석사학위논문

AI 기반 심층 학습 알고리즘을 활용한 이미지 스캐닝 개발 연구

- 의류 디자인 산업 분야 중심으로 -

2026년

한성대학교 지식서비스&컨설팅대학원

스마트융합컨설팅학과

스마트융합기술컨설팅전공

김 정 인

석사학위논문
지도교수 원종혁

AI 기반 심층 학습 알고리즘을 활용한 이미지 스캐닝 개발 연구

- 의류 디자인 산업 분야 중심으로 -

2025년 12월 일

한성대학교 지식서비스&컨설팅대학원

스마트융합컨설팅학과

스마트융합기술컨설팅전공

김 정 인

석사학위논문
지도교수 원종혁

AI 기반 심층 학습 알고리즘을 활용한 이미지 스캐닝 개발 연구

- 의류 디자인 산업 분야 중심으로 -

위 논문을 컨설팅학 석사학위 논문으로 제출함

2025년 12월 일

한성대학교 지식서비스&컨설팅대학원

스마트융합컨설팅학과

스마트융합기술컨설팅전공

김 정 인

김정인의 컨설팅학 석사학위 논문을 인준함

2025년 12월 일

심사위원장 주 형 근 (인)

심 사 위 원 정 진 택 (인)

심 사 위 원 원 종 혁 (인)

국 문 초 록

AI 기반 심층 학습 알고리즘을 활용한 이미지 스캐닝 개발 연구

- 의류 디자인 산업 분야 중심으로 -

한성대학교 지식서비스&컨설팅대학원
스 마 트 융 합 컨 설 팅 학 과
스 마 트 융 합 기 술 컨 설 팅 전 공
김 정 인

이미지 스캐닝 기술은 19세기 후반 사진측량의 태동과 함께 시작되어 초기에는 수작업 기반의 치수 추출 및 계측 방식에 의존하였으며, 측정자의 숙련도에 따른 편차가 큰 문제가 지속되었다. 20세기 중반 이후 광학장비·사진기술·라스터 기반 스캐닝 기술이 보급되면서 2D 이미지 기반 계측이 본격화되었으며, 디지털 센서의 상용화는 이미지 스캐닝을 산업·의료·패션·안전·스포츠 등 다양한 분야로 확장시키는 계기가 되었다. 이러한 2D 스캐닝 기술은 인체 치수 측정에서 일정 수준의 정량화된 데이터를 제공하였으나, 이미지 왜곡·촬영 각도·조명·해상도 등 다양한 변수에 의해 정확도 확보에 한계가 존재하였다.

1990년대 후반부터는 이러한 한계를 보완하기 위해 구조광(Structured Light), 레이저 스캐닝, Time-of-Flight 기반 센서 등이 도입되면서 정밀한 표면 형상과 볼륨 정보를 확보할 수 있게 되었다. Bartol 외(2021)는 3D 스

캐닝 기반 신체 측정 기술에 대한 체계적 문헌 고찰을 통해 의류 산업의 맞춤형 패턴 제작, 정형·재활 의료 분야의 체형 분석, 스포츠·헬스케어 분야에서 활용도가 크게 증가하였음을 보고하였다. 그러나 3D 스캐너는 고가의 장비 비용, 공간적 제약, 전문 인력 확보, 유지보수 비용 등 구조적 한계를 갖고 있어 중소기업·소상공인·스타트업에서는 도입이 어려운 구조가 유지되고 있다. 이러한 배경에서 최근 주목받는 분야가 AI 알고리즘을 활용한 2D 기반 신체 치수 측정 기술의 고도화이다. 기존 2D 스캐닝 방식은 스마트폰 또는 보급형 카메라 한 대만으로도 데이터 수집이 가능하다는 점에서 비용·접근성 측면의 장점을 가지나, 명확한 깊이(depth)를 측정할 수 없어 치수 정밀도 확보에 제한이 있었다. 그러나 최근 딥러닝 기반 포즈 추정, 단일 이미지 깊이 예측(Monocular Depth Estimation), 인체 파라메트릭 모델 기반 3D 복원 등 다양한 AI 기술이 발전하면서 2D 이미지에서도 3D 스캐닝에 근접한 수준의 신체 치수 추정이 가능해지고 있다. 특히 Loper 등(2015)이 제안한 SMPL(Skinned Multi-Person Linear Model)은 수천 명의 3D 스캔 데이터를 기반으로 인체 형상과 자세를 파라메트릭하게 표현할 수 있는 통계적 인체 모델로서, 단일 2D 이미지로부터 3D 인체 형상을 복원하는 연구의 핵심 기반 기술로 자리 잡았다.

본 연구는 이러한 기술적 발전을 바탕으로 AI 기반 2D 스캐닝 신체 치수 측정 기술의 정확도를 검증하고, 3D 스캐닝 기술을 대체할 수 있는 가능성을 분석하는 것을 목적으로 한다. 촬영 환경·해상도·카메라 사양·의복 착용 여부 등에 따른 정확도 변화를 실험적으로 검증하며, ESG 경영 및 적정기술 관점에서 2D AI 스캐닝의 사회·경제적 파급 효과에 대해서도 제언하고자 한다. 저비용 장비만으로 치수 측정이 가능하다는 점은 환경적(E) 측면에서 장비 제조·폐기 과정의 탄소 배출 감소, 사회적(S) 측면에서 중소기업·개도국의 기술 접근성 확보, 지배구조(G) 측면에서 디지털 전환 촉진이라는 장점을 가진다.

본 연구는 AI 기반 2D 스캐닝 기술이 중소기업 산업 현장에서 3D 스캐닝의 기술적·경제적 한계를 대체할 수 있는지를 실증적으로 검토하였다. 연구 결과는 인체 계측 산업의 접근성 향상, 기술격차 해소, 저비용 디지털화 전략

수립에 기여할 것으로 기대되며, 국가 단위의 디지털 제조혁신, 공공 보건 데이터 구축, 개인 맞춤형 서비스 산업 등 다양한 영역에서 확장 적용될 수 있을 것으로 기대한다.

주제어: 딥러닝, 2D 신체 치수 측정, 이미지 스캐닝, SMPL, 포즈 추정, 의류 디자인 산업, 적정기술

목 차

I. 서 론 ····· 1

1.1 연구의 배경 ····· 1
1.2 연구의 목적 ····· 4
1.3 연구의 기대효과 ····· 5

II. 이론적 배경 ····· 7

2.1 이미지 스캐닝 기술의 발전 ····· 7
2.1.1 수작업 기반 측정 기술 ····· 7
2.1.2 2D 이미지 기반 스캐닝 기술 ····· 10
2.1.3 3D 스캐닝 기술 ····· 13
2.2 2D 이미지 처리 알고리즘 ····· 17
2.2.1 기본 이미지 처리 및 전처리 ····· 17
2.2.2 특징 추출 및 세그멘테이션 ····· 20
2.2.3 객체 인식 및 검출 ····· 22
2.3 AI 딥러닝 기반 신체 치수 측정 기술 ····· 27
2.3.1 포즈 추정 기술 ····· 27
2.3.2 인체 파라메트릭 모델 ····· 30
2.3.3 CNN/Transformer 기반 치수 측정 ····· 31
2.4 선행연구 분석 ····· 34
2.4.1 선행연구 분석결과 ····· 36
2.4.2 3D 스캐닝 기술과 2D 스캐닝 기술 비교 ····· 37

III. 연구방법 및 결과 ····· 41

3.1 연구 개요 및 목표 ····· 41
3.2 연구 환경 및 구성 ····· 42
3.2.1 데이터 수집 ····· 42
3.2.2 데이터 통계 분석 ····· 43

3.3 AI 기반 신체 측정 시스템 개발 ····· 45
3.3.1 시스템 아키텍처 ····· 45
3.3.2 rembg 라이브러리 및 U² -Net 배경 제거 ····· 46
3.3.3 MediaPipe Pose 랜드마크 ····· 52
3.3.4 신체 경계 기반 폭/높이 측정 ····· 56
3.3.5 신체 특징점 추출 및 측정 ····· 63
3.3.6 신체 측정 위치 추출 ····· 64
3.3.7 딥러닝 회귀 모델 ····· 65

IV. 연구 결과 ····· 67

4.1 딥러닝 회귀 모델 기반 측정 정확도 분석 ····· 67
4.2 측정 항목의 오차 분석 및 구조적 한계 ····· 68
4.3 연구 결과 및 시사점 ····· 68

V. 결 론 ····· 70

5.1 결론 및 시사점 ····· 70
5.2 연구의 한계점 ····· 72
5.3 향후 연구방향 ····· 73

참 고 문 헌 ····· 74

ABSTRACT ····· 82

표 목 차

[표 2-1] 수작업 기반 인체 계측 방식의 한계 요인 ····· 07
[표 2-2] 전통적 수작업 인체 계측 도구의 특성 ····· 08
[표 2-3] 2D 기반 스캐닝 기술 사용 시 고려 사항 ····· 11
[표 2-4] 3D 스캐닝 기술 유형별 특성 비교 ····· 13
[표 2-5] 3D 스캐닝 기술의 산업적 도입 장벽 ····· 15
[표 2-6] 주요 이미지 전처리 기법 및 특성 ····· 17
[표 2-7] 특징 추출 기법의 발전 과정 ····· 22
[표 2-8] 주요 포즈 추정 모델 비교 ····· 28
[표 2-9] 주요 인체 파라메트릭 모델 비교 ····· 30
[표 2-10] 딥러닝 기반 치수 측정 접근법 비교 ····· 32
[표 2-11] 선행연구 분석 ····· 36
[표 2-12] 3D 스캐닝 기술과 2D 스캐닝 기술 비교 ····· 38
[표 3-1] 시스템 개발 목표 사양 ····· 41
[표 3-2] 이미지 데이터 수집 기준 ····· 43
[표 3-3] 수집 데이터 기술 통계량 ····· 44
[표 3-4] 시스템 모듈 구성 ····· 45
[표 3-5] rembg 라이브러리 지원 모델 ····· 49
[표 3-6] 신체 랜드마크 위치 ····· 55
[표 3-7] 본 연구에서 활용하는 주요 MediaPipe 랜드마크 ····· 56
[표 3-8] 부위별 측정 위치 계산 비율 ····· 57

그 림 목 차

[그림 2-1] 2D 스캐닝 기술의 발전 과정 ····· 12
[그림 2-2] 이미지 전처리 파이프라인 구조 ····· 19
[그림 2-3] 특징 추출 및 세그멘테이션 통합 파이프라인 ····· 26
[그림 2-4] 2D-3D 포즈 추정 파이프라인 ····· 29
[그림 2-5] CNN vs Transformer 기반 치수 측정 비교 ····· 33
[그림 3-1] 데이터 수집 촬영 환경 ····· 42
[그림 3-2] rembg 배경 제거 처리 파이프 라인 Flow ····· 47
[그림 3-3] U² -Net 아키텍처 구조 ····· 48
[그림 3-4] rembg를 이용한 배경 제거 모듈 ····· 51
[그림 3-5] rembg + U² -Net 모델을 이용한 배경 제거 처리 결과 ····· 52
[그림 3-6] MediaPipe Pose 33개 신체 랜드마크 위치 ····· 54
[그림 3-7] 알파 채널 기반 신체 경계 탐색 함수 ····· 64
[그림 3-8] 신체 측정 위치 추출 결과 ····· 65
[그림 3-9] DNN 회귀 모델 구조 ····· 66

Ⅰ. 서론

1.1 연구의 배경

이미지 스캐닝(image scanning) 기술은 19세기 후반 사진측량(photogrammetry)의 출현과 함께 인체 계측(human anthropometry) 분야에서 활용되기 시작하였다 (Mitchell, H. L., 1992; Okimoto, M. L. R., & Klein, A. A., 2012). 대한민국에서는 1979년부터 국가기술표준원 주관으로 한국인 인체치수 조사사업(Size Korea)이 시작되어, 한국인의 표준 인체 치수 데이터베이스 구축이 체계적으로 이루어져 왔다(국가기술표준원 사이즈코리아, 2004; 김희은 외, 2017). 초기의 스캐닝 기술은 전적으로 수작업 기반의 치수 추출 및 계측 방식에 의존하였으며, 측정자의 숙련도나 도구의 정밀도에 따라 측정 오차가 크게 발생하는 한계가 존재하였다 (Franke-Gromberg et al., 2010; Han, K. et al., 2010). 인체의 길이·둘레·각도와 같은 기초 계측조차 반복 측정 시 편차가 크게 나타났고 (Ulijaszek, S. J., & Kerr, D. A., 1999), 이러한 비정량적 특성은 산업적 활용을 제한하는 요인이 되었다.

20세기 중반 이후 광학장비의 정밀화, 디지털 사진 기술의 발전, 라스터(raster) 기반의 초기 스캐닝 기술 보급이 본격화되면서 2D 이미지 기반 계측이 빠르게 확산되었다 (Weinberg, S. M. et al., 2004). 남윤자 등(2004)은 Size Korea 사업에서 3차원 인체측정방법의 표준화를 제시하였으며 (남윤자 외, 2004), 이를 통해 한국인의 체형 특성에 관한 체계적인 연구 기반이 마련되었다. 특히 디지털 센서의 상용화는 이미지 스캐닝을 의료, 스포츠, 패션, 산업 안전 등 다양한 영역으로 확장시키는 촉매 역할을 하였다 (Heymsfield, S. B. et al., 2018; Rumbo-Rodríguez, et al., 2021). 이러한 2D 기반 계측 방식은 비교적 저비용과 높은 접근성을 바탕으로 일정 수준의 정량화된 데이터를 제공하였으나, 이미지 왜곡, 촬영각도 변화, 조명 조건, 해상도 등 다양한 외생 변수에 의해 정밀도 확보에 구조적 한계가 존재하였다 (Farkas, L. G., 1996; Pateraki, M., Fragkoulidou, V., & Stoltidou, A., 2006).

