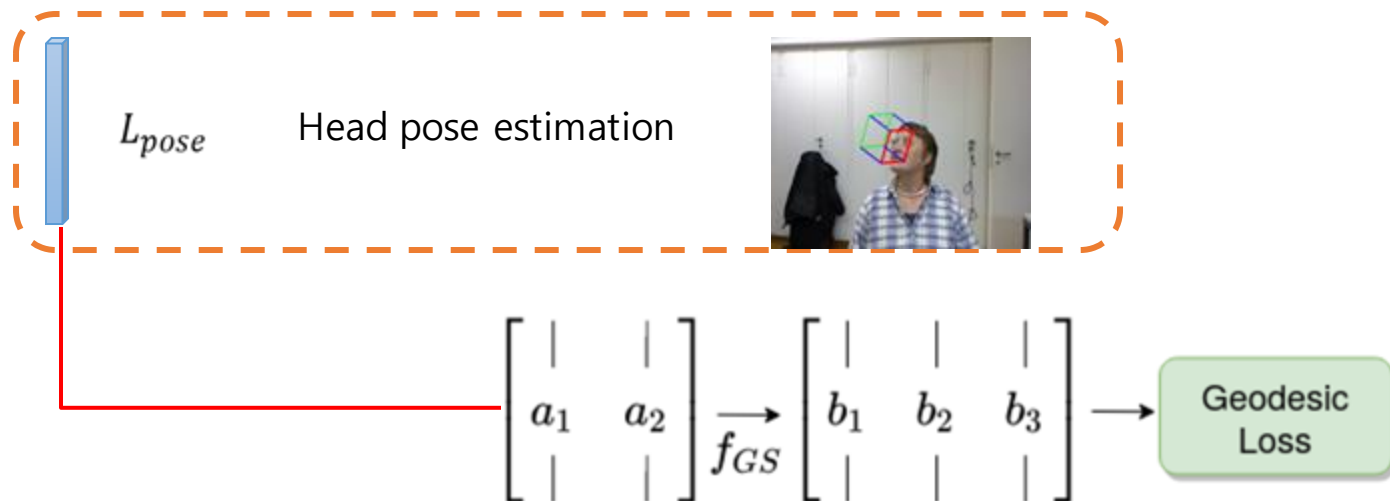
Single stage 기반 face detection 모델

1. Face Detection - 수행
2. Bounding Box - 수행
3. Landmark 수행.
4. Dense face regression 수행

프로젝트 진행 과정 : Architecture



프로젝트 진행 과정 : Architecture



➤ Backbone

- Backbone : **Mobilenet, ResNet50**
- **Pre-processing**을 통한 Training 데이터 ground truth 생성 *i.e.*, face detection, bbox, landmark + 기존 HPE label.
- **Newly training** backbone, pyramid, context module, each head.

프로젝트 진행 과정 : 데이터 세트

데이터 구성 및 실험 구성.

- BIWI – sparse
 - 개수: 15667
 - 훈련 : 70% 검증: 30%
- AFLW2000 – head centralized
 - 개수 : 2000
 - 100% 검증 데이터
- 300W_LP
 - 개수: 61,225
 - AFW: 5,207, IBUG: 1,786, LFPW 16,556, LFPW 37,676



BIWI



AFLW



300W_LP

프로젝트 진행 과정 : Based model 테스트

Model	BIWI	AFLW 2000
BIWI로 학습 전	30.92	47.56
BIWI로 학습 후	1.46	17.98

BIWI - 70% Train, 30%
Test

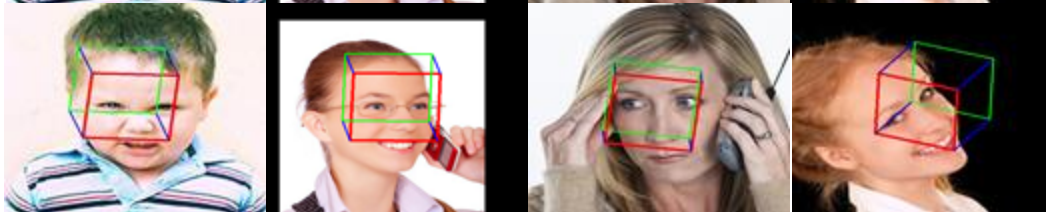
프로젝트 진행 과정 : Based model

검증 데이터 셋 - AFLW 2000

학습 전



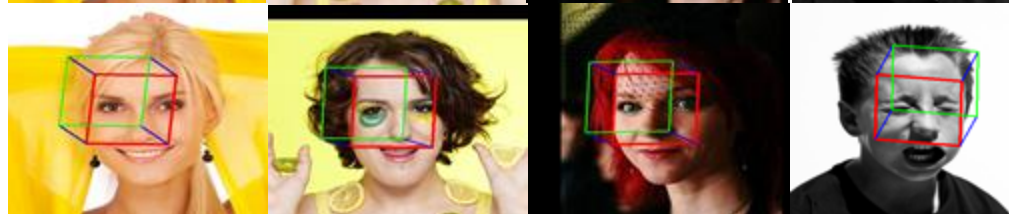
학습 후



학습 전



학습 후



프로젝트 진행 과정 : Based model

검증 데이터 셋 - BIWI (30%)

학습 전



학습 후



학습 전

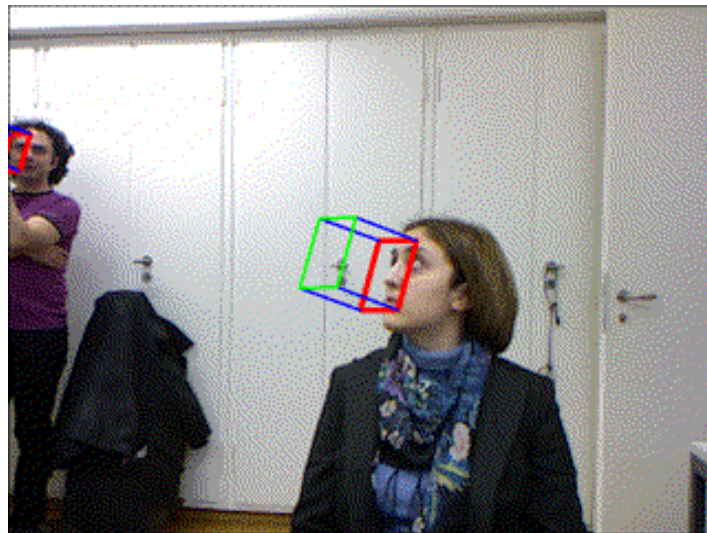
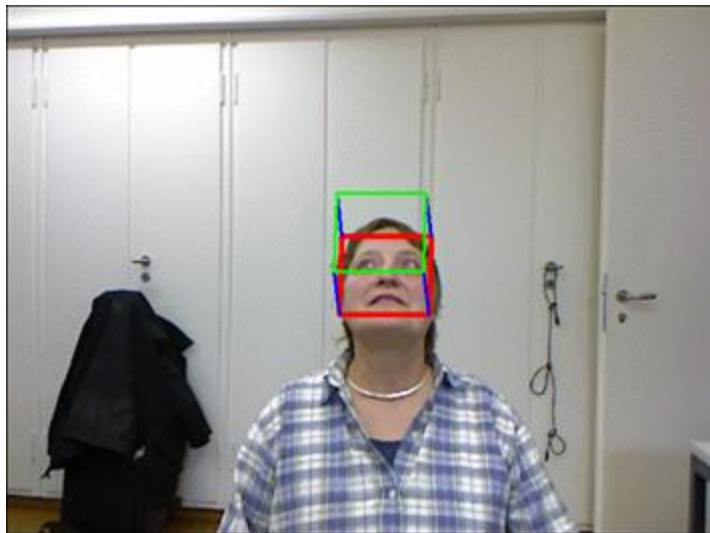


