

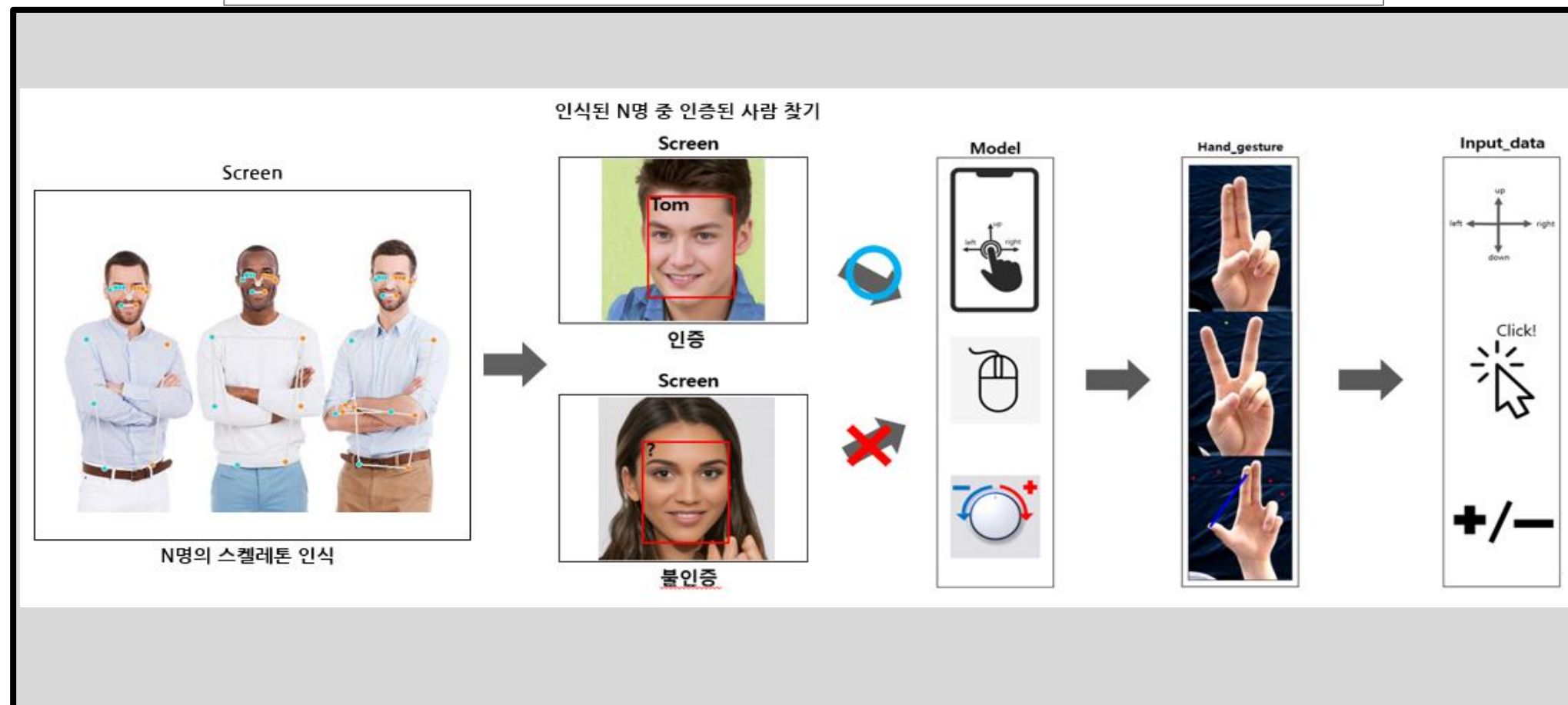


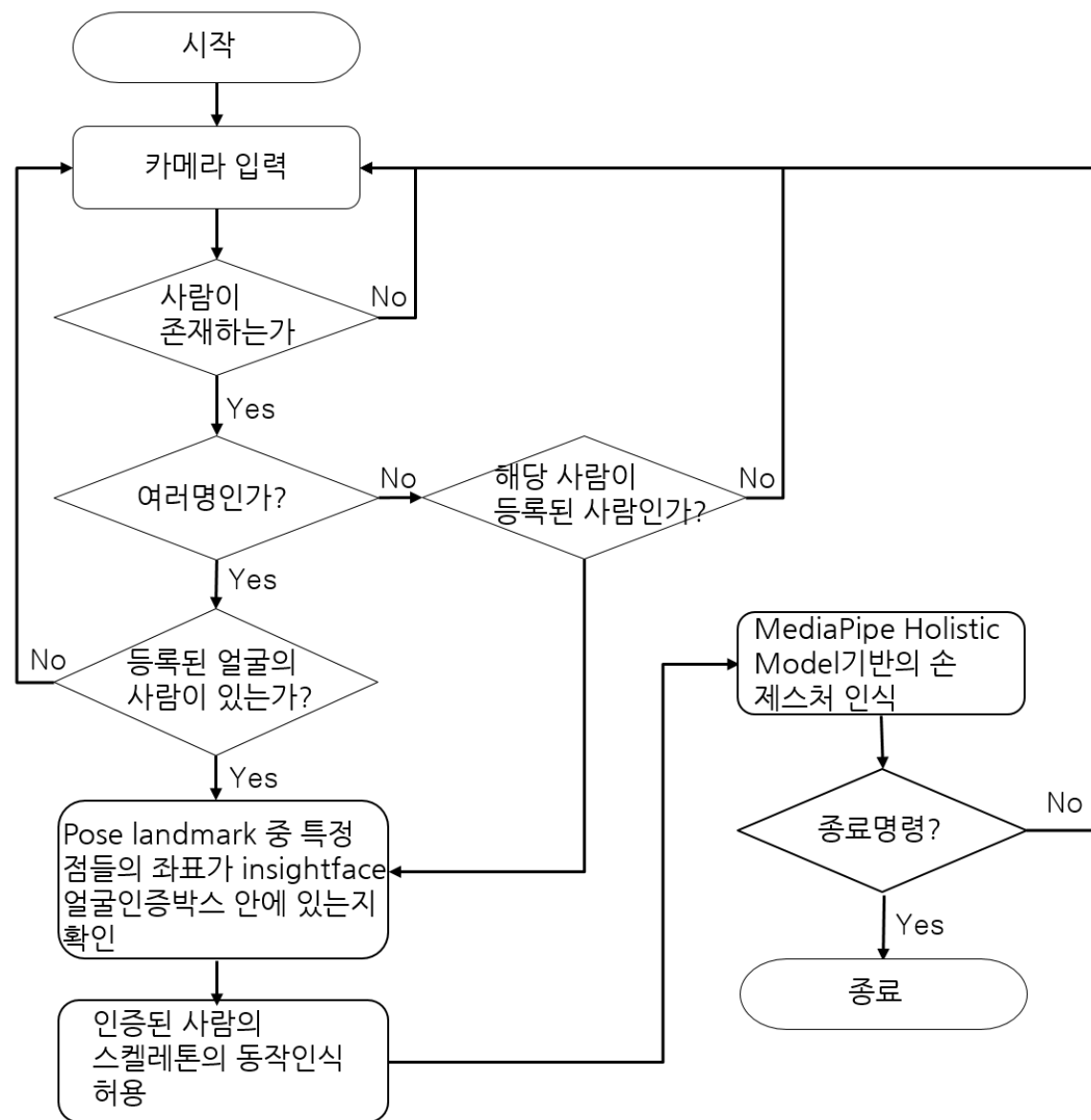
2023. 2. 10.
홍윤아, 송락빈, 곽노윤
백석대학교 컴퓨터공학부

InsightFace 기반의 얼굴 인증과 MediaPipe Holistic 모델 기반의 손 제스처 인식을 이용한 사용자 인터페이스에 관한 연구

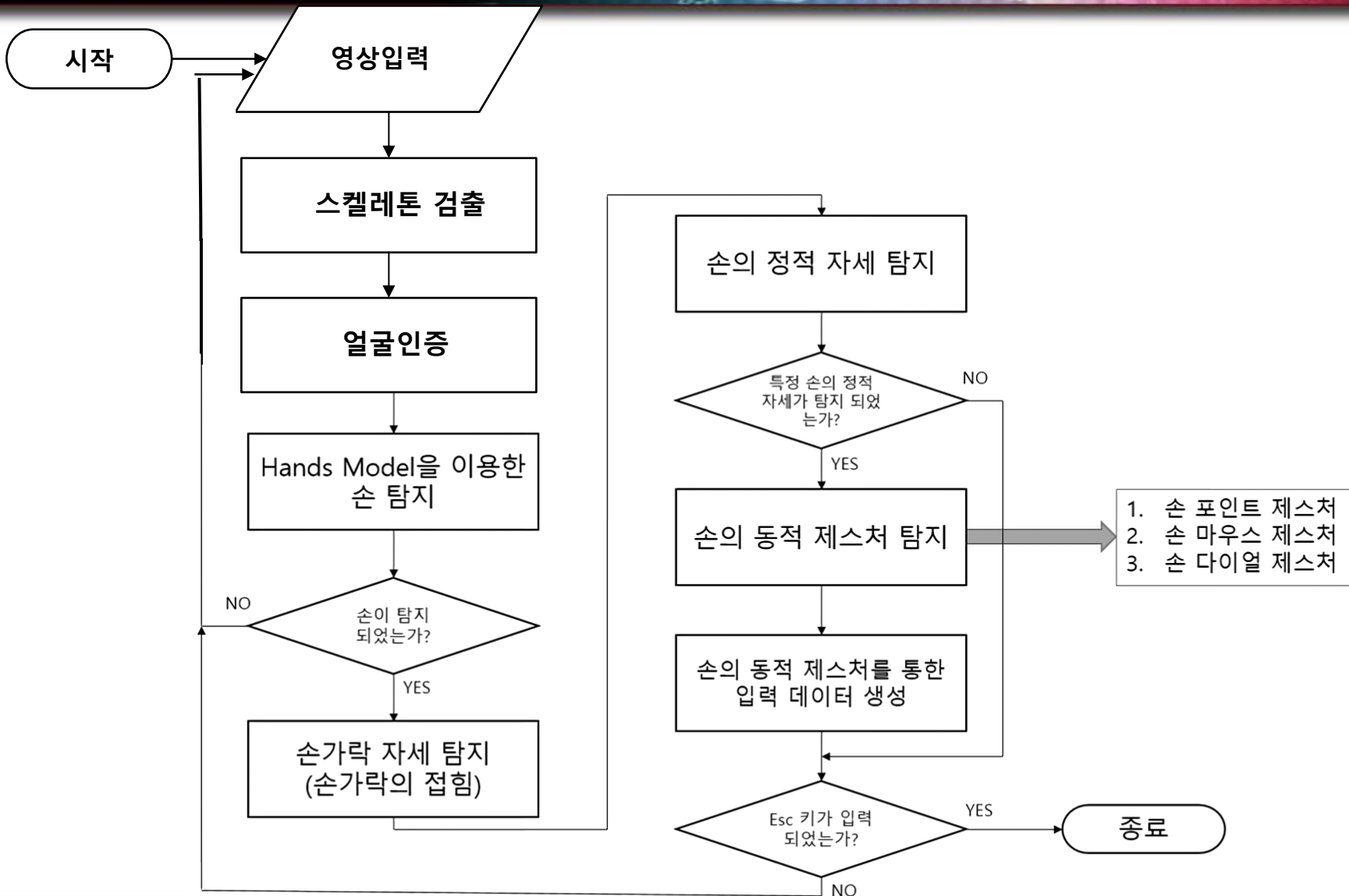


InsightFace 기반의 얼굴 인증과 MediaPipe Holistic 모델 기반의 손 제스처 인식을 이용한 사용자 인터페이스에 관한 연구





제안된 손 제스처 인식 방법의 순서도



□ 얼굴 인증 데이터셋 생성

- 얼굴을 등록하기 위해 동영상의 프레임을 이용하여 데이터를 생성
- 8초의 30fps 동영상 2개 필요 (정면, 측면, 기울어진 얼굴)
- 각 영상마다 프레임 간의 간격을 띄워서 50개의 데이터씩 한 사람당 총 100개의 데이터 생성



<정면>



<측면>



<기울어진 얼굴>

그림 1. 사진 데이터의 얼굴 구도

□ 얼굴 데이터 학습

- InsightFace의 RetinaFace로 얼굴 검출을 하고 Arcface로 얼굴 인식을 함
- Arcface의 DCNN 임베딩 방법을 이용해 얼굴의 임베딩 (n, 512)을 추출 후 Keras로 학습함
- InsightFace에서 사용할 수 있도록 ONNX 파일로 변환함