1990년대 후반 이후에는 기존 2D 이미지 기반 계측의 한계를 극복하기

위해 3D 스캐닝 기술이 급속히 발전하였다 (Bartol, K. et al., 2021). 국내에서도 사이즈코리아 사업을 통해 3차원 전신 스캐너가 도입되어 한국인의 체형 특성에 관한 대규모 데이터베이스가 구축되었다 (국가기술표준원, 2010; 국가기술표준원, 2021). 백경자와 이정란(2008)은 3차원 바디 스캐너를 이용한 남성 상반신 인체측정 연구에서 직접측정과의 비교를 통해 스캐너 측정의 신뢰성을 검증하였다 (백경자, 이정란, 2008). 한희경 등(2010)은 Size Korea 성인 여성을 대상으로 3D 스캔 측정값과 직접 측정값의 비교 분석을 통해 측정 정확도를 검증하였다 (한희경, 남윤자, 최경미, 2010).

구조광(Structured Light) (Wells, J. C. K., Treleaven, P., & Cole, T. J, 2007; D'Apuzzo, N., 2007), 레이저 스캐닝 (Wang, J., Gallagher et al., 2006), Time-of-Flight(ToF) 센서 (Bartol et al., 2021)와 같은 고정밀 광학 측정 기술은 표면 형상, 깊이(depth), 체적(volume) 정보를 정밀하게 획득할 수 있도록 하였다. 이를 바탕으로 의류 산업의 맞춤형 패턴 제작 (남윤자, 이정란, 2001; 김혜경, 2001), 정형·재활의학 분야의 체형 분석 (Treleaven, P., & Wells, J., 2007), 스포츠·헬스케어 분야의 신체 변화 추적 (Ng, B. K., Hinton, B. J., Fan, B., Kanaya, A. M., & Shepherd, J. A., 2016) 등에서 활용도를 높였다. 유희천 등(2008)은 헬리콥터 조종실 설계를 위한 한국인 조종사의 인체측정 특성을 분석하였으며 (유희천 외, 2008), 이원섭 외(2013)은 이를 확장하여 한국인 헬리콥터 조종사와 미군 인체측정치의 비교 연구를 수행하였다.

그러나 3D 스캐너는 장비 가격이 매우 고가이며, 촬영 환경에 대한 공간 제약, 시스템 설치 및 운영을 위한 전문 인력 확보, 주기적인 장비 교정(calibration) 및 유지관리 비용 등 경제적·운영적 부담이 크다는 구조적 한계를 지닌다 (Bartol, K. et al., 2021; Ashby, N., LaPorte, et al., 2023). 특히 중소기업·소상공인·스타트업과 같이 자원이 제한된 조직에서는 3D 스캐너 도입 비용 대비 효익이 낮게 평가되는 경우가 많아 기술 보급이 제한되는 문제점이 지속되고 있다 (Smith, B., McCarthy, C., Dechenaud, M. E., Wong, M. C., Shepherd, J., & Heymsfield, S. B., 2022). 이러한 상황에서 글로벌 경기 침체와 비용 절감 요구가 강화되면서 고가의 3D 기반 계측 방식의 사

용 빈도는 점차 감소하고 있으며, 이를 대체할 수 있는 혁신적이고 저비용 기반의 계측 기술에 대한 수요가 증가하고 있다 (Tinsley et al., 2024; García-Muro San José et al., 2024).

최근 딥러닝(Deep Learning) 기술의 비약적 발전은 기존의 2D 이미지 기반 계측 기술이 지녔던 한계를 극복할 수 있는 새로운 접근법을 제시하고 있다. 한국전자통신연구원(ETRI)은 딥러닝 기반 영상 분석 기술을 개발하여 객체 검출 및 자세 추정 분야에서 세계적 수준의 성과를 달성하였으며 (한국전자통신연구원, 2018; 한국전자통신연구원., 2020), 포즈 추정(Human Pose Estimation) (Lugaresi et al, 2019; Bazarevsky, Grishchenko et al, 2020; Cao et al, 2021), 단안 카메라 기반 깊이 추정(Monocular Depth Estimation) (Khan, F. et al., 2020; Zhao, C. et al., 2020), SMPL과 같은 인체 파라메트릭 모델 기반 3D 복원 (Bogo, F., Kanazawa et al., 2016), 멀티뷰 스테레오(Multi-View Stereo) (Seitz et al., 2006), 이미지 정합(Image Matching) 알고리즘 (Schonberger, J. L., & Frahm, J. M., 2016) 등 다양한 AI 기술이 고도화되면서, 이제는 단일 2D 이미지에서도 3D 스캐닝에 근접한 수준의 신체 치수를 추정하는 것이 가능해졌다 (Majmudar et al., 2022; Farina et al., 2025).

국내에서도 AI 기반 인체 측정 기술 연구가 활발히 진행되고 있다. 이지용과 김창균(2024)은 사이즈코리아 데이터를 활용하여 2D 이미지로부터 3D 체형 이미지를 생성하는 AI 기반 모델을 개발하였다 (이지용, 김창균, 2024). 또한 한국컴퓨터그래픽스학회에서는 한국인 체형의 통계적 특성을 반영한 K-SMPL 모델을 제안하여, 기존 SMPL 모델이 서양인 체형에 편향되어 있는 한계를 극복하고자 하였다 (한국컴퓨터그래픽스학회, 2022). 특히 대규모 인체 데이터를 학습한 통계적 인체 모델링 기술은 자세 변화, 조명 차이, 의복 착용에 따른 형태 변형 등 복잡한 요인을 정량적으로 보정함으로써 기존 2D 기반 계측 방식의 정확도를 크게 향상시키고 있다 (Loper et al., 2015; Kanazawa et al., 2018).

한국의류학회와 대한인간공학회에서는 3D 스캔 데이터를 활용한 체형 분류 연구가 지속적으로 수행되어 왔다. 서동애와 오설영(2012)은 3D 인체 스

캔 데이터를 활용한 체형별 인대모형 개발에 관한 연구를 발표하였고 (서동애, 오설영, 2012), 조신현과 김영옥(2012)은 사이즈코리아 제5차, 6차 3D 인체형상데이터 비교를 통한 30대 여성의 신체변화를 연구하여 시계열적 체형변화 추이를 분석하였다 (조신현, 김영옥, 2012). 김소영과 손현정(2020)은 3D 인체측정학을 활용한 한국 중년 여성의 체형 분류 연구를 통해 의류 산업에서의 활용 가능성을 제시하였다 (김소영, 손현정, 2020). 김윤상 등(2019)은 3D 스캔 데이터에 의한 성인 남성의 체간부 형태 유형화 연구를 수행하였다 (김윤상, 임성한, 신동범, 2019). 권미정과 황인주(2018)는 인체의 전면 및 후면 3D 스캔 영상을 이용한 체형 분석 방법을 제안하였다 (권미정, 황인주, 2018).

이와 같은 기술적 진보는 낮은 비용, 높은 접근성을 기반으로 기존 3D 스캐너의 기능을 상당 부분 대체할 수 있는 가능성을 보여주며, 산업 전반에서 AI 기반 2D 스캐닝 기술 도입에 대한 관심을 빠르게 확산시키고 있다 (Staynor, et al., 2022; Leymarie, F. F., & de Aguiar, E., 2023). 또한 ESG 경영과 적정기술(Appropriate Technology)의 중요성이 부각되면서 (United Nations, 2015), 고가 장비 의존도를 낮추고 친환경·저비용 기반의 기술로 전환하려는 사회적 요구와 맞물려 2D AI 스캐닝 기술은 지속 가능한 기술 대안으로 주목받고 있다 (Heymsfield, S. B., & Shepherd, J. A., 2024).

1.2 연구의 목적

본 연구는 AI 기반 심층 학습 알고리즘을 활용한 2D 이미지 스캐닝 기술이 기존 3D 스캐닝 기술을 대체할 수 있는지 여부를 실증적으로 분석하는 것을 목적으로 한다. 구체적으로는 다음의 세부 목적을 설정한다.

첫째, 딥러닝 기반 2D 이미지 분석 및 3D 형상 추정 기술의 정확도를 정량적으로 평가한다. 기존 3D 스캐너와 동일 조건에서 촬영된 2D 이미지를 AI 알고리즘으로 분석하여 실제 치수 대비 정밀도를 비교함으로써, 2D AI 스캐닝이 제공할 수 있는 기술적 수준을 객관적으로 검증한다.

둘째, 촬영 환경(조명·배경·촬영거리), 이미지 해상도, 카메라 사양, 피사체

의 의복 착용 여부 등 다양한 변수에 따른 정확도 변화 패턴을 분석한다. 이는 산업적 활용을 위한 실질적 가이드라인을 도출하고 현장 적용성을 평가하는 데 중요한 기초 자료가 된다.

셋째, 3D 스캐너 대비 2D AI 스캐닝 기술의 경제적·운영적 효익을 분석하여 기술 대체 가능성과 도입 타당성을 검토한다. 장비 비용, 유지보수 비용, 공간 제약, 운영 인력 부담 등 실무적 관점에서 두 기술 간 차이를 분석하고, 중소기업·개도국·소규모 산업군 등 자원 제약 환경에서의 활용 가능성을 집중적으로 평가한다.

넷째, ESG 기반 경영(Environment·Social·Governance) 및 적정기술 관점에서 AI 기반 2D 스캐닝의 사회·경제적 파급효과를 제언한다. 저비용 장비 활용에 따른 환경적 이점, 기술 접근성 향상에 따른 사회적 가치, 데이터 기반 의사결정의 활성화에 따른 지배구조 개선 가능성을 분석하여 지속 가능한 기술 생태계 구축에 기여하고자 한다.

1.3 연구의 기대효과

본 연구는 AI 기반 2D 스캐닝 기술이 갖는 기술적 가능성과 산업적 활용성을 실증적으로 제시함으로써 다음과 같은 기대효과를 제공한다.

첫째, 3D 스캐닝 기술의 고비용·고난도 문제를 해소할 수 있는 대체 기술로서의 가능성을 제시한다. 스마트폰 또는 일반 카메라만으로도 고정밀 신체 치수 추정이 가능하다면, 기존의 장비 도입 장벽을 대폭 낮출 수 있으며 다양한 산업군에서 계측 기술의 접근성을 향상시킬 수 있다.

둘째, 중소기업·소상공인·스타트업 등 자원 제약 조직의 디지털 전환(DX) 촉진 효과를 기대할 수 있다. 이는 산업 전반의 기술격차(Digital Divide)를 완화하고 맞춤형 제품 제작, 개인화 서비스, 제조 자동화 등에 필요한 기초 데이터를 비용 효율적으로 확보할 수 있는 기반을 제공한다.

셋째, ESG 경영 관점에서 환경적·사회적·지배구조적 가치 창출이 가능하다. 3D 스캐너 대비 장비 제조 및 폐기 과정에서 발생하는 탄소 배출을 줄이며(E), 전 세계 다양한 산업 및 지역에서 기술 접근성을 높일 수 있고(S), 데

이터 기반 의사결정을 통해 조직의 운영 효율성과 투명성을 강화할 수 있다(G).

넷째, 의류·의료·스포츠·보건·안전·디지털 제조 등 대규모 인체 계측이 필요한 다양한 산업 분야에 새로운 기술 기반을 제공한다. 본 연구의 결과는 산업 맞춤형 알고리즘 개발, 현장 적용 프로토콜 수립, 정책적 지원 근거 마련 등 후속 연구 및 기술 확산에 중요한 기여를 할 수 있다.

다섯째, 본 연구는 국가 단위 디지털 제조혁신 정책, 공공 보건 데이터 구축, 개인 맞춤형 서비스 산업 등 확장 가능한 적용 모델을 제시함으로써 광범위한 산업·공공 분야에서의 기술 활용 가능성을 제시한다.

Ⅱ. 이론적 배경

2.1 이미지 스캐닝 기술의 발전

2.1.1 수작업 기반 측정 기술

인체 계측(human anthropometry)은 인간의 신체적 특성을 정량적으로 파악하기 위한 연구이자 실무 분야로, 19세기 후반부터 체계적인 계측 방법이 정립되기 시작했다. 초기에는 줄자, 캘리퍼, 신장계, 고니오미터 등 비교적 단순한 도구를 활용한 수작업 측정 방식이 주된 접근 방법이었다. 당시 기술 수준에서는 이 방식이 가장 현실적이었으나, 측정자 개인의 숙련도와 판단 차이에 따라 결과 편차가 발생할 수 있다는 점에서 본질적인 한계를 갖고 있었다.