학습 후



프로젝트 진행 과정 : Based model

검증 데이터 셋 - BIWI (30%)



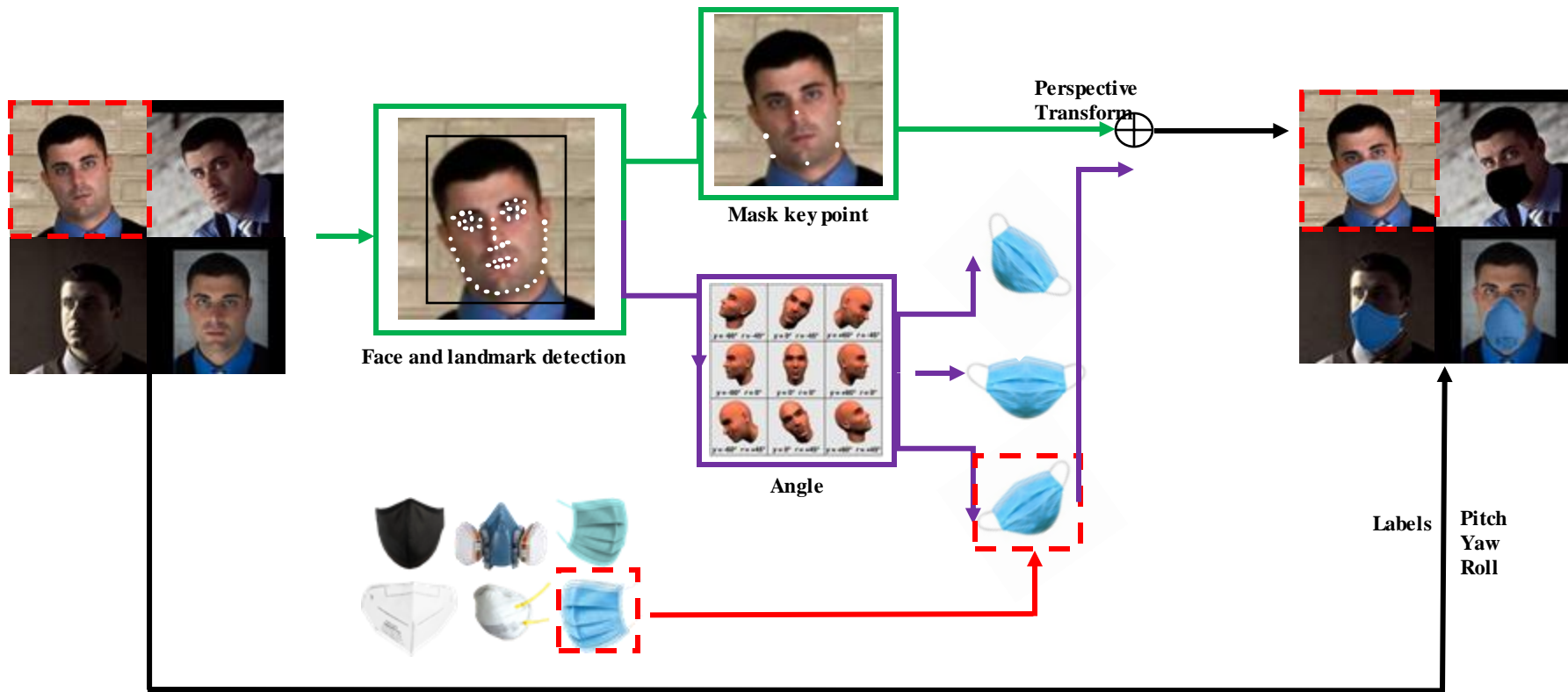
프로젝트 진행 과정 : 마스크 데이터 생성



→
Synthesize

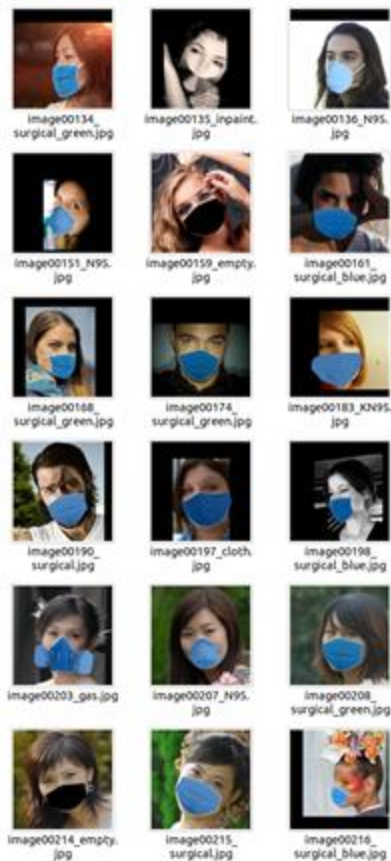


프로젝트 진행 과정 : 마스크 데이터 생성



프로젝트 진행 과정 : 마스크 데이터 생성

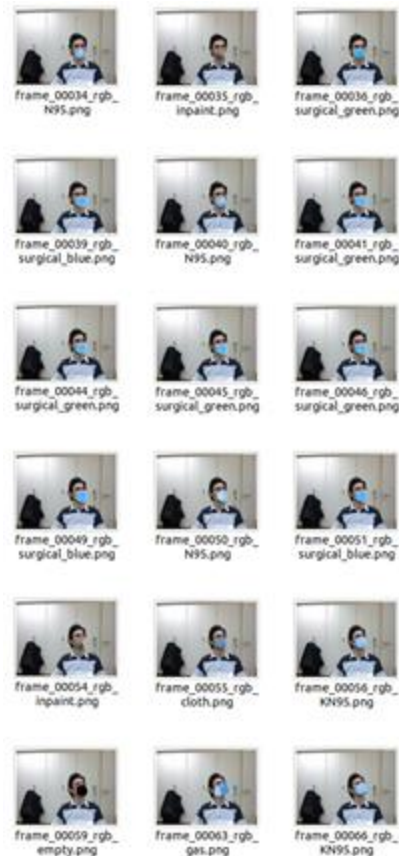
ALFW2000: 1509



300W_LP : 36483 EA



BIWI: 11514 EA



프로젝트 진행 과정 : 마스크 데이터 활용 모델 훈련 / 검증

모델 학습 : - masked data에 대한 robustness 한 모델을 만들기 위해

- **BIWI + 300W_LP 활용 모델 학습.**
 - > Multi-task에서 pose loss trade off tuning [0.2 ~ 0.8]
 - > 훈련된 모델 활용 모델 검증 (**BIWI | BIWI_masked | 300W_LP_masked | AFLW | AFLW_masked**)
- **BIWI + 300W_LP + masked** 데이터 활용 모델 학습
 - > Multi-task에서 pose loss trade off tuning [0.2 ~ 0.8] 및 Hyper parameter tuning.
 - > 훈련된 모델 활용 모델 검증 (**BIWI | BIWI_masked | AFLW | AFLW_masked**)

프로젝트 진행 과정 : 실험 결과

Model1 - Train: BIWI, 300W_LP

Dataset	YAW	PITCH	ROLL	MAE
BIWI	7.50	5.34	4.78	5.87
BIWI_masked	13.95	16.74	8.64	13.11
300W_LP	-	-	-	-
300W_LP_masked	38.30	35.34	8.26	27.30
AFLW2000	27.50	32.22	13.39	24.37
AFLW2000_masked	23.10	28.45	12.55	21.37

- 300W_LP – train 사용.