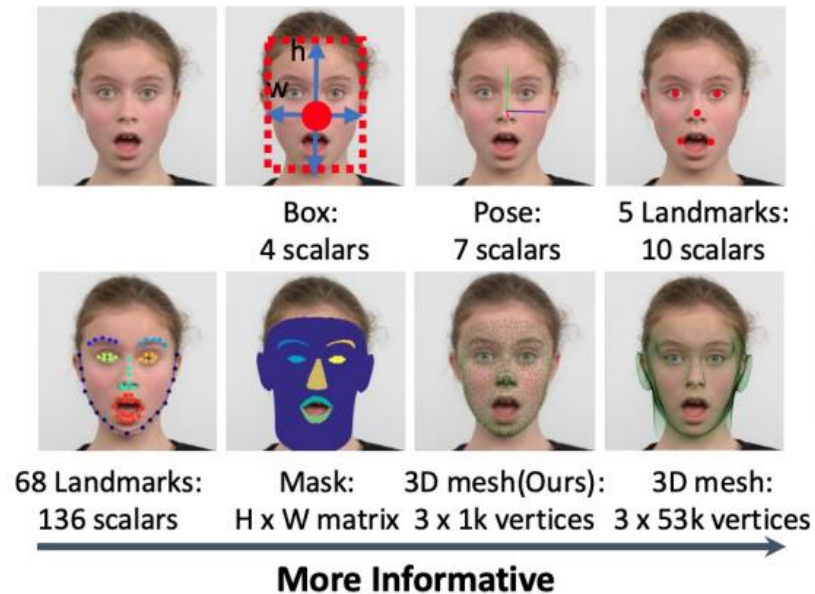


그림 2. RetinaFace논문의 figure1

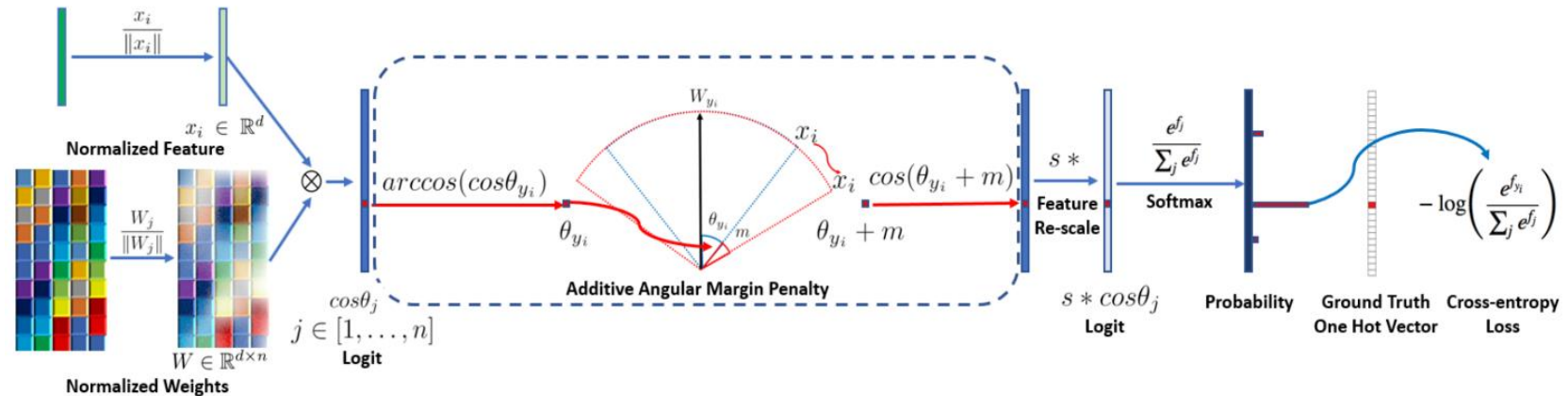


그림 3. Arcface논문의 figure2

❑ InsightFace를 이용한 얼굴 인증

- 미리 학습해 둔 얼굴 리스트에서는 인식 이름으로 인식되고 정확도 0.95 이하로는 인증이 되지 않음
- Arcface로 생성된 임베딩으로 얼굴 인증 진행

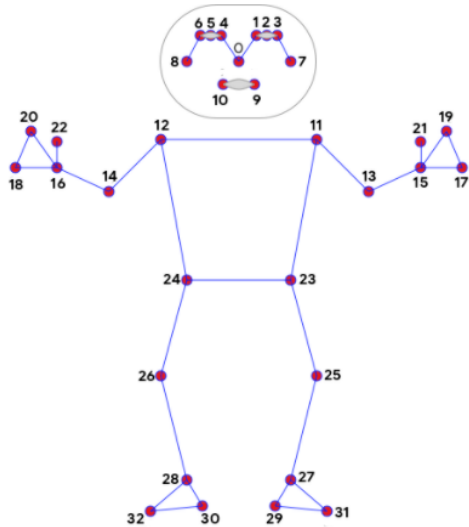
Search Identities



그림 4. 인물 인식 예시

□ Holistic을 이용한 조작권자 인식

- Holistic 모델에서 복수의 제스처 탐지



- | | |
|--------------------|----------------------|
| 0. nose | 17. left_pinky |
| 1. left_eye_inner | 18. right_pinky |
| 2. left_eye | 19. left_index |
| 3. left_eye_outer | 20. right_index |
| 4. right_eye_inner | 21. left_thumb |
| 5. right_eye | 22. right_thumb |
| 6. right_eye_outer | 23. left_hip |
| 7. left_ear | 24. right_hip |
| 8. right_ear | 25. left_knee |
| 9. mouth_left | 26. right_knee |
| 10. mouth_right | 27. left_ankle |
| 11. left_shoulder | 28. right_ankle |
| 12. right_shoulder | 29. left_heel |
| 13. left_elbow | 30. right_heel |
| 14. right_elbow | 31. left_foot_index |
| 15. left_wrist | 32. right_foot_index |
| 16. right_wrist | |