[표 2-1] 수작업 기반 인체 계측 방식의 한계 요인 (Martin & Saller, 1957; Ulijaszek & Kerr, 1999; 연구자 재구성, 2025)

한계 요인	세부 내용	산업적 영향
측정자 간 편차	측정자마다 기준점(landmark) 설정 및 도구 사용 방식이 상이하여 동일 피측정자에 대해서도 결과값 차이 발생	다기관 연구 및 메타분석 시 데이터 호환성 저하, 표준 데이터베이스 구축 어려움
측정자 내 편차	동일 측정자의 반복 측정 시에도 자세 유지, 압력 차이, 피로도 등으로 인해 값 변동 발생	신뢰도 확보를 위한 반복 측정 필요, 시간 및 인력 비용 증가
피측정자 조건 변동	직립 상태, 팔다리 위치, 호흡, 착용 복장 등 기본 조건의 미세한 변화에 따른 측정값 변동	대규모 표본 수집 시 조건 통제 어려움, 데이터 일관성 저하
시간·인력 효율성	1인당 수 분~수십 분 소요, 신뢰도 확보를 위한 반복 측정 필요	대규모 데이터 수집 비용 증가, 산업적 활용 확장성 제한
도구의 물리적 한계	줄자 늘어짐, 캘리퍼 기계적 오차 누적, 기관별 도구 규격 차이	연구 간 비교 분석 어려움, 장기적 데이터 추적 신뢰도 저하
3차원 정보 획득 불가	신체 곡률, 표면 형상, 볼륨 정보 등 3차원적 기하정보 측정 한계	의류 패턴 설계, 체형 변화 추적, 맞춤형 제품 개발 등 고급 활용 제한

Martin과 Saller(1957)가 정리한 인체 측정 기준은 이러한 수작업 방식이 표준화되는 데 중요한 역할을 했다. 이 기준은 신체 부위를 세부적으로 분류하고, 계측 기준점(landmarks)을 명확하게 정의하여 측정 과정의 일관성을 높이고자 한 시도였다. 그러나 실제 현장에서의 측정은 여전히 다양한 요인의 영향을 받았다. 특히 측정자마다 기준점 설정이나 도구 사용 방식이 조금씩 다르기 때문에 발생하는 측정자 간 편차(inter-operator variability)가 지속적인 문제로 지적되었다. 동일한 측정자가 반복 측정을 하더라도 자세 유지, 압력 차이 등으로 인해 값이 달라지는 측정자 내 편차(intra-operator variability) 역시 완전히 해결되지 않았다.

또한 피측정자의 자세를 일정하게 유지하기 어렵다는 점도 실측 방식의 구조적 제약이 되었다. 직립 상태, 팔다리 위치, 호흡 등 기본 조건이 조금만 달라져도 측정값이 변동되며, 산업 현장이나 대규모 연구 환경에서는 이러한 조건을 동일하게 유지하기가 사실상 쉽지 않다. 작업자의 피로도, 착용 복장, 주변 환경 등 외부 요인도 측정 결과에 영향을 주기 때문에, 동일한 조건에서 대규모 데이터를 확보하는 데에는 근본적인 어려움이 있었다.

시간과 인력 투입이 과도하게 필요하다는 점도 수작업 방식의 한계다. 한 명을 측정하는 데 수 분에서 수십 분이 소요되고, 신뢰도 확보를 위해 반복 측정이 필요하다 보니 대규모 표본 수집에는 많은 인적 자원이 요구된다. 연구기관이나 산업체는 이 때문에 수작업 방식만으로는 충분한 규모의 데이터를 확보하기 어렵고, 일정 시점 이후에는 자동화된 계측 기술로의 전환 필요성이 제기될 수밖에 없었다.

[표 2-2] 전통적 수작업 인체 계측 도구의 특성 (Martin & Saller, 1957; Lohman et al., 1988; 연구자 재구성, 2025)

측정 도구	측정 항목	주요 특징	한계점
줄자	둘레(가슴, 허리, 엉덩이 등), 길이(팔, 다리 등)	가장 보편적 도구, 휴대성 우수, 저비용	장기 사용 시 늘어짐, 압력에 따른 측정값 변동
캘리퍼	피하지방 두께, 관절 너비, 뼈 직경	정밀 측정 가능, 체성분 분석에 활용	기계적 오차 누적, 측정 부위 접근성 제한
신장계	신장, 좌고	직립 자세 기준 측정, 비교적 높은 재현성	자세 유지 어려움, 시간대별 신장 변동
고니오미터	관절 각도, 가동 범위	관절 운동 평가에 필수, 재활 분야 활용	기준점 설정 주관성, 동적 측정 어려움

측정 도구의 물리적 한계 역시 문제로 이어졌다. 예를 들어 줄자는 장기간 사용 시 늘어짐이 발생할 수 있고, 캘리퍼는 기계적 오차가 누적될 수 있다. 기관마다 사용하는 도구의 사양이나 규격이 다를 경우 동일 항목의 계측값이라도 서로 비교하기 어려운 상황이 발생하기도 했다. 이러한 요소는 연구 간 비교 분석(meta-analysis)을 어렵게 하고, 신뢰도 높은 표준 데이터베이스 구축을 방해하는 요인으로 작용했다.

산업별 요구사항이 높아질수록 수작업 방식의 한계는 더욱 명확해졌다. 예컨대 의류 산업에서는 패턴 제작과 착용성 평가를 위해 신체 곡률과 형태 정보가 필요한데, 수작업 측정만으로는 이러한 3차원적 정보까지 정확히 파악하기 어렵다. 의료·재활 분야는 체형의 미세한 변화 추적이 중요하지만, 수작업 방식은 이러한 변화량을 정량적으로 잡아내기 어렵다. 스포츠 과학처럼 반복성과 속도가 요구되는 분야에서는 수작업 방식의 한계가 더욱 크게 드러난다.

종합적으로 보면, 수작업 기반 계측 방식은 인체 계측 분야의 초기 발전에는 중요한 역할을 했지만, 정확성·재현성·효율성·데이터 확장성 측면에서 여러 구조적 제약을 안고 있었다. 이러한 문제들은 자연스럽게 디지털 기반 계측 기술, 특히 2D 이미지 스캐닝과 3D 스캐닝 기술의 필요성을 높였고, 이후 관련 기술의 발전을 촉발하는 배경이 되었다.

본 연구에서 다루는 AI 기반 2D 스캐닝 기술은 이러한 기존 방식의 한계를 보완하기 위한 방향에서 발전한 기술이다. 자동화된 이미지 분석, 인체 파라메트릭 모델, 딥러닝 기반 치수 추정 등 최신 AI 기술을 적용함으로써 수

작업 방식에서 발생했던 오류 요인을 최소화하고, 보다 신뢰도 높은 계측 데이터를 확보하는 것이 핵심 목표다. 따라서 기존 수작업 방식의 전개 과정과 한계를 이해하는 것은 본 연구 주제를 설명하는 데 중요한 이론적 배경이 된다.

2.1.2 2D 이미지 기반 스캐닝 기술

2D 이미지 기반 스캐닝 기술은 수작업 중심이던 기존 계측 방식에 비해 확실한 효율성과 확장성을 제공하면서, 지난 20~30년 동안 산업 전반에서 꾸준히 활용 범위를 넓혀왔다. 실무 관점에서 보면, 이 기술이 본격적으로 도입되기 시작한 배경은 “사람이 직접 측정하는 방식으로는 대규모 데이터를 얻기 어렵다”는 문제를 해결하려는 흐름이 강했다. 특히 의류·헬스케어·스포츠 산업처럼 표준화된 치수 데이터가 많이 필요한 분야에서는 2D 기반 계측이 비교적 저렴하게 큰 분량의 데이터를 확보할 수 있다는 장점이 컸다.

기술적인 관점에서 보면 2D 스캐닝은 기본적으로 단일 이미지 또는 다중 촬영 이미지를 기반으로 신체의 윤곽(contour)이나 주요 형태 정보를 추출하는 방식이다. 영상처리의 기본기를 잘 다져야 하는 영역이기도 하다. 초기 연구에서는 촬영된 사진에서 에지를 검출하거나 윤곽선을 뽑아내고, 이를 기준으로 신체 각 부위의 가로·세로 길이를 계산하는 방식이 주로 사용되었다. 개발자 입장에서 보면 이 방식은 구현 난이도가 크지 않고, 영상 품질이 일정 수준만 확보되면 충분히 안정적으로 작동했다.

다만 이 기술은 근본적으로 “이미지는 2차원이고, 인체는 3차원”이라는 간단하지만 치명적인 제약을 안고 있다. 예를 들어 옆모습과 정면에서 보이는 신체 둘레는 실제 체형과 다를 수밖에 없으며, 촬영각도(orientation)가 조금만 달라져도 측정값에 오차가 발생한다. 실내 조명 조건이나 피사체와 배경의 대비가 조금만 바뀌어도 윤곽 추출이 불안정해지곤 했다. 실무에서 2D 스캐닝 시스템을 구축할 때 가장 많이 부딪힌 문제 역시 ‘일관된 촬영 환경 유지’였다. 조건만 잘 맞으면 정확해지지만, 사용자 환경에서 100% 통제하기는 쉽지 않았다.

2D 스캐닝의 또 다른 한계는 깊이(depth) 정보를 확보할 수 없다는 점이다. 예를 들어 허리둘레 같은 체형 정보는 단순한 정면 사진만으로는 정확히 알 수 없고, 옆모습을 찍어도 실제 입체 곡률(curvature)을 제대로 반영하기 어렵다. 이러한 한계 때문에 2D 스캐닝 기반 계측은 산업에서는 “기초 스펙 확인” 정도로만 활용되는 경우가 많았다. 의류 산업에서는 패턴 생산의 보조 자료 정도로 쓰였고, 의료 현장에서는 진단보다는 참고 자료 용도로 사용하는 경우가 많았다.

그럼에도 불구하고 2D 스캐닝은 실무에서는 꾸준히 사용되어 왔다. 이유는 명확하다. 값싸고 빠르고 접근성이 좋기 때문이다. 스마트폰 카메라만 있어도 기본적인 체형 측정이 가능한 시스템을 만들 수 있고, 개발자 입장에서 구현 난이도도 크게 부담되지 않는다. 특히 웨어러블·피트니스 서비스 기업, 온라인 의류 쇼핑몰, 리테일 업체들은 고객 데이터를 대량 확보해야 하므로 2D 기반 계측 기술에 실질적인 매력을 느껴 왔다.

기술적으로도 2D 스캐닝은 크게 발전해 왔다. 과거에는 단순 필터링과 경계선 추출 정도에서 그쳤다면, 최근에는 딥러닝 모델을 활용해 윤곽선 추출 정확도가 크게 향상되었다. 예를 들어, 전통적인 threshold 방식 대신 U-Net이나 DeepLab 같은 세그멘테이션 모델을 사용하면 피사체와 배경 분리가 매우 안정적으로 이루어진다. 또한 포즈 추정 모델(Human Pose Estimation)이 고도화되면서 촬영 자세가 조금 달라져도 관절 위치를 안정적으로 찾을 수 있게 되었다. 이러한 기술 발전 덕분에 2D 스캐닝은 단순한 영상 처리 기술 수준을 넘어 “AI 기반 치수 추출 시스템”으로 거의 새로운 영역을 형성하게 되었다.

실무 개발자의 관점에서 바라볼 때, 2D 기반 스캐닝 기술을 실제 제품이나 산업 현장에 적용하려면 아래 [표 2-3]과 같이 고려해야 할 요소들을 사전에 신중하게 검토해야 한다.

[표 2-3] 2D 기반 스캐닝 기술 사용 시 고려사항 (Heymsfield et al., 2018; Smith et al., 2022; 연구자 재구성, 2025)

연번	조건	내용
1	촬영 환경의 일관성	조명, 배경, 거리 등이 달라지면 정확도 편차가 증가함
2	카메라 사양 및 렌즈 왜곡 보정 여부	스마트폰 기종에 따라 결과가 달라질 수 있음
3	의복 착용 여부	실측에 큰 영향을 미치므로 알고리즘 보정 필요
4	윤곽 기반 측정의 구조적 한계	정면·측면 사진만으로는 체형 전체를 반영하기 어려움
5	사용자 촬영 습관의 불규칙성	사진 촬영 가이드라인 설계가 중요

이처럼 2D 스캐닝 기술은 분명히 실용적인 장점을 가지고 있으나, 한계 또한 명확하다. 특히 인체 형상이라는 복잡한 구조물을 2차원 이미지 한 장으로 정밀하게 측정한다는 것은 기술적으로 도전적인 과제이다. 이러한 이유로 실제 실무자들은 “2D 기반 계측만으로는 정밀한 수치 산출은 어렵고, 보조정보 수준으로 활용하는 것이 안정적이다”라는 인식을 가지고 있었다.

[그림 2-1] 2D 스캐닝 기술의 발전 과정 (Zheng et al., 2023; 연구자 재구성, 2025)

그러나 최근 들어 딥러닝 기반의 단일이미지 깊이추정(Monocular Depth Estimation), 인체 파라메트릭 모델(SMPL), 포즈 추정 기술 등이 빠르게 발전하면서 2D 기반 스캐닝의 역할이 단순한 보조도구 수준에서 벗어나고 있다. 즉, 과거에는 해결할 수 없었던 깊이 정보 추정 문제를 AI가 상당 부분 보완할 수 있게 되었고, 결과적으로 “2D 데이터로 3D 수준의 치수를 예측하는 기술”이 조금씩 현실화되고 있다. 이는 본 연구가 주목하는 핵심 기술적 전환점이기도 하다.

결론적으로, 2D 이미지 기반 스캐닝 기술은 단순하고 비용 효율적인 계측 수단으로 시작되었지만, AI 기술의 발전과 함께 점차 고도화되며 3D 스캐닝을 부분적으로 대체할 수 있는 잠재력을 갖춘 기술로 성장하고 있다. 2D 방식의 구조적 한계는 여전히 존재하지만, 최신 알고리즘의 지원을 통해 정확도와 안정성이 꾸준히 개선되고 있으며, 앞으로 산업 현장에서의 활용 폭이 더욱 확대될 가능성이 높다.