프로젝트 진행 과정 : 실험 결과

Model1 - 학습 결과 BIWI 시각화

face
detection



head pose
estimation



GT



프로젝트 진행 과정 : 실험 결과

Model1 - 학습 결과 BIWI_masked 시각화

face
detection



head pose
estimation



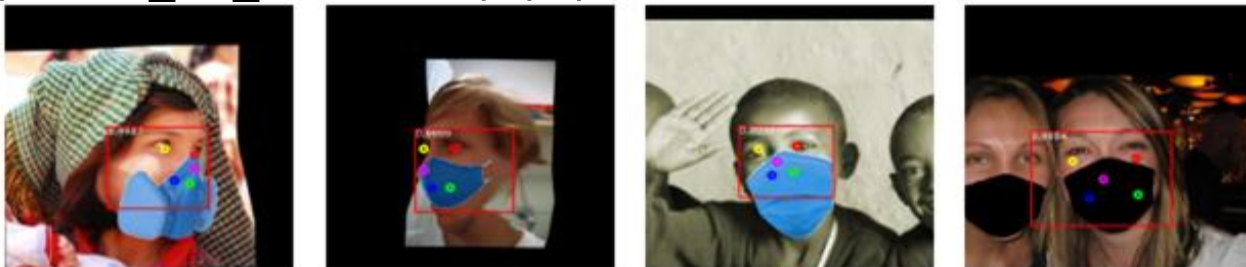
GT



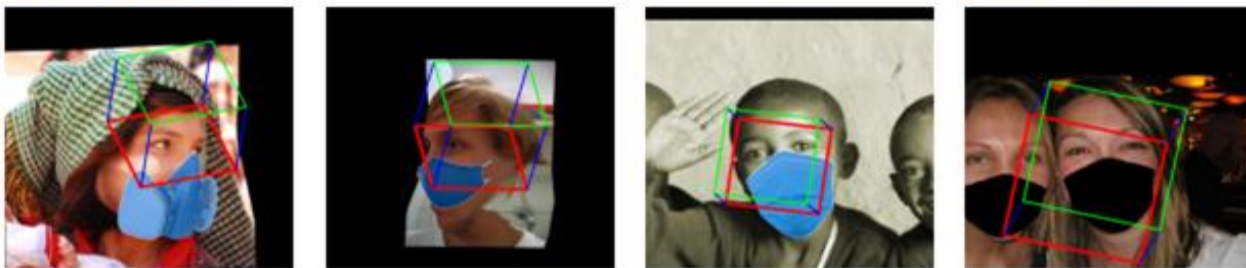
프로젝트 진행 과정 : 실험 결과

Model1 - 학습 결과 300W_LP_masked 시각화

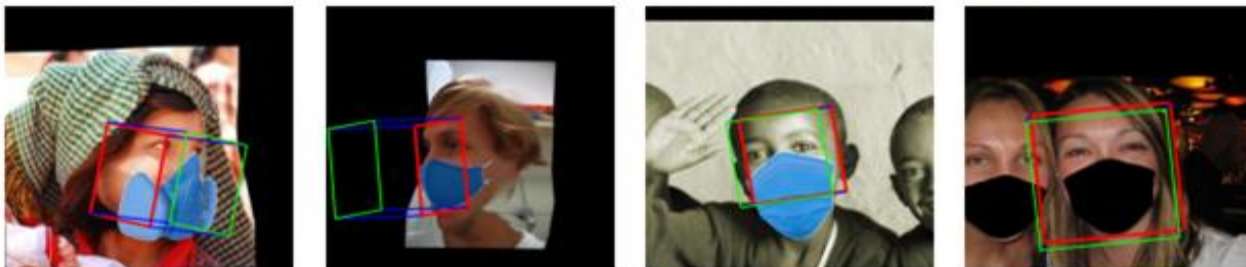
face
detection



head pose
estimation



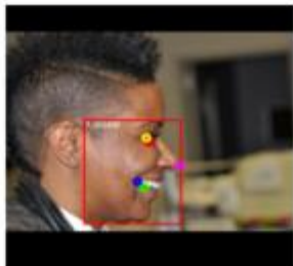
GT



프로젝트 진행 과정 : 실험 결과

Model1 - 학습 결과 AFLW2000 시각화

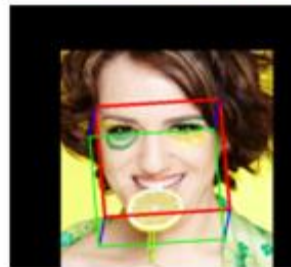
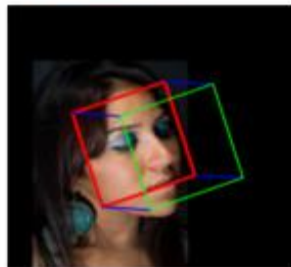
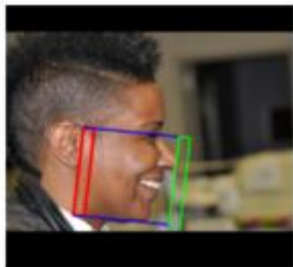
face
detection



head pose
estimation



GT



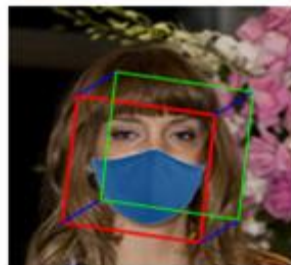
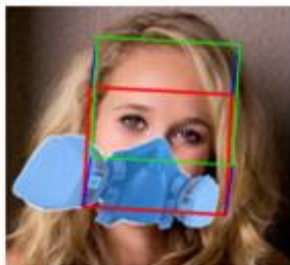
프로젝트 진행 과정 : 실험 결과