그림 5. Holistic의 Pose landmark



그림 6.화면의 인식된 사람들의 스켈레톤

그림 7.인식된 사람은 다시 인식되지 않도록 제거



□ Holistic을 이용한 조작권자 인식

- Holistic 모델에서 복수의 제스처 탐지

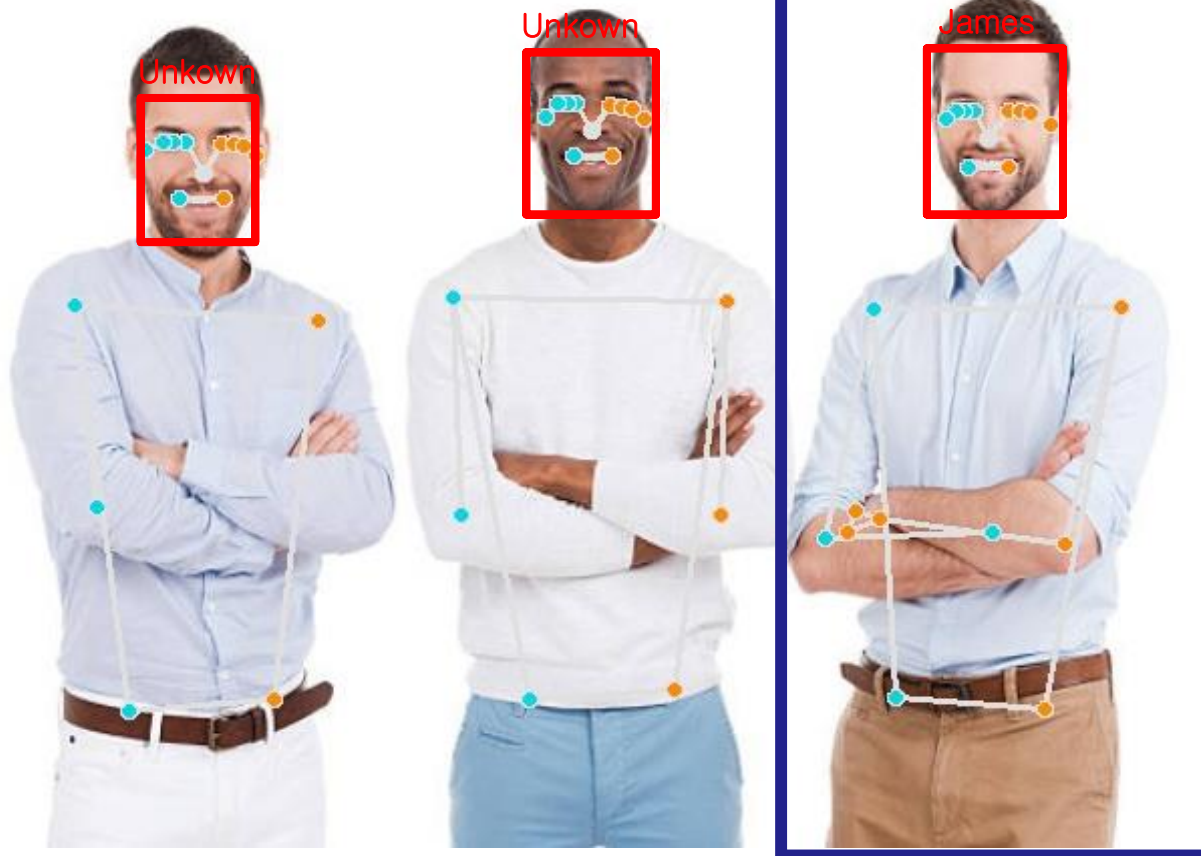


그림 8. 얼굴 인증된 사람의 스켈레톤의 특정 값들이
얼굴인증 영역 안에 들어가 있는지 확인

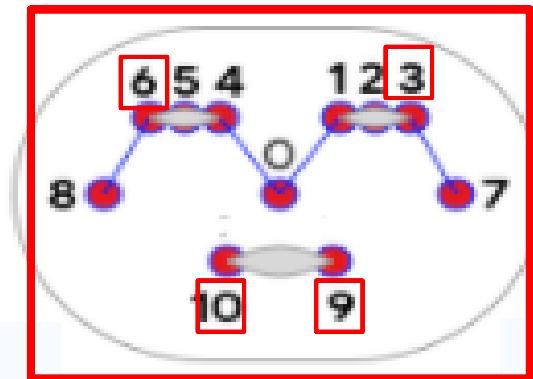
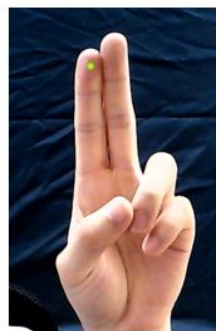


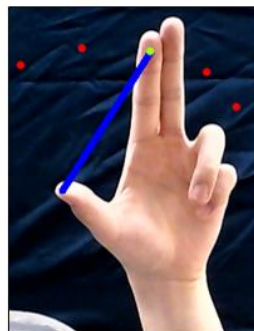
그림 9. 인증된 사람의 얼굴이 맞는지 확인
3,6,9,10 랜드마크 값 사용

□ 손의 정적 자세 탐지

- 손의 정적 자세란 손의 제스처를 입력받기 위한 손의 모양으로, 특징점과 손의 정적 자세 정보를 이용해 구분함
- 손의 정적 자세를 탐지하기 위해서는 손가락의 접힘 상태와 손가락 특징점들 사이의 거리 사용
- 그림 1(a)의 **손 포인트 자세**, 그림 1(b)의 **손 다이얼 자세**, 그림 1(c)~그림 1(e)의 **손 마우스 자세**



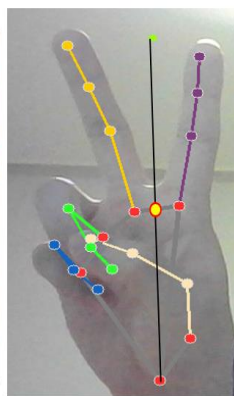
(a)



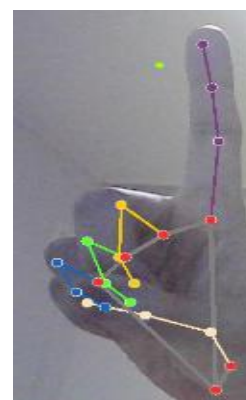
(b)



(c)



(d)



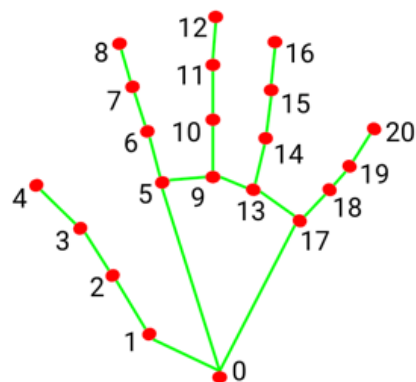
(e)

- (a) 손 포인트 자세
- (b) 손 다이얼 자세
- (c) 손 마우스 자세(좌 클릭)
- (d) 손 마우스 자세(마우스 커서 모드)
- (e) 손 마우스 자세(우 클릭)