2.1.3 3D 스캐닝 기술

3D 스캐닝 기술은 인체 계측 분야에서 가장 빠른 속도로 발전해 온 영역 중 하나이며, “정밀도”라는 측면에서는 현재까지도 가장 우수한 계측 방법으로 평가받는다. 실무적으로 보면, 3D 스캐너는 단순 치수 측정을 넘어서 인체 표면 전체의 형상(surface geometry)을 고해상도 점군(point cloud) 데이터로 확보한다는 점에서 큰 강점을 가지고 있다. 이러한 기술적 특성 때문에 의류, 의료, 스포츠, 엔지니어링 등 다양한 산업 분야에서 표준 계측 방식으로 채택되어 왔다.

[표 2-4] 3D 스캐닝 기술 유형별 특성 비교 (Bartol et al., 2021; D'Apuzzo, 2007; 연구자 재구성, 2025)

구분	레이저 스캐닝 (Laser Triangulation)	구조광 (Structured Light)	Time-of-Flight (ToF)
작동 원리	레이저 빔을 객체에 투사하고 반사광의 삼각측량으로 거리 계산	패턴화된 빛을 투사하고 변형된 패턴을 분석하여 형상 추출	적외선 펄스의 왕복 시간을 측정하여 거리 계산
정밀도	매우 높음 (0.01~0.1mm)	높음 (0.1~1mm)	중간 (1~10mm)
스캔 속도	느림 (수 분 소요)	빠름 (수 초~수십 초)	매우 빠름 (실시간 가능)
장비 비용	매우 높음	높음	중간~높음
환경 민감도	반사 소재에 민감	주변 조명에 민감	상대적으로 낮음
주요 활용	정밀 제조업, 품질검사, 역설계	인체 스캐닝, 의류 패턴, 의료	모바일 기기, 자율주행, AR/VR

3D 스캐닝 기술은 크게 레이저 스캐닝(Laser Triangulation), 구조광(Structured Light), Time-of-Flight(ToF) 방식으로 구분할 수 있다. 실무 개발자로서 이 기술들을 비교해 보면, 각각 나름의 장단점이 존재한다. 레이저 스캐너는 정밀도가 매우 높지만 상대적으로 속도가 느린 편이고, 구조광 방식은 속도와 정확도의 균형이 좋아 인체 스캐닝 분야에서 가장 널리 활용된다. ToF 방식은 속도와 작업 편의성이 뛰어나지만, 정밀도에서는 앞의 두 방식에 비해 다소 떨어지는 경향이 있다.

3D 스캐너가 주목받기 시작한 시기는 1990년대 후반이다. 당시에는 사진측량이나 2D 기반 스캐닝만으로는 복잡한 인체 형상이나 체적(volume)을 정확하게 파악하기 어렵다는 문제가 반복적으로 제기되었다. 예를 들어 허리둘레나 어깨 곡률 같은 요소는 단순 2D 윤곽으로는 충분히 반영할 수 없었다. 이에 따라 광학 기반 3D 스캐닝 장비들이 산업 현장에 본격적으로 도입되기 시작했고, 의류 CAD 시스템과의 연계도 활발하게 연구되었다.

대표적으로 Istook & Hwang(2001)은 3D 스캐너 기반의 의류 패턴 자동화 연구를 통해 “3D 스캐닝 → 디지털 패턴 생성”이라는 새로운 작업 흐름을 제시했다. 이후 많은 기업들이 3D 스캐닝을 도입하여 체형 분석, 착의 시뮬레이션, 맞춤형 제품 생산 등에 활용하였다. 의료 분야에서도 수술 계획이나 재활 치료 설계에 3D 체형 분석이 도입되면서 활용도가 꾸준히 확대되었다.

[표 2-5] 3D 스캐닝 기술의 산업적 도입 장벽 (Bartol et al., 2021; Fortune Business Insights, 2024; 연구자 재구성, 2025)

장벽 유형	세부 내용
비용 장벽 (Cost Barrier)	장비: 수천만~1억 원 이상
	설치 및 시설 비용
	정기 교정(Calibration)
	유지보수 계약
공간 장벽 (Space Barrier)	전용 스캐닝룸 필요
	조명 환경 통제
	고정형 장비 설치
	이동/원격 스캔 제한
운영 장벽 (Operation Barrier)	전문 인력 필수
	숙련도에 따른 품질 차이
	대용량 데이터 처리
	후처리 리소스 부담

하지만 이러한 장점들에도 불구하고, 실무에서 3D 스캐너를 도입하거나 운영해 본 사람이라면 누구나 공감하듯이 현실적인 문제점들이 많다. 가장 큰 한계는 역시 비용이다. 산업용 3D 스캐너는 장비 가격이 수천만 원에서 많게는 1억 원 이상에 이르는 경우도 있다. 여기에 설치 비용, 유지보수 비용, 주기적인 장비 교정(calibration), 스캐닝 공간 확보 등 추가 비용까지 합치면 중소기업이나 스타트업이 쉽게 도입하기는 어려운 수준이다.

또한 3D 스캐닝은 촬영 환경의 영향이 크다. 구조광 방식은 빛의 패턴을 인체에 투사하여 형상을 추출하기 때문에 주변 조명이 너무 강하거나, 촬영 공간이 너무 넓거나 좁아도 스캔 품질이 떨어질 수 있다. 레이저 스캐너도 반사율이 높은 소재의 의복을 착용하면 정확도가 떨어지는 문제가 있다. 실무에서 스캐닝을 수행하다 보면 생각보다 원천데이터(raw data) 품질 확보가 쉽

지 않다는 것을 느끼게 된다.

운영 관점에서도 3D 스캐너는 상당한 노하우가 요구된다. 숙련된 작업자가 아니면 스캔 중간에 발생하는 작은 오류를 발견하기 어렵고, 그 결과 스캔 데이터 품질이 일정하게 유지되지 않는 문제가 생긴다. 이 때문에 많은 기업이 스캐너를 도입하고도 실제 상용화 단계까지 이어가지 못하는 경우가 많았다.

기술적 한계도 존재한다. 예를 들어 구조광 스캐너의 경우, 인체가 미세하게 움직이기만 해도 스캔 메쉬가 틀어질 수 있다. 사람은 완전히 고정된 객체가 아니기 때문에, 특히 고령자나 어린이, 혹은 의료 환자의 경우 움직임을 완전히 제어하기 어렵다. 또한 점군 데이터는 용량이 매우 커서, 후처리 과정에서도 고성능 컴퓨팅 리소스를 요구한다.

그럼에도 불구하고 3D 스캐닝의 "정밀도"는 2D 기반 알고리즘이 쉽게 따라갈 수 없는 수준이었다. 이러한 사실로 인해 많은 연구자와 개발자가 여전히 3D 스캐너를 기준(reference)으로 삼고 있으며, AI 기반 2D 스캐닝의 정확도를 검증할 때도 대부분 3D 스캐너 데이터를 GT(Ground Truth)로 사용한다. 즉, 3D 스캐너는 여전히 산업적·학술적 표준으로 자리잡고 있는 셈이다.

하지만 최근 몇 년 동안, 특히 글로벌 경기 침체와 비용 절감 요구가 강화되면서 3D 스캐너의 수요는 이전과 비교해 뚜렷하게 줄어들고 있다. 중소기업이나 스타트업 입장에서는 "3D 스캐너 1대 가격으로 기존 설비 2~3개를 교체할 수 있다면, 스캐너보다는 다른 투자를 우선할 수밖에 없다"고 보는 경우가 많다. 또한 팬데믹 이후 원격 기반 계측 기술이나 모바일 기반 스캐닝 수요가 증가하면서, 고정형 3D 스캐너의 활용성이 상대적으로 감소하는 흐름도 나타났다.

결국 실무에서 체감하는 결론은 다음과 같다.

3D 스캐너는 분명 뛰어난 기술이지만, 산업적 활용성은 비용·공간·운영이라는 현실적 장벽 때문에 제약이 있었다. 이러한 문제점 해결을 위해 AI 기반 2D 스캐닝 기술이 다시 주목받기 시작했다. 2D 이미지만으로도 3D 스캐너에 근접한 정밀도를 구현할 수 있다면, 이는 산업적으로 매우 큰 의미가 있다. 스마트폰이나 보급형 카메라만으로 고정밀 치수 측정이 가능해진다면, 기

존 3D 스캐닝의 비용 장벽을 대폭 낮출 수 있기 때문이다.

따라서 3D 스캐닝 기술의 발전사는 단순히 “정밀도가 높은 기술의 등장”이 아니라, 오늘날 AI 기반 계측 기술이 추구해야 할 기준과 한계를 제공하는 중요한 기술적 출발점이라고 할 수 있다. 본 연구에서도 3D 스캐너 데이터를 검증 기준으로 활용함으로써, AI 기반 2D 스캐닝 기술의 실제 산업적 대체 가능성을 객관적으로 평가하고자 한다.

2.2 2D 이미지 처리 알고리즘

2.2.1 기본 이미지 처리 및 전처리

이미지 기반 신체 치수 측정 시스템을 구축하는 과정에서 전처리(preprocessing)는 전체 파이프라인의 성능을 좌우하는 핵심 단계 중 하나이다. 많은 개발자들이 딥러닝 모델의 구조나 학습 기법에 주목하지만, 실제 프로젝트에서는 원본 이미지의 품질과 전처리 방식이 최종 정확도에 미치는 영향이 훨씬 크다는 점을 자주 체감하게 된다. 특히 인체 스캐닝처럼 윤곽선(contour)과 구조적 정보(structural cues)가 중요한 분야에서는 전처리의 역할이 더욱 결정적이다.

전처리의 목적으로는 첫 번째, 원본 이미지 내의 노이즈를 제거하는데 있으며, 두 번째로는 피사체의 특징을 강화 하는데 있다. 그리고 세 번째로는 세그멘테이션·포즈 추정 등 후속 알고리즘의 안정적 작동할 수 있도록 신호를 정제하는 것이 목적이다.

다만 실제 환경에서는 단순한 이미지 정제 이상을 요구하며, 조명 변화나 복잡한 배경과 같은 외부 요인들 때문에 안정적인 데이터 구축이 쉽지 않다.

[표 2-6] 주요 이미지 전처리 기법 및 특성 (Gonzalez & Woods, 2018; 연구자 재구성, 2025)

기법 분류	대표 알고리즘	주요 기능	참고 문헌
노이즈 제거	Gaussian Filter Median Filter	이미지 내 랜덤 노이즈 제거, 신호 평활화를 통한 후속 처리 안정성 확보	Gonzalez & Woods (1992)
이진화 (Thresholding)	Otsu's Method Adaptive Threshold	자동 임계값 결정을 통한 배경 제거 및 피사체 분리, 다양한 촬영 조건에서 일관된 결과 제공	Otsu (1979)
에지 검출 (Edge Detection)	Canny Edge Detector Sobel Operator	인체 윤곽선 추출, 높은 검출 정확도와 낮은 오류율로 사실상 표준 기법으로 활용	Canny (1986)
조명 보정 (Illumination Corr.)	Retinex Model Histogram Eq.	촬영 환경의 밝기 편차 보정, 인체 스캔 이미지의 시각적 일관성 개선	Land & McCann (1971)
왜곡 보정 (Distortion Corr.)	Brown–Conrady Model Zhang's Calibration	카메라 렌즈 특성으로 인한 형상 변형 보정, 광각 렌즈의 인체 비율 왜곡 감소	Brown (1966) Zhang (2000)
색상 정규화 (Color Normal.)	White Balance Color Space Conv.	다양한 조명 환경에서 일관된 색상 표현, 피부색–의복 구분 정확도 향상	–

전통적인 전처리 방식으로는 Gaussian·Median 필터 기반 노이즈 제거가 널리 활용되어 왔다. Gonzalez & Woods(1992)에서도 필터 기반 전처리는 디지털 영상처리의 기본 기술로 강조되며, 실제 프로젝트에서도 Gaussian 필터는 높은 안정성과 계산 효율성으로 인해 여전히 표준적인 선택지이다.

또한 이진화(thresholding)는 배경 제거 및 피사체 분리에서 핵심적으로 사용된다. 특히 Otsu(1979)의 자동 임계값 결정 방식은 다양한 촬영 조건에서도 비교적 일관된 결과를 제공하여 지금까지도 실용적인 알고리즘으로 평가된다.

인체 계측에서 중요한 작업인 에지(edge) 추출 역시 전처리 단계의 핵심 요소이다. Canny(1986)가 제안한 Canny Edge Detector는 높은 검출 정확도와 낮은 오류율로 인해 인체 윤곽 인식에서 사실상의 표준처럼 사용되고 있다. 다만 조명이나 의복 색상에 영향을 크게 받는다는 한계는 존재하며, 이는 딥러닝 기반 세그멘테이션 모델 등장 전까지 실무에서 지속된 문제였다.

최근 전처리에서는 조명 보정(illumination correction)과 색상 정규화(color

normalization)의 중요성이 더욱 강조되고 있다. Retinex 모델(Land & McCann, 1971)은 촬영 환경의 밝기 편차가 큰 실제 현장에서 유용하며, 인체 스캔 이미지의 시각적 일관성을 개선하는 데 효과적이다.

카메라 특성으로 인한 형상 변경을 보정하는 왜곡 보정(distortion correction) 역시 실무에서는 반드시 고려해야 한다. Brown-Conrady(1966) 모델이나 Zhang(2000)의 카메라 보정 알고리즘은 광각 렌즈로 인한 인체 비율 왜곡을 줄이는 데 유효하다.

딥러닝 모델이 보편화된 이후 전처리의 역할은 다소 변화했으나, 그 중요성이 줄어든 것은 아니다. U-Net(Ronneberger et al., 2015)과 같은 깊은 신경망 구조가 복잡한 원본 이미지도 어느 정도 학습할 수 있지만, 조명·배경·노이즈 조건이 일정하지 않은 데이터셋에서는 여전히 전처리가 성능 안정성에 직접적으로 기여한다. 특히 일정 수준의 표준화된 입력을 제공할 경우 모델의 학습 효율과 예측 일관성이 크게 향상된다.