Model1 - 학습 결과 AFLW2000_masked 시각화

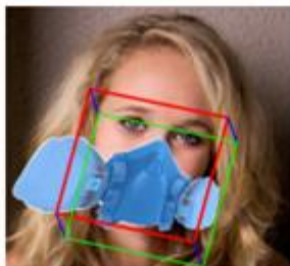
face
detection



head pose
estimation



GT



프로젝트 진행 과정 : 실험 결과

Model2 - Train: BIWI, BIWI_masked, 300W_LP, 300W_LP_masked

Dataset	YAW	PITCH	ROLL	MAE
BIWI	5.58	3.93	4.46	4.66
BIWI_masked	5.41	4.74	3.39	4.51
300W_LP	-	-	-	-
300W_LP_masked	-	-	-	-
AFLW2000	27.45	28.52	14.73	23.57
AFLW2000_masked	21.26	22.48	10.84	18.19

300W_LP – train 사용.
300W_LP_masked – train 사용

프로젝트 진행 과정 : 실험 결과

Model2 - 학습 결과 BIWI 시각화

face
detection



head pose
estimation



GT



프로젝트 진행 과정 : 실험 결과

Model2 - 학습 결과 BIWI_masked 시각화

face
detection



head pose
estimation



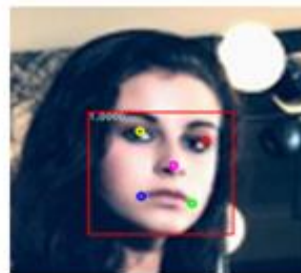
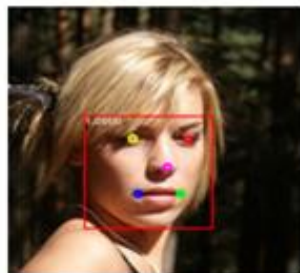
GT



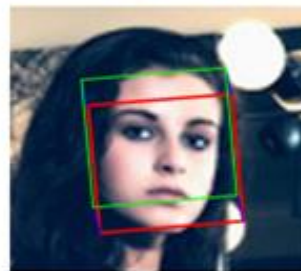
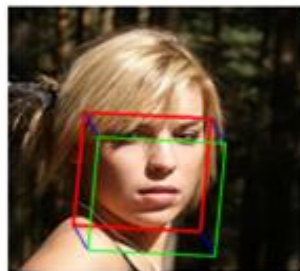
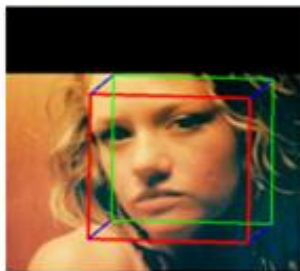
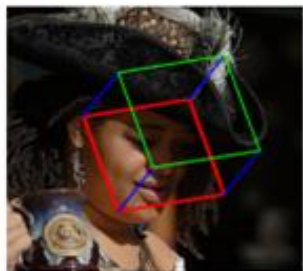
프로젝트 진행 과정 : 실험 결과

Model2 - 학습 결과 AFLW2000 시각화

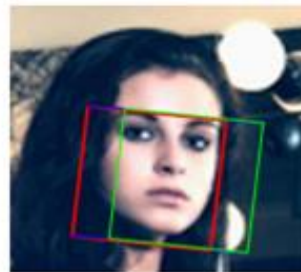
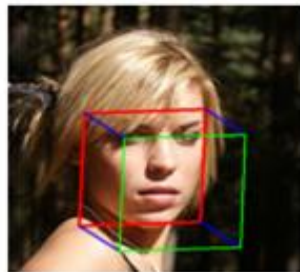
face
detection



head pose
estimation



GT



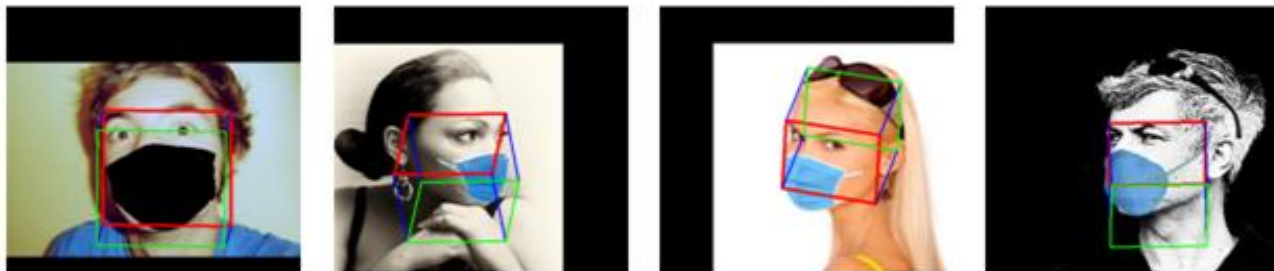
프로젝트 진행 과정 : 실험 결과

Model2 - 학습 결과 AFLW2000_masked 시각화

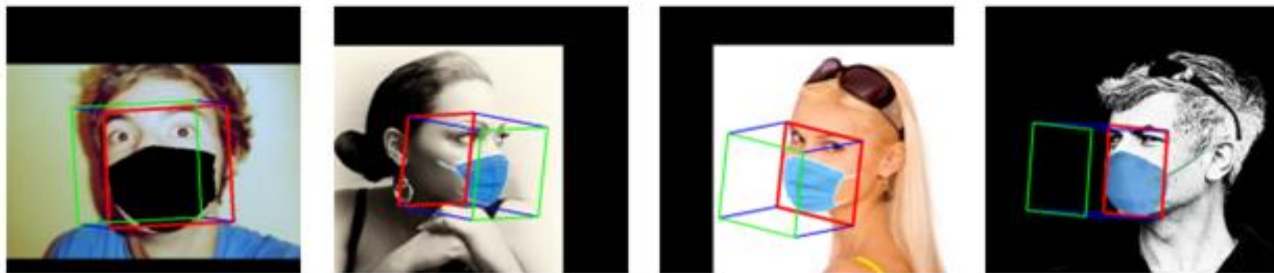
face
detection



head pose
estimation



GT

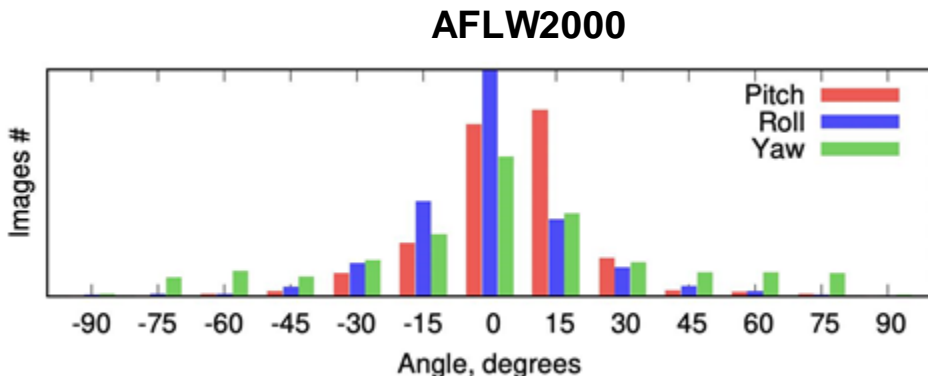
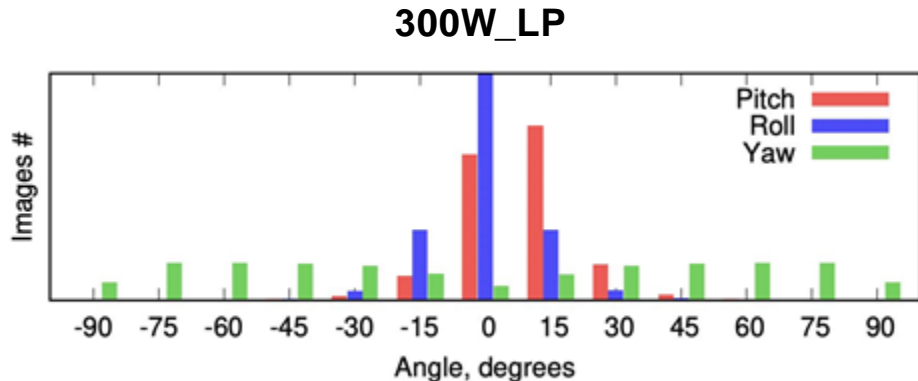