그림 5. 손의 정적 자세 유형

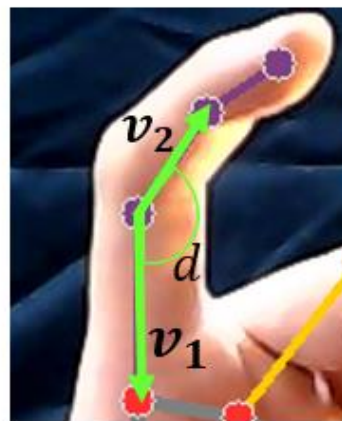
손의 정적 자세 탐지

- 통상의 손의 퍼짐을 확인하는 방법은 TIP과 DIP의 좌표 비교하는데, 간편하지만 손이 돌아가 아래를 향하면 탐지 할 수 없다는 단점이 존재함
- 제안된 손 제스처 인식에서는 모델에서 출력되는 3차원 공간에서의 특징점을 활용함. 2개의 손가락 마디에서 3차원 벡터를 추출하고 내적을 통해 각도를 계산함



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP

11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP



$$d = \cos^{-1} \left(\frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{|v_1| |v_2|} \right)$$

where $v_1 = (x_1, y_1, z_1)$, $v_2 = (x_2, y_2, z_2)$

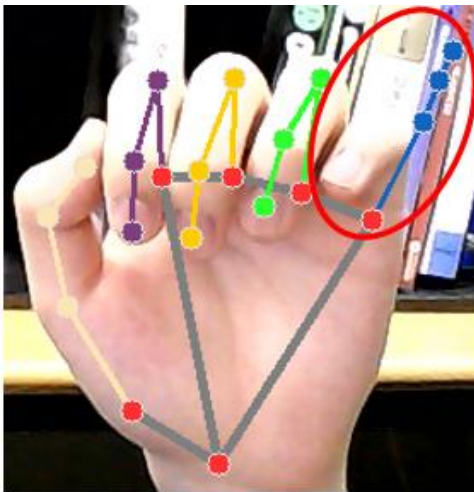
그림 6. 손의 정적 자세 탐지

□ 손의 동적 제스처 탐지

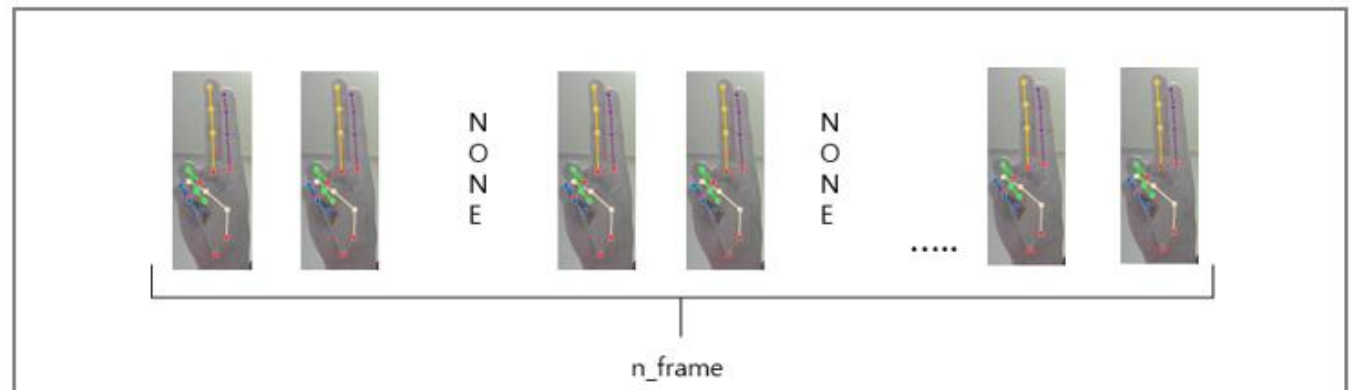
- 손의 동적 제스처란 손의 정적 자세의 변화를 탐지하여 최종적인 입력 데이터를 생성하는 제스처
- 동적 제스처를 탐지하는 과정은 특정 프레임 동안 같은 종류의 정적 자세 입력, 손의 정적 자세에서 특정 조건들을 만족하는 변화 탐지 2개의 단계로 구성됨

1. 특정 프레임 동안 같은 종류의 정적 자세 입력(손 포인트 제스처)

- MediaPipe Hands Model은 높은 성능의 손 탐지를 보여주지만 손의 각도 및 이동에 따라 그림 3(a)와 같이 오작동을 일으킴
- 따라서 그림 3(b)와 같이 특정 프레임 동안 Model의 오차로 인해 정적 자세가 탐지되지 않는 경우가 발생함
- 이 문제를 해결하기 위해 신뢰도 개념을 적용하여 같은 정적 자세가 80% 이상 입력되면 동적 제스처가 입력되었다고 탐지함



(a)



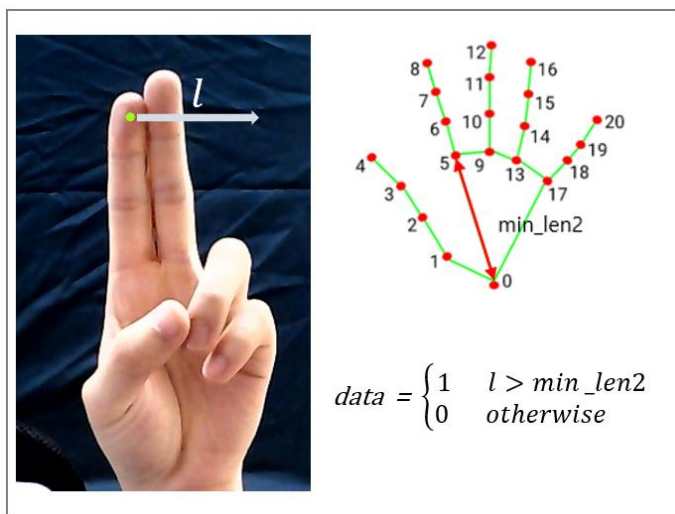
(b)