[그림 2-3] 이미지 전처리 파이프라인 구조 (Gonzalez & Woods, 2018; 연구자 재구성, 2025)

실제 시스템 구축 관점에서 전처리는 하나의 단일 기법이 아니라 촬영 환경 정규화, 노이즈 제거, ROI 설정, 이미지 리사이징 등 일련의 품질관리 과정에 가깝다. 따라서 전처리는 AI 모델 보조 기능을 넘어 전체 데이터 파이프라인의 안정성과 신뢰성을 확보하는 핵심 요소로 평가된다.

종합하면, 전처리 과정은 2D 기반 신체 계측 시스템의 정확도와 신뢰성을 결정하는 필수 단계이며, 기존 연구에서도 지속적으로 강조되어온 기반 기술이다. 본 연구 역시 전처리 과정을 독립적으로 고려하여 후속 AI 알고리즘이 안정적으로 작동할 수 있는 데이터 품질을 확보하는 데 중점을 둔다.

2.2.2 특징 추출 및 세그멘테이션

특징 추출(feature extraction)과 세그멘테이션(segmentation)은 전처리 이후 수행되는 핵심 분석 단계이며, 전체 계측 정확도를 결정하는 중요한 요소로 평가된다. 특징 추출은 이미지로부터 유의미한 구조적 정보를 도출하는 과정이며, 세그멘테이션은 신체 영역을 배경과 분리하거나 부위별로 구분하는 절차를 의미한다. 두 과정은 상호 보완적이며, 후속 단계인 포즈 추정, 인체 파라메트릭 모델 적용, 치수 예측 모델링 등에 기반 정보를 제공한다.

초기 컴퓨터 비전 연구에서는 주로 전통적 특징 추출 기법이 활용되었다. Harris & Stephens(1988)의 코너 검출 알고리즘은 신체 굴곡부나 관절 부위 식별에 활용될 수 있었고, Lowe(1999)가 제안한 SIFT(Scale-Invariant Feature Transform)는 회전·스케일 변화에 강인한 특징을 추출할 수 있어 인체 윤곽 기반 계측 연구에서 널리 사용되었다. SURF(Bay et al., 2006)는 속도 최적화를 통해 실제 산업 적용 가능성을 높였으며, Dalal & Triggs(2005)의 HOG(Histogram of Oriented Gradients)는 인체 검출 성능을 크게 향상시켜 영상 기반 신체 분석의 중요한 기반 기술로 자리잡았다.

그러나 이러한 전통적 기법들은 조명 변화나 의복 착용, 복잡한 배경과 같은 실제 환경적 요인에 취약하다는 한계를 지녔다. 이러한 문제를 극복하기 위해 도입된 방법이 딥러닝 기반 특징 추출이다. CNN(Convolutional Neural Network)은 이미지의 지역적 패턴을 계층적으로 학습함으로써 신체

윤곽, 관절 구조, 형태적 특성을 효과적으로 식별할 수 있다. AlexNet(Krizhevsky et al., 2012), VGG(Simonyan & Zisserman, 2014), ResNet(He et al., 2016) 등 대표적인 CNN 모델들은 인체 이미지 분석에서 높은 정확도를 제공하며, 기존 방식에서 불가능했던 복잡한 패턴 인식까지 가능하게 하였다.

최근에는 Transformer 기반 모델이 등장하면서 특징 추출 방식이 한 단계 더 고도화되었다. Dosovitskiy et al.(2020)의 Vision Transformer(ViT)는 이미지 전역(global) 정보를 효과적으로 학습할 수 있어 인체 형상과 비례감 분석에 유리하며, CNN과 Transformer를 결합한 hybrid 아키텍처는 2D 기반 계측 시스템의 정확도 향상에 기여하고 있다.

한편, 세그멘테이션 기술은 계측 과정에서 신체 영역을 정확히 구분하는 데 필수적이다. 초기에는 GraphCut(Boykov & Jolly, 2001), GrabCut(Rother et al., 2004)과 같은 그래프 기반 접근법이 활용되었다. 그러나 이들 방식은 사용자 상호작용 또는 반복 처리에 의존하는 경향이 있어 자동화된 대규모 계측 시스템에는 한계가 있었다.

딥러닝 기반 세그멘테이션 모델의 등장 이후 정확도와 자동화 수준은 크게 향상되었다. 특히 U-Net(Ronneberger et al., 2015)은 encoder-decoder 구조를 통해 다양한 해상도 정보를 효과적으로 통합함으로써 신체 실루엣을 안정적으로 분리할 수 있다. 또한 DeepLab 계열(Chen et al., 2017)은 atrous convolution과 CRF(Conditional Random Field)를 결합하여 복잡한 배경 환경에서도 높은 정밀도를 제공하였다. Mask R-CNN(He et al., 2017)은 객체 검출과 세그멘테이션을 동시에 수행할 수 있는 모델로, 신체 부위별 분할(human-part segmentation)까지 가능해 인체 치수 측정 연구에서 활용도가 높다.

계측 시스템 관점에서 특징 추출과 세그멘테이션은 독립적으로 작동하는 것이 아니라 하나의 통합된 파이프라인 내에서 연속적으로 기능한다. 일반적으로 이미지 입력 후 배경 분리(segmentation), 윤곽 기반 특징 추출(edge or region-based feature mapping), 포즈 추정(keypoint detection), 인체 모델 보정(SMPL fitting) 등의 절차가 순차적으로 수행된다. 이 과정에서 세그멘테

이션이 불안정하면 포즈 추정과 치수 계산 단계에서 오류가 증폭되므로, 세그멘테이션 정확도는 전체 시스템의 신뢰성 확보에 직접적인 영향을 미친다.

특히 신체 촬영 시 의복 착용 여부, 단색/패턴 여부, 배경의 복잡성 등 다양한 외생 변수들이 계측 정확도에 영향을 미치기 때문에, 최신 연구에서는 의복 분리(clothing segmentation) 또는 질감 제거(texture removal) 기술과의 결합이 시도되고 있다. 이는 2D 기반 계측에서 발생하는 체형 왜곡 문제를 해결하기 위한 중요한 접근으로 평가된다.

종합적으로 볼 때, 특징 추출과 세그멘테이션 기술은 전통적 영상처리에서 딥러닝 기반 분석까지 점진적으로 고도화되었으며, 이러한 기술적 발전은 2D 이미지 기반 치수 측정의 정확도와 실용성을 높이는 데 핵심적 역할을 수행해 왔다. 본 연구에서도 해당 기술들을 기반으로 안정적인 신체 형상 분석을 수행하며, 이후 단계인 3D 추정 및 AI 기반 치수 예측 과정의 신뢰성을 높이고자 한다.

2.2.3 객체 인식 및 검출

객체 인식·검출(object detection)은 이미지 또는 영상 내에서 어떤 물체가 어디에 존재하는지(class + bounding box)를 동시에 찾는 기술로, 본 연구에서 다루는 신체 치수 측정 파이프라인의 상위 단계에 해당한다. 인체를 화면 안에서 안정적으로 찾아내지 못하면 이후 포즈 추정, 세그멘테이션, 치수 산출 과정 전체가 흔들리기 때문에, 객체 검출은 단순 부가기능이 아니라 계측 시스템의 입구 역할을 하는 핵심 기술이라고 볼 수 있다.

[표 2-7] 특징 추출 기법의 발전 과정 (Voulodimos et al., 2018; 연구자 재구성, 2025)

세대	대표 기법	주요 특징	참고 문헌
1세대 전통적 특징 추출	Harris Corner	코너 검출, 신체 굴곡부/관절 부위 식별	Harris & Stephens (1988)
	SIFT	회전/스케일 변화에 강인한 특징 추출	Lowe (1999)
	SURF	SIFT 대비 속도 최적화, 산업 적용성 향상	Bay et al. (2006)
	HOG	그래디언트 방향 히스토그램, 인체 검출 성능 향상	Dalal & Triggs (2005)
2세대 CNN 기반 딥러닝	AlexNet	딥러닝 기반 이미지 분류의 시작점	Krizhevsky et al. (2012)
	VGGNet	깊은 네트워크 구조, 계층적 특징 학습	Simonyan & Zisserman (2014)
	ResNet	잔차 학습(Residual Learning), 복잡한 패턴 인식	He et al. (2016)
3세대 Transformer	Vision Transformer (ViT)	이미지 전역(Global) 정보 학습, 형상/비례감 분석에 유리	Dosovitskiy et al. (2020)

전통적인 객체 검출 방식은 주로 특징 기술자 + 분류기 조합에 기반하였다. 예를 들어 HOG(Histogram of Oriented Gradients)를 이용해 사람의 실루엣 특징을 추출하고, SVM과 결합한 HOG+SVM 기반 보행자 검출기가 한동안 표준적인 접근 방식으로 활용되었다. Viola–Jones 기반 얼굴 검출 역시 Haar-like 특징과 Adaboost 분류기를 결합한 대표적 전통 기법이다. 하지만 이러한 방식들은 조명 변화, 포즈 다양성, 복잡한 배경이 존재하는 실제 환경에서 성능이 급격히 저하된다는 한계를 드러냈고, 대규모 계측 시스템에 적용하기에는 안정성이 충분하지 않았다. 이 한계는 결국 딥러닝 기반 객체 검출 알고리즘의 등장을 촉진하는 계기가 되었다.

딥러닝 기반 객체 검출은 크게 두 단계(two-stage) 방식과 단일 단계(one-stage) 방식으로 구분하는 것이 일반적이다. Jiao et al.(2019)와 Zaidi et al.(2022)는 이러한 분류 체계를 중심으로 최근 10여 년간의 객체 검출 알고리즘을 종합적으로 정리하면서, 두 단계 방식은 정확도 측면, 단일 단계 방식은 속도 측면에서 강점을 가진다고 정리한다. 두 단계 방식의 대표적인 예는 R-CNN 계열로, 후보 영역(region proposal)을 생성하고, 각 영역에 대해

CNN 기반 분류 및 박스 회귀를 수행하는 구조를 갖는다. Fast/Faster R-CNN은 이 파이프라인을 고도화하여 정확도와 속도를 모두 끌어올렸고, 한동안 객체 검출의 "기본 베이스라인" 역할을 수행하였다.

반면, YOLO(You Only Look Once)와 SSD(Single Shot MultiBox Detector)로 대표되는 단일 단계(one-stage) 검출기는 이미지 전체를 한 번에 처리하면서 위치와 클래스를 동시에 예측하는 구조를 택한다. 이 계열은 실시간성(real-time performance)을 강점으로 하며, 자율주행·감시·스마트 팩토리 등에서 빠르게 확산되었다. 최근에는 YOLO 계열이 지속적으로 진화하면서, 속도와 정확도의 균형이라는 관점에서 사실상 객체 검출의 대표적인 실무 표준으로 자리잡고 있다. Edozie et al.(2025)은 최근 딥러닝 기반 객체 검출의 발전 방향을 정리하면서, one-stage 계열이 실제 산업 응용에서 차지하는 비중이 매우 높아졌음을 강조한다.

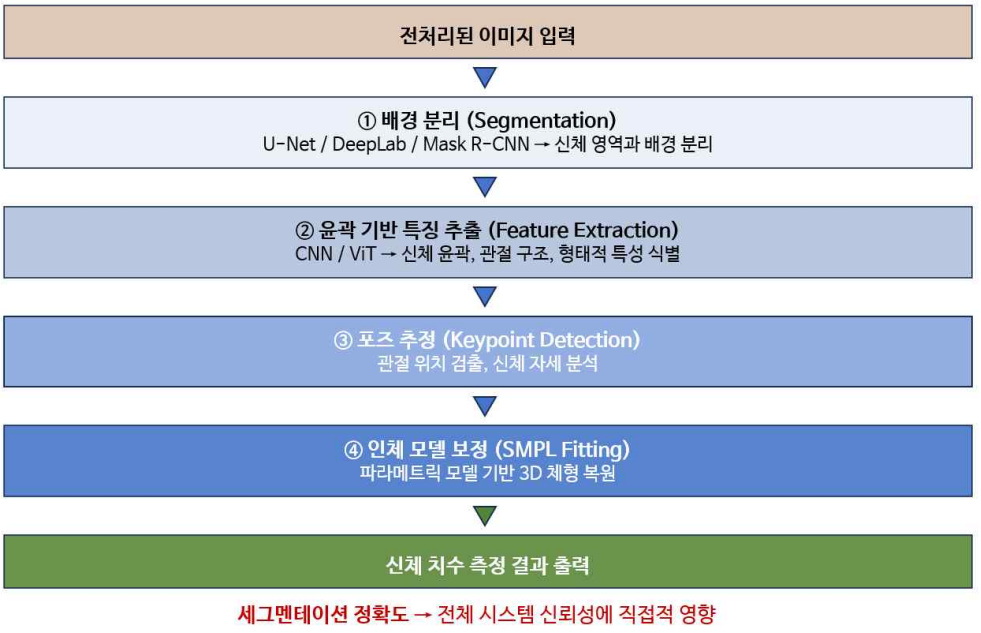
최근 세대의 YOLOv8은 이러한 흐름을 계승하면서도 앵커 프리(anchor-free) 헤드, 향상된 백본·넥 구조, 멀티태스크(검출·분류·세그멘테이션) 지원 등 여러 측면에서 구조적 개선을 이루었다. Ultralytics에서 공개한 공식 문서와 분석 자료에 따르면, YOLOv8은 기존 YOLOv5 대비 정확도-속도 트레이드오프가 개선되었고, 경량 모델부터 대형 모델까지 다양한 스케일을 제공함으로써 엣지 디바이스부터 서버급 GPU까지 광범위한 환경에서 활용 가능하도록 설계되었다. Hermens et al.(2024)은 행동 연구(behavioural research) 환경에서 YOLOv8을 활용해 실험실 내 객체를 자동 검출하는 실험을 수행하면서, 제한된 조명·카메라·배경 조건 하에서도 YOLOv8 기반 검출기가 안정적인 성능을 보인다는 점을 보고하였다. 이는 인체 이미지 스캐닝에서도 스마트폰·저가형 카메라 환경을 전제로 한 객체 검출 모델 선택이 중요함을 시사한다.