프로젝트 진행 과정 : 성능 고도화를 위한 분석

Public 데이터셋에서 angles bias가 존재 [그림 1, 그림 2]

다양한 [Yaw, Pitch, Roll] 분포를 가진 데이터에서는 낮은 성능을 보이며 AFLW와 같은 데이터에서도 낮은 성능.

- 문제점 1 – 훈련 데이터의 분포상 다양한 각도를 학습하기 어려운 단점이 있음.
- 문제점 2 – 기존에는 데이터 flipping 으로 다양성을 주려는 노력이 있었음. [yaw \rightarrow -yaw 등]
- 문제점 3 – Rotation을 HPE에 적용시 ground truth 라벨 자체를 수정해야하는 문제점이 있음.

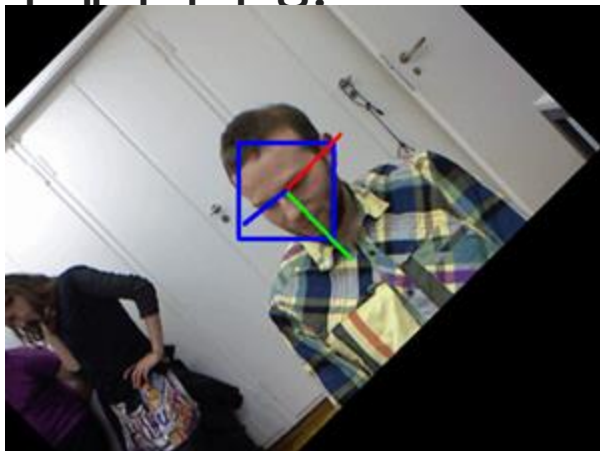
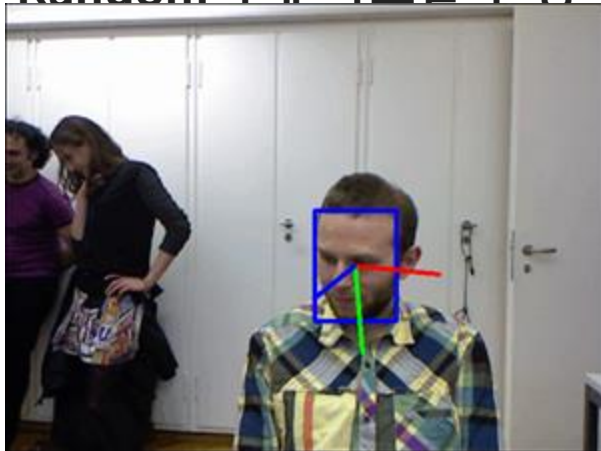


프로젝트 진행 과정 : 방법론 탐색 및 구현

Head Pose Estimation에서 다양한 각도의 **rotation**을 활용한 **augmentation** 논문 탐색.

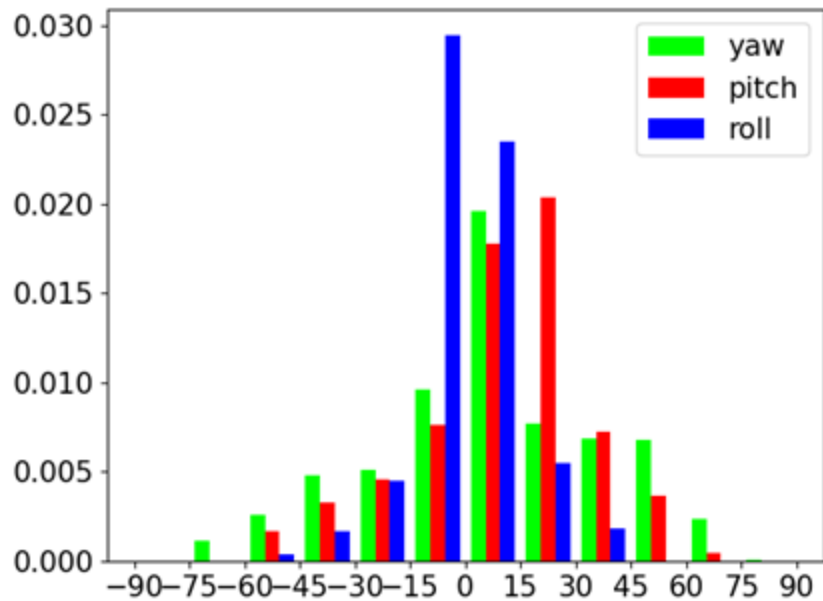
Sheka et al. 제안한 특정 각도로 이미지가 **rotation**시 변하는 **[yaw, pitch, roll]**을 계산 하는 방법을 제안하며 해당 데이터를 활용 모델 검증한 논문을 구현.

Random하게 각도를 선정하여 데이터 생성.



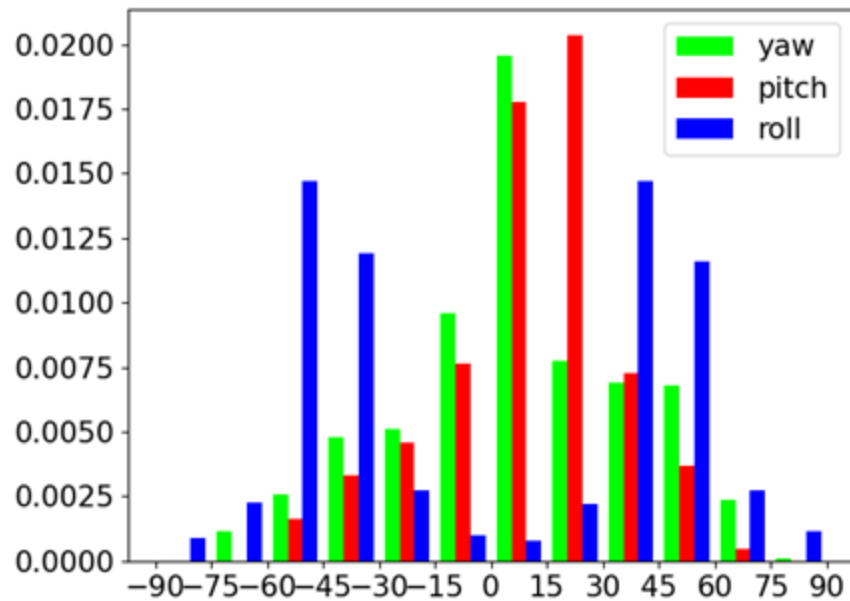
프로젝트 진행 과정 : Rotation 적용

Original BIWI



Angles, Degrees

Rotated BIWI



Angles, Degrees

프로젝트 진행 과정 : 실험 결과

Model3 - Train: BIWI, BIWI_rotated

Dataset	YAW	PITCH	ROLL	MAE
BIWI	5.43	4.17	3.74	4.45
BIWI_masked	4.45	2.97	2.90	3.44
300W_LP	-	-	-	-
300W_LP_masked	-	-	-	-
AFLW2000	23.82	12.70	13.95	16.82
AFLW2000_masked	20.24	13.08	10.22	14.51