□ 손의 동적 제스처 탐지

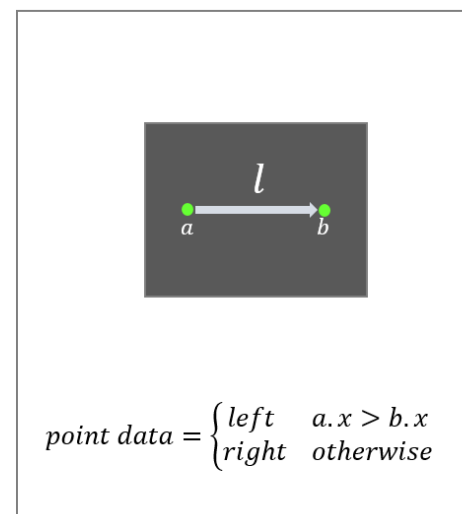
- 손의 동적 제스처란 손의 정적 자세의 변화를 탐지하여 최종적인 입력 데이터를 생성하는 제스처
- 동적 제스처를 탐지하는 과정은 다음과 같이 크게 2개의 단계로 구성됨

2. 손의 정적 자세에서 특정 조건들을 만족하는 변화 탐지(손 포인트 제스처)

- 그림 3(c)와 같이 검지 TIP에 녹색 포인트의 이동거리를 통해 포인트 동적 제스처 입력 확인
- 그림 3(d)와 같이 포인트 이동의 시작 점과 끝 점 좌표 비교를 통해 방향 확인



(c)



(d)

그림 8. 손의 동적 제스처 탐지

손의 동적 제스처 탐지

2. 정적 자세에서 특정 조건들을 만족하는 변화 탐지

그림 3(e)와 같이 전체 n_frame 중 앞의 절반은 마우스 커서 모드이 정적 자세이며 뒤의 최근 $\frac{1}{4}$ 은 클릭 정적 자세를 확인함

(2) 손 마우스 제스처

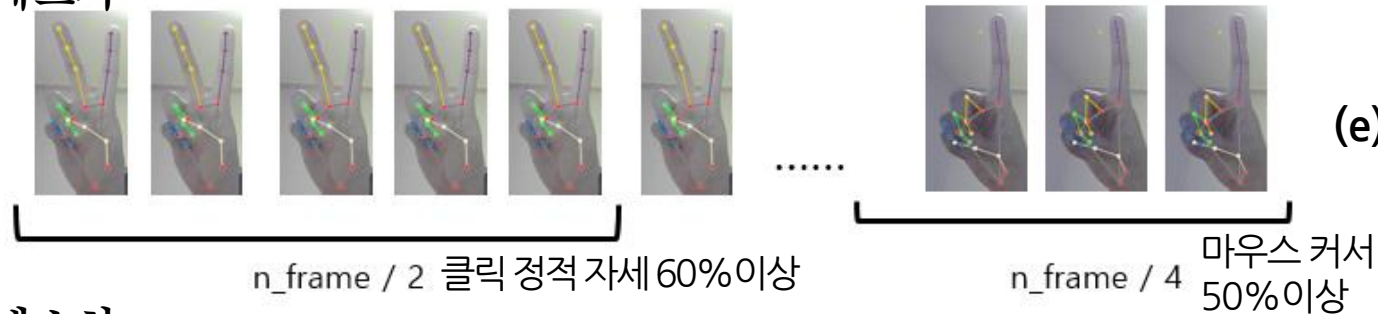
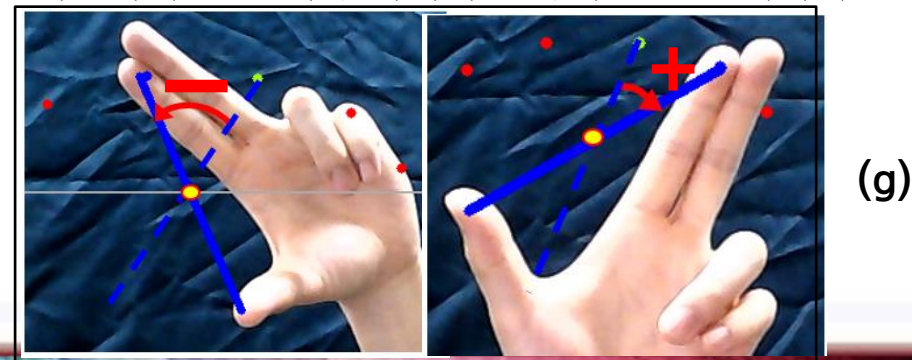
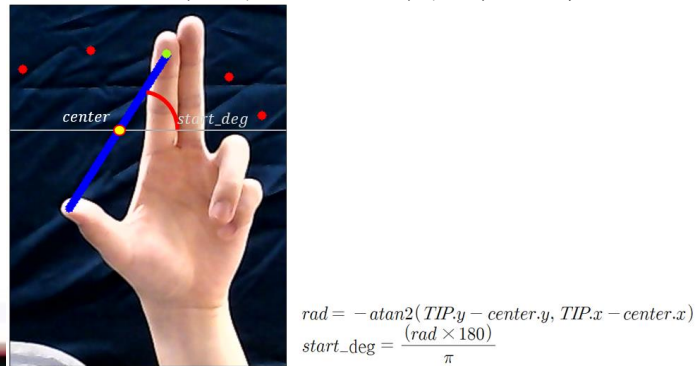