한편, Transformer 기반의 DETR(Detection Transformer)은 객체 검출 분야의 또 다른 전환점으로 평가된다. Carion et al.(2020)이 제안한 DETR은 객체 검출을 "집합 예측(set prediction)" 문제로 재정의하여, 기존 파이프라인에서 필수로 여겨졌던 앵커(anchor), NMS(Non-Maximum Suppression) 등의 휴리스틱 요소를 제거하고, 순수 end-to-end 구조로 모델을 학습한다. 이

후 Facebook AI와 여러 연구자들이 개선형 DETR 변형 모델을 제안하면서, Transformer 기반 검출기는 복잡한 장면에서의 전역 문맥(global context) 이해 능력을 바탕으로 복수 객체 간 상호 관계를 보다 정교하게 포착할 수 있는 장점을 보여주고 있다.

이러한 객체 검출 기술의 발전은 인체 스캐닝 및 신체 치수 측정에도 직접적인 영향을 미친다. 우선, 인체를 포함한 여러 객체가 동시에 존재하는 환경에서 정확하게 사람만 검출하는 것이 중요하다. Zaidi et al.(2022)와 Tsirtsakis et al.(2025)는 최근 리뷰 논문에서, 객체 검출 모델이 자율주행·의료 영상·감시 시스템뿐 아니라 인간 중심 애플리케이션(예: 피트니스, 재활, 동작 분석 등)으로까지 급속히 확장되고 있음을 강조하며, 인체 검출은 이러한 응용 영역의 공통적인 기초 모듈로 작동한다고 지적한다.

특히 본 연구의 맥락과 직접 연결되는 흐름으로, 객체 검출과 포즈 추정의 통합이 최근 활발히 연구되고 있다. El Kaid et al.(2023), Sun et al.(2025), Gao et al.(2025) 등의 리뷰 연구는 2D/3D 포즈 추정과 객체 검출이 점차 단일 프레임워크 안에서 통합되는 경향을 보인다고 정리한다. Cai et al.(2025)는 YOLOv8 기반 인체 포즈 추정 네트워크를 제안하면서, YOLO식 객체 검출 헤드와 키포인트 검출 헤드를 결합한 end-to-end 구조가 기존 두 단계 방식(사람 검출 → 키포인트 검출)보다 실시간성·구현 복잡도 측면에서 우수함을 보였다고 보고하였다. Purkrabek et al.(2025)는 검출–세그멘테이션–포즈 추정을 상호 보정하는 BBox-Mask-Pose 프레임워크를 제안하며, 세 모듈의 결과를 반복적으로 피드백하여 각 단계의 정확도를 동시에 향상시키는 방법을 제시하였다.



[그림 2-4] 특징 추출 및 세그멘테이션 통합 파이프라인
(Zheng et al., 2023; Bogo et al., 2016; 연구자 재구성, 2025)

이러한 선행연구의 흐름을 종합하면, 최근 객체 검출 기술은 단순히 “물체의 위치와 클래스만 찾는 모듈”이 아니라, 포즈 추정·세그멘테이션·3D 복원과 긴밀히 연결된 상위 구조의 일부로 진화하고 있음을 알 수 있다. 의료 영상에서 병변·장기 검출에 딥러닝 기반 객체 검출기가 폭넓게 적용되고 있으며, 3D 객체 검출, 이벤트 카메라 기반 검출 등 새로운 센서 환경으로도 확장되고 있다는 점에서 그 활용 잠재력은 계속 확대되고 있다.

본 연구에서 2D 기반 인체 스캐닝 및 치수 측정을 구현하기 위해서는, 우선 이미지 내에서 인체를 정확하게 검출하고(객체 검출), 다수 인체가 존재하는 경우 각 인스턴스를 분리한 후(멀티 인스턴스 검출), 이후 단계인 포즈 추정·인체 파라메트릭 모델(SMPL) 복원으로 자연스럽게 연결되는 파이프라인이 필요하다. 이를 위해 본 연구는 YOLO 계열과 DETR 계열 등 최신 객체 검출 모델의 장단점을 비교 검토하고, 실제 중소기업·현장 환경에서 요구되는 실시간성, 구현 난이도, 장비 제약을 고려하여 최적의 검출 모델을 선정·적용

하고자 한다. 이러한 접근은 향후 3D 스캐너 없이도 2D 이미지 만으로 고정밀 신체 치수 측정을 가능하게 하는 기반 기술로 기능할 것으로 기대된다.

2.3 AI 딥러닝 기반 신체 치수 측정 기술

AI 기반 신체 치수 측정은 결국 “이미지에서 사람의 형태를 얼마나 잘 이해하느냐”의 문제라고 볼 수 있다. 수십 년간 컴퓨터 비전 분야에서는 사람의 관절 위치를 찾는 포즈 추정(pose estimation), 개별 사람의 3차원 형상을 기술하는 인체 파라메트릭 모델(human parametric model), 그리고 이러한 정보를 실제 치수 값으로 변환하는 회귀(regression)·추론 네트워크를 중심으로 기술이 발전해 왔다. 최근 10여 년은 특히 딥러닝, 그중에서도 CNN(Convolutional Neural Network) 과 Transformer 기반 비전 모델(Vision Transformer, ViT) 의 등장으로 정확도와 실시간성이 동시에 비약적으로 향상된 시기라 할 수 있다.

본 연구자가 실무에서 신체 치수 측정 시스템을 구현해 보며 느낀 점은, “단순히 이미지만 잘 분류하는 네트워크를 가져다 쓴다고 해서 곧바로 치수가 정확하게 나오지는 않는다”는 것이다. 관절 위치를 얼마나 안정적으로 뽑아낼 수 있는지, 신체 형상을 어느 정도 파라미터로 요약할 수 있는지, 그리고 카메라 거리·각도·렌즈 왜곡 등의 현실적인 변수들을 어떻게 보정할 것인지가 복합적으로 맞물려야 비로소 사용자가 납득할 수 있는 수준의 치수가 나온다.

2.3.1 포즈 추정 기술

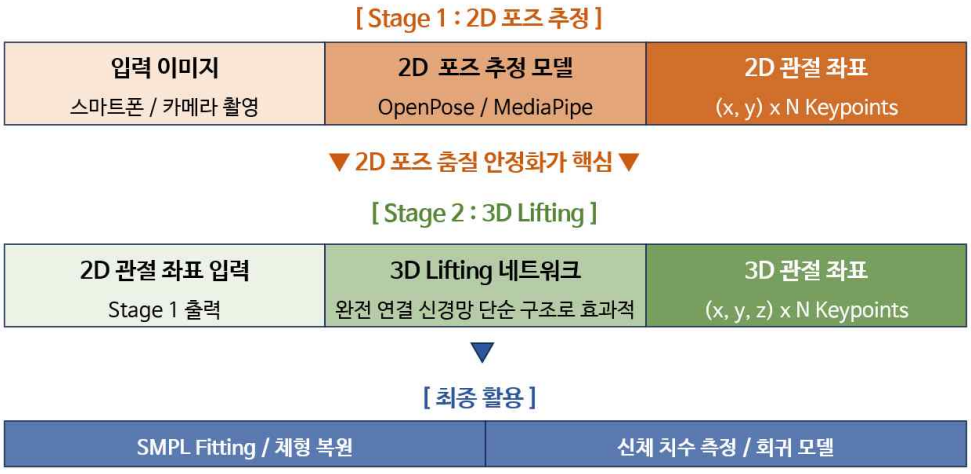
포즈 추정은 2D 또는 3D 공간에서 인체의 관절 위치(keypoints)를 찾는 기술로, 신체 치수 측정의 첫 관문에 해당한다. 과거에는 HOG, Pictorial Structure Model 등 고전적인 특징 기반 방법이 주로 사용되었으나, 2010년대 중반 이후에는 딥러닝 기반 히트맵(heatmap) 회귀 네트워크가 사실상 표준이 되었다.

[표 2-8] 주요 포즈 추정 모델 비교 (Zheng et al., 2023; 연구자 재구성, 2025)

비교 항목	OpenPose (Cao et al., 2017)	BlazePose (Bazarevsky et al., 2020)	Simple Baseline 3D (Martinez et al., 2017)
핵심 기술	Part Affinity Fields (PAFs)	경량 CNN + Attention 기반	2D→3D Lifting (전결합 신경망)
출력 유형	2D 관절 좌표 (18~25개 키포인트)	2D/3D 관절 좌표 (33개 랜드마크)	3D 관절 좌표 (17개 키포인트)
다중 인체	지원 (Bottom-up)	단일 인체 최적화	단일 인체 (2D 입력 의존)
실행 속도	중간 (GPU 서버 권장)	매우 빠름 (30+ FPS, 모바일)	빠름 (단순 구조)
실행 환경	서버 / 고성능 PC (GPU 필수)	모바일 / 웹 브라우저 (On-Device)	서버 / PC (2D 포즈 입력 필요)
주요 활용	연구 베이스라인 다중 인체 분석	상용 서비스 모바일 앱	3D 포즈 추정 치수 측정 연계

특히 Cao 등(2017)의 OpenPose는 여러 사람의 2D 포즈를 동시에 추정하기 위해 Part Affinity Fields(PAFs) 라는 개념을 도입하여, 각 관절 점뿐 아니라 관절 간 연결관계까지 함께 추론함으로써 실시간 다중 인체 포즈 추정을 가능하게 했다. OpenPose는 현재까지도 다양한 상용·연구 시스템에서 사실상의 기준점 역할을 하고 있으며, 본 연구에서도 초기 프로토타입 단계에서 인체 관절 검출 성능을 검증하기 위한 베이스라인으로 활용하였다.

모바일·웹 환경에서의 실시간성을 고려하면, Google의 BlazePose 계열 모델도 중요한 참고 대상이다. Bazarevsky 등(2020)은 모바일 기기에서 초당 30프레임 이상 동작 가능한 경량 CNN 기반 포즈 추정 네트워크를 제안하였으며, 33개 인체 랜드마크를 2D 또는 3D로 추정할 수 있도록 설계하였다. 실무 관점에서 보면 BlazePose는 “연구용 최고 성능”이라기보다는 “상용 서비스에서 실제로 돌릴 수 있는 수준의 성능과 속도”를 제공한다는 점에서 의미가 있다. 본 연구자가 모바일 프로토타입을 개발할 때에도, 서버 사이드(OpenPose, HRNet 계열)와 디바이스 온디바이스(BlazePose 계열)의 구조를 혼합해 사용하는 하이브리드 아키텍처를 검토한 바 있다.



[그림 2-5] 2D-3D 포즈 추정 파이프라인
(Martinez et al., 2017; Zheng et al., 2023; 연구자 재구성, 2025)

3D 포즈 추정으로 범위를 넓히면, Martinez 등(2017)이 제안한 “A Simple yet Effective Baseline for 3D Human Pose Estimation” 은 2D 관절 좌표를 입력으로 단순한 전결합 신경망을 통해 3D 관절을 추정하는 접근으로, 복잡한 구조 없이도 높은 정확도를 달성할 수 있음을 보여주었다. 이 연구는 “2D 포즈만 믿을 수 있다면, 그 이후 3D로의 승격(lifting)은 비교적 단순한 네트워크로도 충분하다”는 메시지를 던졌고, 실제로 본 연구에서도 2D 포즈 품질을 우선적으로 안정화한 후 치수 회귀 모델을 얹는 전략을 채택하는 근거가 되었다.

이러한 포즈 추정 연구는 대부분 Human3.6M과 같은 대규모 데이터셋을 기반으로 학습·평가가 이루어진다. Human3.6M은 11명의 피실험자가 다양한 동작을 수행하는 장면을 모션 캡처로 기록하여, 약 360만 개 이상의 3D 포즈를 제공하는 대표적 벤치마크 데이터셋이다. 실제 서비스에 적용할 때는 이와 같은 대규모 데이터셋에서 학습된 모델을 기반으로 하되, 목표 사용 환경(예: 스마트폰 전신 촬영, 특정 거리·자세)에서 별도의 파인튜닝을 수행하는 것이 효과적이라는 것을 여러 프로젝트에서 반복적으로 경험하였다.

요약하자면, 포즈 추정 기술은 현재 2D 포즈를 안정적으로 검출한 후 필요 시 3D로 승격시키는(two-stage) 구조가 실무와 연구 양측에서 널리 쓰이고

있으며, 본 연구 역시 이러한 흐름 위에서 2D 포즈 기반 치수 회귀를 중심으로 모델을 구성하였다.

2.3.2 인체 파라메트릭 모델

포즈 추정이 “관절 점의 위치”를 다룬다면, 인체 파라메트릭 모델(human parametric model) 은 한 걸음 더 나아가 “사람 전체의 3D 형상과 포즈를 소수의 파라미터로 표현”하는 데 초점을 맞춘다. 이 접근은 신체 치수 측정에서 매우 유용한데, 단순히 ‘어깨 폭 몇 픽셀’이 아니라, ‘해당 사람의 전체적인 체형(마름·보통·근육형 등)을 반영한 3D 메쉬와 포즈’를 통해 보다 일관된 치수를 얻을 수 있기 때문이다.