프로젝트 진행 과정 : 실험 결과

Model3 - 학습 결과 BIWI_masked 시각화

face
detection



head pose
estimation

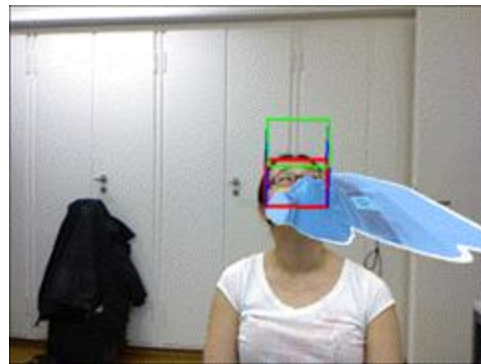
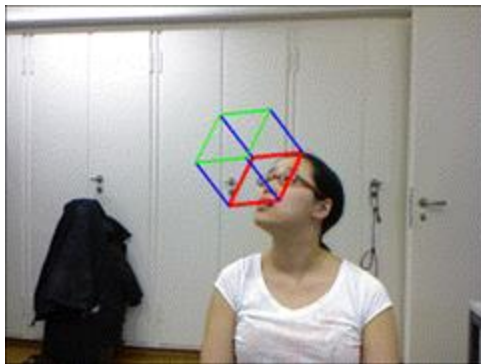


GT



프로젝트 진행 과정 : 실험 결과

Model3 - 학습 결과 BIWI_masked 시각화



프로젝트 진행 과정 : 성능 비교

각 모델 별 성능 비교

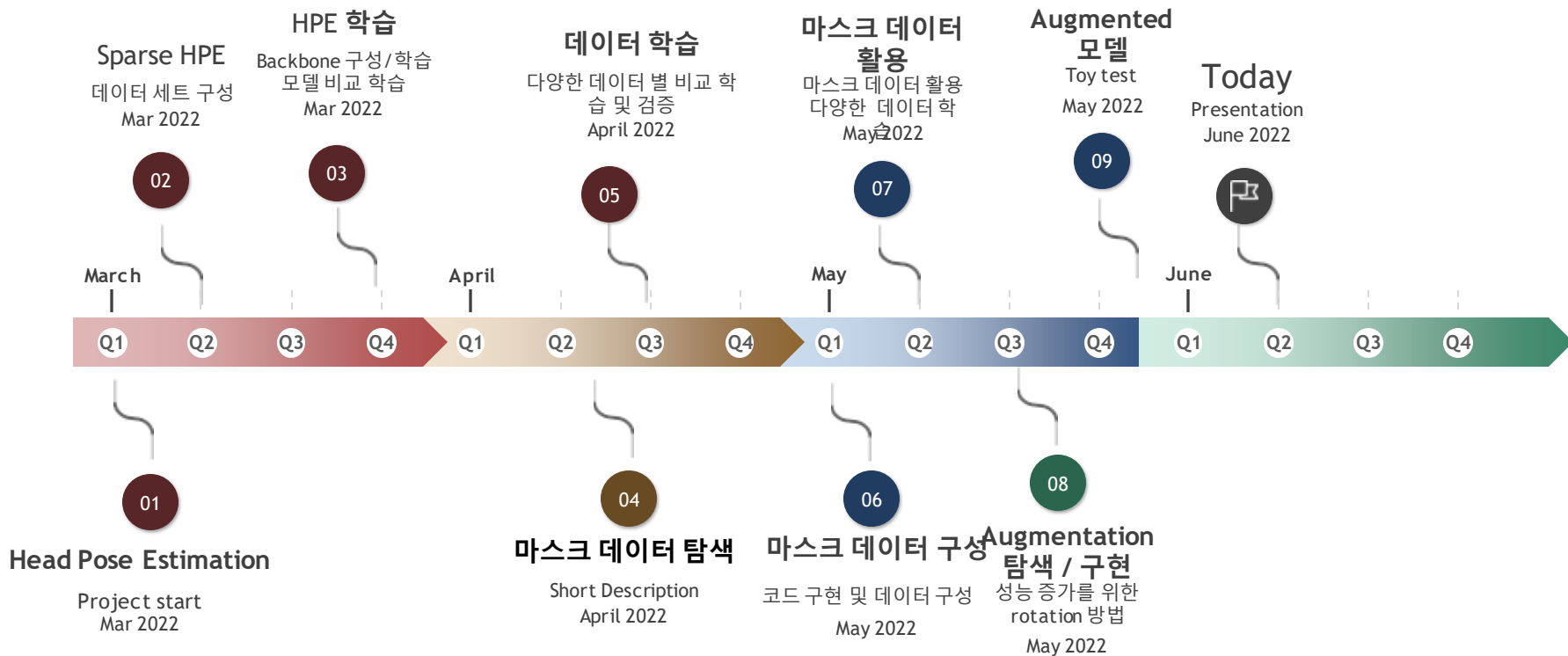
Model	Dataset	YAW	PITCH	ROLL	MAE
Model1	BIWI	7.50	5.34	4.78	5.87
	BIWI_masked	13.95	16.74	8.64	13.11
Model2	BIWI	5.58	3.93	4.46	4.66
	BIWI_masked	5.41	4.74	3.39	4.51
Model3	BIWI	5.43	4.17	3.74	4.45 (-1.45)
	BIWI_masked	4.45	2.97	2.90	3.44 (-9.67)

프로젝트 진행 과정 : 성능 비교

각 모델 별 성능 비교

Model	Dataset	YAW	PITCH	ROLL	MAE
Model1	AFLW2000	27.50	32.22	13.39	24.37
	AFLW2000_masked	23.10	28.45	12.55	21.37
Model2	AFLW2000	27.45	28.52	14.73	23.57
	AFLW2000_masked	21.26	22.48	10.84	18.19
Model3	AFLW2000	23.82	12.70	13.95	16.82 (-7.55)
	AFLW2000_masked	20.24	13.08	10.22	14.51(-6.86)

프로젝트 : history



프로젝트: 목표 달성 및 향후 방향

- Single stage Head Pose Estimation의 성능 개선 (MAE <= 4.5) – **4.45 달성**
- Occlusion 상황에서의 성능 고도화 (MAE <= 5.0) – **3.45 달성**
- HPE 마스크 관련 IEIE 대한전자공학회 하계학술대회 논문 게재

마스크가 씌어진 헤드 포즈 추정 데이터 생성 및 검증

원작: 이호준, 홍승은
 인제대학교 전기컴퓨터공학부
 e-mail: hoo0227@inje.ac.kr, hoojunlee1999@gmail.com, oosung@inje.ac.kr
 Data Generation and Validation for Masked Head Pose Estimation
 Kyoungtaek Han, Hojun Lee and Sungeun Hong
 Department of Electrical and Computer Engineering
 Inha University

Abstract

After the pandemic of the coronavirus disease (COVID-19), many changes have taken place in the field of artificial intelligence. In particular, due to the demand for assets that increased the demand for face recognition and detection in occlusion such as mask. As a strong attention, internationally, various public datasets were used in the field of face recognition, but there were not found in the field of head pose estimation. In this paper, we provide new masked head pose estimation benchmark datasets that can be widely used in validation of occluded head pose method or training occluded situation.