그림 8. 손의 동적 제스처 탐지

(3) 손 다이얼 제스처

- 다이얼 정적 자세가 들어오면 그림 3(f)와 같이 검지의 TIP과 엄지 TIP을 이용해 시작 각도를 계산
- $n_frame/4$ 동안 손 다이얼 자세의 비율은 80% 이상이어야 함
- 그 후 새로운 프레임이 들어 오면 $start_deg$ 와 현재 각도를 비교하여 30° 이상 차이가 발생하면 +/- 데이터 발생



□ 가상의 스마트 홈 기기 제어 시뮬레이션

- 제안한 손 제스처 인식 방법을 이용한 가상의 스마트 홈 제어 시뮬레이션 수행
- 그림 4(a)와 같이 가상의 스마트홈 기기 제어 메뉴 구성, 그림 4(b)와 같이 초기 상태에서 down을 통해 메인 메뉴를 활성화 후 left/right를 통해 원하는 메뉴로 이동한 뒤 down을 통해 기기의 서브 메뉴 활성화
- 그림 4(c)는 선택한 TV 메뉴에서 다이얼 입력을 통해 TV 채널을 변경

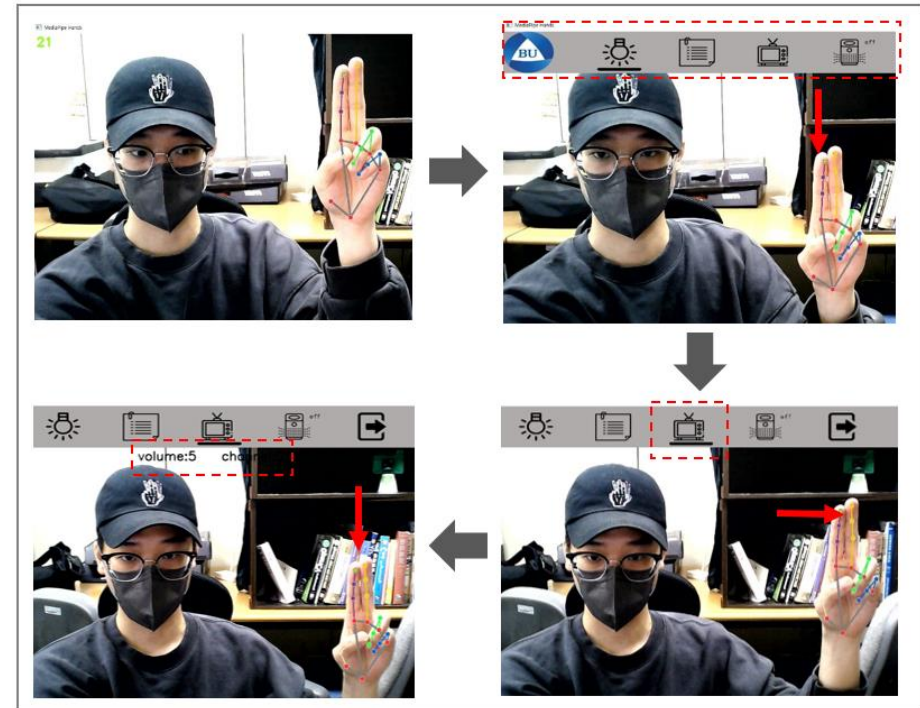
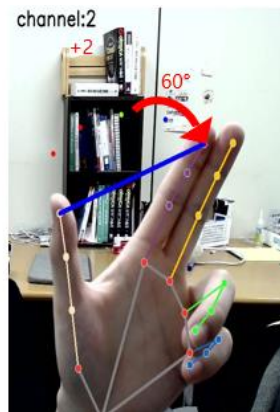
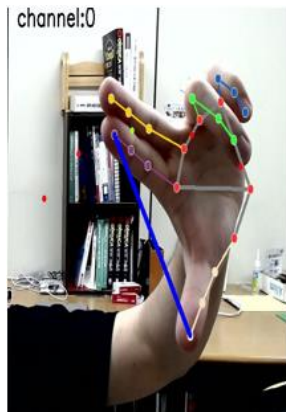


그림 9. 가상의 스마트 홈 기기 제어 시뮬레이션



Image Processing and
Computer Graphics Lab.

2023. 2. 10.
홍윤아, 송락빈, 곽노윤
백석대학교 컴퓨터공학부학부

Thank You !!

InsightFace 기반의 얼굴 인증과 MediaPipe Holistic 모델 기반의 손 제스처 인식을 이용한
사용자 인터페이스에 관한 연구