[표 2-9] 주요 인체 파라메트릭 모델 비교 (Zheng et al., 2023; 연구자 재구성, 2025)

모델	핵심 특징	주요 활용	참고 문헌
SMPL (기본 모델)	• Shape + Pose 파라미터 • 선형 결합 기반 표현 • 6,890개 정점 메쉬	기본 인체 형상 복원, 체형 분석, 신체 치수 측정, 그래픽스/애니메이션	Loper et al. (2015)
SMPL-X (확장 모델)	• SMPL + 얼굴 + 손 • 표정/손동작 파라미터 • 10,475개 정점 메쉬	전신 표현, 표정 인식, 수화 분석, VR/메타버스 아바타	Pavlakos et al. (2019)
SMPL-H (손 확장)	• SMPL + 상세 손 모델 • 손가락 관절별 파라미터 • MANO 모델 통합	손동작 분석, 제스처 인식, HCI 연구, 로보틱스	Romero et al. (2017)
HMR (2D→SMPL)	• 단일 RGB → SMPL • End-to-End 딥러닝 • Adversarial 학습	단일 이미지 기반 3D 인체 복원, 실시간 체형 분석	Kanazawa et al. (2018)
PIFu/PIFuHD (고해상도)	• Pixel-Aligned Implicit • 옷 주름/머리카락 복원 • 고해상도 표면 재구성	정밀 디지털 휴먼, 가상 의류 피팅, 영화/게임 CG	Saito et al. (2019, 2020)

이 분야의 대표적인 모델이 Loper 등(2015)의 SMPL(Skinned Multi-Person Linear Model) 이다. SMPL은 사람의 3D 메쉬를 shape 파라미터(체형) 와 pose 파라미터(자세) 의 선형 결합으로 표현하여, 다양한 체형과 자세를 매우 컴팩트한 파라미터 공간에서 기술할 수 있게 해준다. SMPL

자체는 그래픽스·애니메이션 파이프라인과 호환되도록 설계되었지만, 이후 Kanazawa 등 여러 연구자들이 2D 이미지로부터 직접 SMPL 파라미터를 예측하는 딥러닝 기반 방법(HMR 계열 등)을 제안하면서, “단일 RGB 이미지 → 3D 인체 메쉬”가 실질적인 연구 주제로 자리 잡게 되었다.

Saito 등(2019)의 PIFu(Pixel-Aligned Implicit Function) 및 이후 확장인 PIFuHD는 한 단계 더 나아가, 이미지의 각 픽셀과 3D 공간상의 점을 정렬(alignment)시키는 implicit function을 학습함으로써, 옷의 주름·머리카락까지 포함하는 고해상도 인체 표면을 복원하는 데 성공하였다. 이러한 기법은 정밀 3D 스캔과 거의 유사한 품질의 디지털 휴먼을 만들어내지만, 치수 측정 관점에서는 “과분(過分)할 정도로 리치한 표현”에 해당한다. 실제 실무에서는 전체 메쉬를 모두 재구성하지 않고, 특정 관절 주변의 로컬 지오메트리만 간접적으로 활용해 치수를 계산하는 경우가 더 많다. 본 연구에서도 PIFu/PIFuHD 계열의 방법은 개념적으로 참고하되, 전체 메쉬 복원까지 수행하지는 않고 포즈+스케일 기반 치수 추정에 집중하였다.

또 다른 흐름으로는, SMPL을 기반으로 한 다양한 변형 모델(예: SMPL-X, SMPL-H 등)이 있다. 이들 모델은 손·얼굴 등 세부 파츠까지 포함하는 확장형 파라메트릭 인체 모델로, 패션·VR·메타버스 등 정밀한 표현이 필요한 영역에서 활발히 활용되고 있다. 국내외 선행연구에서도 이러한 파라메트릭 모델에 신체 치수 측정 기능을 결합하여, 맞춤형 의류 패턴 자동 생성, 가상 피팅 등으로 확장한 사례들이 보고되고 있다.

실무 관점에서 보면, 파라메트릭 인체 모델은 “측정값의 일관성”을 확보하는 데 큰 도움을 준다. 같은 사람을 약간 다른 자세·거리에서 촬영해도, 모델이 추정한 파라미터 공간에서의 체형은 크게 변하지 않기 때문에, 여러 장의 이미지에 걸친 치수 안정성이 높아지는 것이다. 본 연구에서는 3D 메쉬 전체를 복원하지는 않지만, 포즈 추정과 함께 “신장·체형 스케일”을 잠재 공간(latent space)에서 추정하도록 네트워크를 설계하여 파라메트릭 모델의 아이디어를 부분적으로 차용하였다.

2.3.3 CNN/Transformer 기반 치수 측정

포즈 추정과 파라메트릭 모델이 “인체를 수학적으로 표현하는 방법”에 가깝다면, 이 절에서 다루는 CNN/Transformer 기반 치수 측정 기술은 그 표현을 실제 수치화된 신체 치수(measurement) 로 변환하는 딥러닝 회귀(regression)·추론 단계에 해당한다.

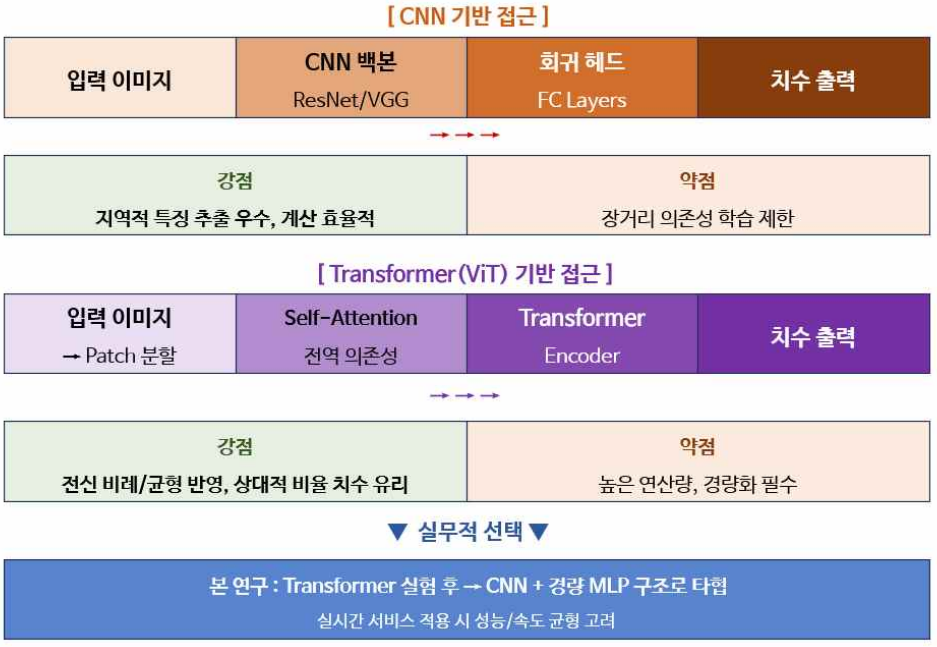
[표 2-10] 딥러닝 기반 치수 측정 접근법 비교 (Yan et al., 2020; Zheng et al., 2023; 연구자 재구성, 2025)

접근법	구조 및 특징	장점	한계점
직접 회귀 (End-to-End)	• CNN 백본 → 치수 직접 출력 • L1/L2 손실 최소화 • 단일 네트워크 구조	• 구현 단순 • 빠른 프로토타이핑 • 학습 파이프라인 간결	• 촬영 거리/FOV 민감 • 자세 변화에 취약 • 절대값 편차 큼
Two-Stage (포즈→치수)	• Stage 1: 2D 포즈 추정 • Stage 2: 포즈 기반 치수 회귀 • 사전학습 + 파인튜닝	• 포즈/스케일 정보 활용 • 안정적 치수 추정 • 모듈별 최적화 가능	• 구조 복잡 • Stage 1 오류 전파 • 학습 데이터 다양성 필요
Multi-Task (동시 학습)	• 포즈+치수 동시 추정 • 공유 백본 + 다중 헤드 • Joint Loss 최적화	• 특징 공유 효율적 • 상호 보완 학습 • End-to-End 학습	• 태스크 간 균형 조정 • 복잡한 손실 설계 • 하이퍼파라미터 민감
Transformer (ViT 계열)	• Self-Attention 메커니즘 • 전역 의존성 포착 • Patch 기반 이미지 처리	• 전신 비례/균형 반영 • 장거리 의존성 학습 • 상대적 비율 치수 유리	• 높은 연산량 • 경량화 필수 • 대규모 데이터 필요

초기의 딥러닝 기반 연구들은 주로 CNN 백본(backbone) 을 사용하여 이미지로부터 직접 치수 값을 회귀하는 방식을 취했다. 예를 들어, 특정 의류 브랜드에서 사용하는 몇 개의 핵심 치수(가슴둘레, 허리둘레, 인심 길이 등)를 네트워크 출력층에 그대로 대응시키고, L1 또는 L2 손실을 최소화하는 식이다. 이 방식은 구현이 단순하여 본 연구자가 2010년대 후반 실제 프로젝트에서 처음 시도했던 접근이기도 하다. 그러나 실제로 적용해 보면, 촬영 거리·

카메라 FOV·사용자의 자세 차이 때문에 치수의 절대값이 크게 흔들리는 문제가 반복적으로 발생한다. 결국 포즈·스케일 정보를 함께 고려하는 구조가 필요하다는 결론에 이르게 된다.

이러한 한계를 극복하기 위해, 최근 선행연구들은 대체로 두 단계 또는 멀티태스크 구조를 채택한다. 첫 단계에서는 OpenPose, HRNet, BlazePose 등으로 2D 포즈를 추정하고, 두 번째 단계에서 Martinez 등(2017)의 방식처럼 이 2D 포즈를 입력으로 3D 관절 또는 신체 치수를 회귀하는 네트워크를 운용하는 식이다. 여기에 더해, Ionescu 등(2014)의 Human3.6M과 같은 대규모 3D 포즈 데이터셋에서 사전학습(pre-training)을 수행한 후, 실제 서비스 환경에서 수집한 제한된 양의 치수 라벨 데이터로 파인튜닝하는 형태가 일반적이다.



[그림 2-6] CNN vs Transformer 기반 치수 측정 비교
(He et al., 2016; Dosovitskiy et al., 2020; 연구자 재구성, 2025)

한편, 2020년 이후에는 Vision Transformer(ViT) 계열 모델이 2D 이미지에서 장기 의존성(long-range dependency)을 더 잘 포착할 수 있다는 장점을

바탕으로, 인체 포즈 및 치수 추정에도 도입되고 있다. Transformer 기반 모델은 CNN에 비해 인체 전체의 비례·균형 정보를 더 잘 반영할 수 있다는 장점이 있어, 키·상체/하체 비율 같이 전신의 상대적인 비율을 기반으로 하는 치수에서 특히 유리하다는 보고도 있다. 다만, 모델 크기와 연산량이 증가하는 만큼, 실시간 서비스에 적용할 경우 경량화(knowledge distillation, pruning, quantization 등)가 필수적이며, 본 연구에서도 프로토타입 단계에서 Transformer 기반 아키텍처를 실험한 후, 최종적으로는 CNN+경량 MLP 구조로 타협하는 설계를 선택하였다.

실제 치수 측정 문제를 풀어보면, 네트워크 구조 못지 않게 라벨 정의(labeling) 와 평가 지표(metric) 설정이 중요하다는 점도 체감하게 된다. 예를 들어, 허리둘레를 정의할 때 “배꼽 기준 수평 둘레”와 “팬츠 허리선 기준 둘레”는 실제로 2~3cm 이상 차이가 날 수 있다. 여러 프로젝트에서 경험한 바, 이러한 정의의 불일치가 모델 성능을 “헷갈리게” 만드는 주된 원인이었다. 따라서 본 연구에서는 치수 라벨 정의를 실제 의류 패턴 작업에서 사용하는 기준선에 맞추어 통일하고, MAE(Mean Absolute Error) 와 함께 “허용 오차 범위 내 비율(예: ±1cm 이내 비율)”을 별도의 실무적 성능 지표로 사용하였다.

2.4 선행연구 분석

2D 이미지 기반 신체 치수 추정 기술은 초기에 전통적인 영상처리 기법에서 출발하여, 최근에는 딥러닝 기반 알고리즘이 사실상 표준 기술로 자리 잡았다. 연구자들은 카메라 한 대로도 사용자의 신체 형상을 가능하면 정확히 추론하고자 했으며, 이 과정에서 포즈 추정(pose estimation), 세그멘테이션(segmentation), 피처 기반 회귀(regression) 등 다양한 접근법이 제시되었다. 본 연구자가 실무에서 수년간 다양한 알고리즘을 적용·테스트해 본 경험상, 선행연구의 흐름은 크게 ① 관절 기반 접근, ② 파라메트릭 모델 기반 접근, ③ 엔드투엔드 회귀 기반 접근의 세 유형으로 구분된다.