1. 서론

2020년 초 코로나바이러스의 유행과 함께 코로나 바이러스는 곧 세계를 휩쓸고 해외여행금지, 휴교, 휴업 등 다양한 사회적 생활의 변화가 일어났다. 이에 따라 세계보건기구(WHO)와 질병관리청(CDC)에서는 마스크 착용을 통한 코로나바이러스의 유행을 예방하기 위해 마스크 착용을 권고하였다. 이러한 결과, 마스크 착용이 헤드 포즈 추정에 미치는 영향에 대한 연구가 진행되고 있다.

1.1. 연구 배경

2020년 초, 세계보건기구(WHO)에서는 코로나바이러스의 유행을 예방하기 위해 마스크 착용을 권고하였다. 이에 따라 세계보건기구(WHO)와 질병관리청(CDC)에서는 마스크 착용을 통한 코로나바이러스의 유행을 예방하기 위해 마스크 착용을 권고하였다. 이러한 결과, 마스크 착용이 헤드 포즈 추정에 미치는 영향에 대한 연구가 진행되고 있다.

1.2. 연구 목적

본 연구의 목적은 마스크 착용이 헤드 포즈 추정에 미치는 영향을 분석하고, 마스크 착용 상황에서의 헤드 포즈 추정 성능을 향상시키는 방법을 제안하는 것이다.

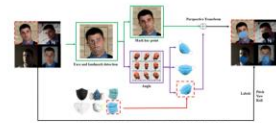


Figure 1. Masked Head Pose Estimation Architecture



Figure 2. Masked Head Pose Estimation Results



Figure 3. Masked Head Pose Estimation Results

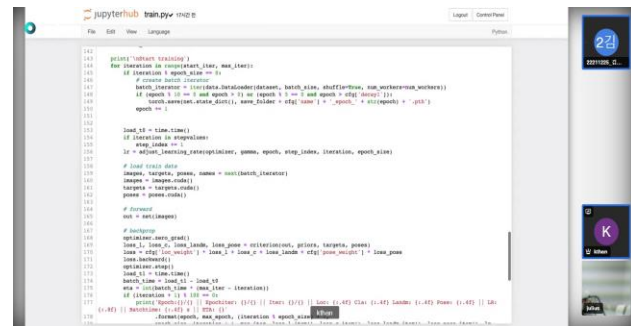
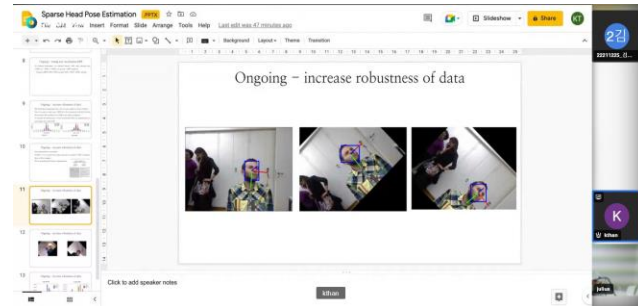


Figure 4. Masked Head Pose Estimation Results

향후 프로젝트

- Single stage Head Pose Estimation의 성능 고도화 ex) AFLW2000, AFLW2000_masked.
- rotation [-45,45] 에서 더 큰 각도 변화 적용, 데이터 구성 및 training 시 augmentation 단계로 발전.
- 훈련 모델 real-time 테스트 및 robustness 테스트 (다양한 각도 및 마스크 etc).
- 프로젝트 지도에 따라서 신규 논문 구상 (Toward Sparse Head Pose Estimation without face

프로젝트 보조 설명 : 멘토와의 정기 회의



프로젝트 보조 설명 : 플랫폼 사용 실적

jupyterhub

Login Control Panel

Files Running Clusters

Select items to perform actions on them.

Upload New

	Name	Last Modified	File size
<input type="checkbox"/>	work		
<input type="checkbox"/>	..	방문 전	
<input type="checkbox"/>	datasets	방문 전	
<input type="checkbox"/>	media	방문 전	
<input type="checkbox"/>	reports	방문 전	
<input type="checkbox"/>	utils	방문 전	
<input checked="" type="checkbox"/>	generation.pynt	방문 전	35.1 KB
<input type="checkbox"/>	requirements.txt	방문 전	0 B

In [16]:

```
for i, entry in enumerate(mask_codes):
    mask_dict = {}
    mask_color = ""
    mask_texture = entry.split("-")[0]
    if len(entry.split("-")) != 2:
        mask_variation = entry.split("-")[1]
        if "2" in mask_variation:
            mask_color = mask_variation
        else:
            mask_texture = mask_variation
    mask_dict["img"] = mask_type
    mask_dict["color"] = mask_color
    mask_dict["texture"] = mask_texture
    img_mask_dict["img"] = mask_texture
    img_mask_dict["img"] = mask_color

# Check if path is file or directory or none
is_directory, is_file, is_other = check_path(args.path)
display_mask(texture)

if is_directory:
    path, dirs, files = os.walk(args.path, topdown=True)
    file_count = len(files)
    dir_count = len(dirs)
    if len(files) > 0:
        print_orderly("Masking image files", n)

# Process files in the directory if any
for f in tqdm(files):
    image_path = path + "/" + f

    write_path = path + " masked"
    if not os.path.exists(write_path):
        os.makedirs(write_path)

    if is_image(image_path):
        # Process if file is image
        if args.verbose:
            str_p = "Processing: " + image_path + args.verbose
            tqdm.write(str_p)

        split_path = f.split("/")
        masked_image, mask_binary_array, original_image = mask_image(
            image_path, args
        )
        for i in range(len(mask_codes)):
            w_path = f
            write_path = f
            split_path[i] = f
            split_path[i] = mask[i]
            split_path[i] = f
            img = masked_image[i]
            cv2.imwrite(write_path, img)

        print_orderly("Masking image directories", n)
```



프로젝트 보조 설명 : 플랫폼 사용 실적

	Name	Last Modified	File size
..		몇 초 전	
data		3분 전	
datasets		4일 전	
layers		5분 전	
model		6분 전	
reports		8일 전	
utils		7분 전	
weights		6분 전	
run.py		2분 전	1.21 KB
training.py		몇 초 전	17.3 KB



```
195 if args.resume_epoch > 0:
196     start_iter = args.resume_epoch * epoch_size
197 else:
198     start_iter = 0
199
200 for iteration in range(start_iter, max_iter):
201     if iteration % epoch_size == 0:
202         batch_iterator = iter(torch.utils.data.DataLoader(
203             dataset_type=Dataset,
204             batch_size=batch_size,
205             shuffle=True,
206             num_workers=1))
207     if (epoch % 10 == 0 and epoch > 0) or (epoch % 5 == 0 and epoch > cfg.decay1):
208         torch.save(net.state_dict(), save_folder + "hpe_mobilenet_0.25*_epoch_" + str(epoch) + ".pth")
209         epoch += 1
210
211 load_t0 = time.time()
212 if epoch % 10 == 0 and ((iteration % epoch_size) + 1) % 10 == 0:
213     print("===== update step index =====")
214     step_index += 1
215
216 lr = adjust_learning_rate(optimizer, gamma, epoch, step_index, iteration, epoch_size)
217
218 # load train data
219 images, gt_mat, .. = next(batch_iterator)
220 # bbox, classification, ldn regressions
221 images = images.cuda()
222 .., hpe = net(images)
223 g_result = []
224 hpe_loss = 0
225 for i in range(1):
226     trade_off = ((i+1) * 0.1) + 0.8
227     pred = hpeutil.compute_rotation_matrix_from_ortho4d(hpe[i])
228     hpe_loss += trade_off * criterion(gt_mat.cuda(), pred)
```



프로젝트 보조 설명 : 플랫폼 사용 실적

```
In [10]: !python3 train.py --dataset BIWI --batch_size 32 --epoch 50 --num_workers 1

Loading Dataset...
BIWI 10149

Save weights at ./weights/BIWI/BIWI_

Start training
/opt/conda/lib/python3.8/site-packages/torch/nn/functional.py:718: UserWarning: Named tensors and all their associated APIs are an experimental feature and subject to change. Please do not use them for anything important until they are released as stable. (Triggered internally at /pytorch/c10/core/TensorImpl.h:1156.)
  return torch.max_pool2d(input, kernel_size, stride, padding, dilation, ceil_mode)
Epoch:1/50 || Epochiter: 100/318 || Iter: 100/15900 || Loc: 0.4238 Cla: 0.4732 Landm: 1.3545 Pose: 0.4357 || LR: 0.00
100000 || Batchtime: 1.0484 s || ETA: 4:36:05
Epoch:1/50 || Epochiter: 200/318 || Iter: 200/15900 || Loc: 0.2013 Cla: 0.4963 Landm: 0.8653 Pose: 0.3771 || LR: 0.00
100000 || Batchtime: 1.0319 s || ETA: 4:30:02
Epoch:1/50 || Epochiter: 300/318 || Iter: 300/15900 || Loc: 0.2623 Cla: 0.4502 Landm: 1.1124 Pose: 0.4561 || LR: 0.00
100000 || Batchtime: 1.0368 s || ETA: 4:29:34
100%[#####] 169/169 [02:49<00:00, 1.00s/it]
Epoch:1/50 || BIWI Validation Errors [Yaw: 28.2986, Pitch: 7.7107, Roll: 9.6093, MAE: 15.2962]
```

```
In [18]: !python3 test.py -m ./weights/BIWI/BIWI_Resnet50_Final.pth --dataset BIWI --num_workers 1

Loading pretrained model from ./weights/BIWI/BIWI_Resnet50_Final.pth
remove prefix 'module.'
Missing keys:0
Unused checkpoint keys:0
Used keys:462
Finished loading model!
0%[#####] 0/85 [00:00<7, 7it/s]/opt/conda/lib/python3.8/site-packages/torch/nn/functional.py:718: UserWarning: Named tensors and all their associated APIs are an experimental feature and subject to change. Please do not use them for anything important until they are released as stable. (Triggered internally at /pytorch/c10/core/TensorImpl.h:1156.)
  return torch.max_pool2d(input, kernel_size, stride, padding, dilation, ceil_mode)
100%[#####] 85/85 [02:19<00:00, 1.44s/it]
Yaw: 4.2494, Pitch: 3.2705, Roll: 4.5611, MAE: 4.0270
```

```
img_raw = io.imread(img_path)

rgb_mean_normalize = [123, 117, 104]
img = img_raw - rgb_mean_normalize
img = img / 255.
img = torch.tensor(img).unsqueeze(0).permute(0, 3, 1, 2).float().to(device)

dets, landms, poses = eval_single(net, img, cfg)

pose = torch.tensor(poses[0][1])
pred_poses = utils.compute_euler_angles_from_rotation_matrices(pose, use_gpu=False) * 180 / np.pi

img_det_ldmk = img_raw.copy()
img_hpe = img_raw.copy()
img_det_ldmk = get_det_landm_image(img_det_ldmk, dets[0], landms[0], vis_thres=cfg['vis_thres'])
img_hpe = get_hpe_image(img_hpe, dets[0], pred_poses[0].reshape(1, 3).numpy(), vis_thres=cfg['vis_thres'])

fig, ax = plt.subplots(1, 2)
ax[0].imshow(img_det_ldmk); ax[0].set_title('face detection'); ax[0].axis('off')
ax[1].imshow(img_hpe); ax[1].set_title('head pose estimation'); ax[1].axis('off')
plt.show()
```

```
Loading pretrained model from ./weights/BIWI/BIWI_Resnet50_Final.pth
remove prefix 'module.'
Missing keys:0
Unused checkpoint keys:0
Used keys:462
Finished loading model!

/opt/conda/lib/python3.8/site-packages/torch/nn/functional.py:718: UserWarning: Named tensors and all their associated APIs are an experimental feature and subject to change. Please do not use them for anything important until they are released as stable. (Triggered internally at /pytorch/c10/core/TensorImpl.h:1156.)
  return torch.max_pool2d(input, kernel_size, stride, padding, dilation, ceil_mode)
```



프로젝트 보조 설명 : 플랫폼 문제점 및 개선사항

- 신규 프로젝트 개설시 shared memory가 낮게 잡혀서 데이터 로더 또는 배치 사이즈 등 조정에 한정적인 문제점이 있다.

Solution : Docker 개설시 shm-size 를 최대로 설정해주면 해결 가능. Ex) shm-size : 128G

- Remote 사용시 가끔 마이크 작동이 안되는 상황이 빈번히 발생.

참조

Berg, Axel, Magnus Oskarsson, and Mark O'Connor. "Deep ordinal regression with label diversity." *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021.

Deng, Jiankang, et al. "Retinaface: Single-shot multi-level face localisation in the wild." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

Hempel, Thorsten, Ahmed A. Abdelrahman, and Ayoub Al-Hamadi. "6D Rotation Representation For Unconstrained Head Pose Estimation." arXiv preprint arXiv:2202.12555 (2022).

Sheka, Andrey, and Victor Samun. "Rotation Augmentation for Head Pose Estimation Problem." *2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*. IEEE, 2021.

경청해 주셔서 감사합니다.

Menti : 김수정, 한경탁



suprema
SECURITY & BIOMETRICS