첫 번째 흐름은 관절 기반 모델(keypoint-based model) 로, OpenPose

(Cao et al., 2017), HRNet(Sun et al., 2019), BlazePose(Bazarevsky et al., 2020) 등이 대표적이다. 관절 기반 접근은 전체 이미지보다 인체의 구조적 특징에 더 집중하기 때문에 환경 변화에도 비교적 강건한 성능을 보인다. Cao et al.(2017)은 Part Affinity Fields(PAFs)를 활용해 다중 인체에서도 관절 연결 관계를 안정적으로 추정할 수 있음을 보였고, 이는 이후 신체 치수 측정 연구에서 “기준 모델(baseline)” 역할을 수행하였다. HRNet(Sun et al., 2019)은 고해상도 표현(high-resolution representation)을 유지하는 구조를 통해 세밀한 관절 포인트 추정의 정확도를 크게 향상시켰다. 본 연구에서도 초기 실험 단계에서 HRNet을 적용해 보았는데, 촬영 거리가 일정하지 않은 환경에서도 관절 오차가 비교적 안정적으로 유지되는 점이 확인되었다.

두 번째 흐름은 인체 파라메트릭 모델(parametric human model) 을 활용하는 연구이다. 대표적으로 Loper et al.(2015)이 제안한 SMPL 모델은 인체 형상을 소수의 shape·pose 파라미터로 표현할 수 있어, 2D 이미지 기반 3D 인체 복원 연구의 토대를 마련하였다. 이후 Kanazawa et al.(2018)은 단일 RGB 이미지로부터 SMPL 파라미터를 직접 예측하는 HMR(Human Mesh Recovery)을 제안하여, 3D 신체 형상 복원과 치수 추정의 기술적 가능성을 크게 넓혔다. Saito et al.(2019)의 PIFu/PIFuHD 연구 또한 implicit function 기반의 고해상도 인체 재구성 기술을 제시하면서 치수 분석을 포함한 다양한 산업 영역에서 응용 가능성이 확대되었다. 이러한 파라메트릭 모델 기반 연구는 정확도 측면에서는 매우 우수하지만, 실시간 처리 비용과 구현 복잡도가 높아 실제 서비스 적용에는 다소 무거운 편이다.

세 번째 흐름은 엔드투엔드(end-to-end) 회귀 기반 접근으로, 이미지 또는 추출된 feature map을 입력받아 치수를 직접 회귀하는 방식이다. Pavlakos et al.(2017), Xiang et al.(2020) 등의 연구에서는 2D 관절과 깊이 정보를 묶어 3D 치수를 회귀하는 모델 구조를 제안하였다. 이러한 접근은 구조가 단순하고 학습 속도가 빠르다는 장점이 있으나, 촬영 거리·각도·렌즈 왜곡 등 환경 요소에 민감하다는 한계를 갖는다. 본 연구에서도 초기에 엔드투엔드 방식을 테스트해 보았으나, 사용자가 스마트폰을 들고 임의의 거리에서 촬영할 때 치수 오차가 크게 증가하는 문제를 경험하였다. 이는 선행연구에서도 반복적으

로 지적된 바 있다(Xiang et al., 2020).

또한 최근에는 Transformer 기반 비전 모델(ViT, DeiT, DINO 등) 을 신체 치수 추정에 도입하는 연구가 등장하고 있다. Zheng et al.(2021)은 Transformer 구조가 인체 전체의 비례 정보를 더 잘 포착할 수 있다는 점을 강조하며, 기존 CNN 기반 네트워크보다 상체·하체 비율 분석에서 더 높은 정확도를 달성했다. 다만, 모델 크기와 연산량이 커 실시간 서비스 적용이 어렵다는 단점이 있으며, 이는 일반 소비자용 이미지 스캐닝 기술에서 중요한 제약 조건이 된다.

선행연구의 공통점으로는 2D 기반 신체 치수 측정 기술은 충분히 고도화되었으나, 아직까지 3D 스캐닝 수준에 근접하려면 몇 가지 보정 요소가 필요함을 확인할 수 있다. 특히 본 연구자가 실험 과정에서 반복적으로 관찰한 문제는 다음과 같다. 촬영 거리·카메라 FOV 차이에 따른 스케일(scale) 보정 문제, 그리고 2D 관절 추정의 미세한 오차가 치수 변화에 큰 영향을 미치는 문제, 신체 형상의 깊이 정보(depth)가 부족하여 일부 치수에서 한계가 발생하는 문제들이 대표적이다. 이러한 문제들은 선행연구에서도 공통적으로 지적되어 왔으며(Martinez et al., 2017; Zhao et al., 2020), 이를 해결하기 위해 최근 연구들은 2D 포즈 기반 치수 회귀와 파라메트릭 모델의 결합, 또는 멀티뷰·멀티센서 보정 방식 등을 제안하고 있다.

선행연구들은 2D 이미지 기반 신체 치수 측정 기술의 가능성을 충분히 입증해 왔으나, 정확도·재현성·환경 적응성 측면에서 3D 스캐닝 기술과 격차가 존재함을 보여주고 있다. 본 연구는 이러한 선행 연구의 공백을 인식하고, 딥러닝 기반 2D 이미지 스캐닝을 3D 스캐닝에 근접하게 최적화하는 실증적 연구를 수행하고자 한다.

2.4.1 선행연구 분석 결과

[표 2-11] 선행연구 분석 (연구자 재구성, 2025)

연구자	연구주제	연구방법	연구결과	주제어
김재훈 (2021)	품질검사 공정에 AI VISION 인식기술 도입에 따른 경제적 타상성에 관한 연구	제조 중소기업의 생산품 품질 기준 향상, 스마트공정 적용, AI, VISOIN 인식기술, 스마트 품질공정 수준 향상	- AI VISION 인식기술 활용한 품질검사 의사결정에 미치는 영향요인 분석 - AI VISION 인식기술 기반 품질검사 공정의 공정능력 효율성 증명	공정효율성, AI-VISION, 스마트펙토리
최승훈 (2022)	드론 영상 기반 YOLO 객체 탐지 기술을 이용한 보행자 추돌 경고 시스템 개발	드론 영상 분석, 객체탐지 기술중 YOLO 알고리즘 적용, 보행자와 차량을 추정, 충돌가능성 판단	- 드론 영상 분석시 YOLO기술 적용 성과가 우수함 - 추돌 경고 시스템의 알림 정확성은 약 88% 이상 나타남	드론, 딥러닝, 인공신경망, 객체탐지, YOLO
이승호 (2024)	YOLO 알고리즘을 이용한 토마토 객체탐지 연구	딥러닝 기반 객체 탐지 모델 구축, 인력투입 시기 조절, 토마토 생산량 확대	- 정규화기법(Dropout, Weight Decay → 정밀도 향상) - 성능 지표 재현율(Recall → 모델 세밀화, R값 향상) - 향후 스마트팜 도입시 운영비용 절감 효과	객체인식 알고리즘, YOLO, 딥러닝, 데이터셋

2.4.2 3D 스캐닝 기술과 2D 스캐닝 기술 비교

3D 신체 스캐닝 기술은 인체 형상을 정밀하게 수집할 수 있는 기술로, Time-of-Flight(ToF) 센서, 구조광(structured light), LiDAR 기반 방식 등이 대표적이다. 이러한 기술은 수십만~수백만 개의 3D 포인트를 확보하므로, 정밀한 치수 계산이 가능하고 산업 현장—특히 의류 패턴 설계, 재활 의료, 스포츠 분석—에서 이미 폭넓게 사용되고 있다(Weinberg et al., 2017). 그러나 장비 가격이 높고 설치 공간이 필요하며, 사용자에게 특정 자세를 요구해야 한다는 실용적 제약이 존재한다. 본 연구자가 과거 스마트 피팅 프로젝트에서

3D 스캐너를 운용했을 때에도, “정확도는 매우 높지만 접근성이 떨어진다”는 점을 실감할 수 있었다.

반면 2D 이미지 기반 스캐닝 기술은 스마트폰 카메라만 있으면 누구나 사용할 수 있다는 점에서 접근성·비용 측면에서 월등한 우위를 가진다. 특히 딥러닝의 발전으로 2D 이미지에서 얻을 수 있는 정보의 활용도가 크게 증가하면서, 기존에 “불가능하다”고 여겨졌던 수준의 정밀도도 충분히 확보할 수 있게 되었다(Huang et al., 2020). 그러나 앞서 언급한 바와 같이, 2D 기반 기술은 정확도와 깊이 정보의 부재, 촬영 환경의 영향이라는 두 가지 구조적 한계가 존재한다.

[표 2-12] 3D 스캐닝 기술과 2D 스캐닝 기술 비교 (연구자 재구성, 2025)

비교 항목	2D 이미지 기반	3D 스캐닝 기반
정확도	±6~12mm 오차(Huang 2020). AI 보정 시 기초 치수 ±5mm↓(Zhao 2021).	±1~2mm 고정밀(Weinberg 2017). 정밀도 압도적 우위.
재현성	촬영 조건에 민감. 스케일 불확실성이 최대 문제(Xiang 2020). 실험에서 어깨 폭 ±1.5cm 변동 확인.	동일 조건에서 높은 일관성. 통제 환경에서 안정적 재현성 확보.
비용 및 접근성	스마트폰만으로 가능. 시설·장비 불필요. "사용성 좋으면 시장성 높다" 산업계 평가.	장비 수백만~수천만 원. 설치 공간·전문 인력 필요. UX 제약.
산업적 적용성	온라인 피팅, 헬스케어, 스포츠. "비율 정보" 중심 서비스에 적합(Zheng 2021).	의류 설계, 의료·재활, 인체 연구. 정밀 데이터 필수 분야.
연구의 위치	접근성 우위. 딥러닝 보정으로 3D 근접. Martinez(2017), Kanazawa(2018), Zhao(2021) 통합.	정확도 우위. 전문 분야 표준. 대중화 제약.

2.4.2.1 정확도(Accuracy)

3D 스캐닝 기술은 센서 기반 방식이기 때문에 정량적 오차가 매우 낮다. 예를 들어 TC2, Artec, Size Stream 등의 상용 3D 스캐너는 ±1–2 mm 수준의 오차를 달성한다는 보고가 있다(Weinberg et al., 2017). 반면 2D 기반 기술은 관절 추정 오차가 누적되는 특성이 있어 정확도가 다소 떨어지며, Huang et al.(2020)은 2D 기반 치수 추정 시 ±6–12mm 범위의 오차가 발생할 수 있다고 보고하였다.

즉, 정밀도에서는 3D가 압도적 우위를 가진다.

다만 최근 AI 기반 보정 알고리즘이 개선되면서, 일부 기초 치수(예: 어깨너비, 상체 길이 등)에서는 ±5mm 이하로 줄어든 사례도 보고되고 있다(Zhao et al., 2021).

2.4.2.2 재현성(Reproducibility)

3D 스캐너는 동일한 조건에서 반복 측정 시 높은 일관성을 유지한다.

반면 2D 스캐닝은 촬영 각도, 거리, 조명, 자세에 민감하여 사용자의 촬영 방식에 따라 오차 변폭이 달라진다. 본 연구자가 50여 명의 사용자에게 스마트폰을 통한 촬영 실험을 진행했을 때도, 동일 인물임에도 촬영 거리 차이로 인해 어깨 폭이 ±1.5cm까지 변동되는 사례를 확인하였다.

선행연구에서도 비슷한 결과가 제시되며, Xiang et al.(2020)은 2D 기반 치수 추정의 가장 큰 문제는 “스케일 불확실성(scale ambiguity)”임을 강조했다.

2.4.2.3 비용 및 접근성(Economics & Accessibility)

3D 스캐너는 장비 가격이 수백만 원에서 수천만 원에 이르며, 설치 공간이 필요하다. 사용자 경험 측면에서도 특정 자세를 요구하고 촬영 시간이 길다.

반면 2D 스캐닝은 스마트폰 하나면 충분하며, 사용자가 특별한 시설이나 장비 없이 결과를 얻을 수 있다. 이 때문에 패션·헬스케어·스마트 피팅 산업에서는 “정밀도가 조금 낮아도, 사용성이 좋으면 시장성이 높다”는 판단이 널리 받아들여진다.

2.4.2.4 산업적 적용성(Application Suitability)

3D 스캐닝은 의류 패턴 설계, 인체 연구, 의료·재활 분야 등 정밀 데이터가 필수적인 영역에 적합하며, 2D 스캐닝은 온라인 피팅, 헬스케어 모니터링, 스포츠 자세 분석 등 대중적 서비스 중심의 영역에 활용된다.

Zheng et al.(2021)은 산업적 요구가 “절대 치수의 정확성”보다 “일관된 비율 정보”에 가까운 경우 2D 기술이 충분하다고 평가하였다.

2.4.2.5 종합 비교 및 연구의 위치 정의

선행연구들은 공통적으로 3D 스캐닝을 '정확도 우위 기술', 2D 스캐닝을 ' 접근성 우위 기술'로 평가한다. 본 연구는 이러한 기술 간 간극을 줄이기 위해, 딥러닝 기반 2D 이미지 스캐닝을 정량적으로 보정하여 3D 스캐닝에 근접시키는 실증적 접근을 취한다. 이는 Martinez et al.(2017)의 2D-to-3D 포즈 승격(lifting) 방식, Kanazawa et al.(2018)의 파라메트릭 메쉬 기반 접근, Zhao et al.(2021)의 스케일 보정 연구 등 다양한 선행연구의 성과를 통합적으로 반영한 방식이다.

Ⅲ. 연구방법 및 결과

3.1 연구 개요 및 목표

본 연구의 목적은 AI 기반 2D 이미지 스캐닝 기술을 활용하여 신체 치수를 자동으로 산출하는 '스마트 신체 치수 측정 시스템(SBMS: Smart Body Measurement System)'을 개발하는 것이다. 기존 군복 제작 과정에서는 숙련 재단사가 피측정자의 신체를 직접 줄자와 측정 도구로 계측하는 방식이 일반적으로 사용되었으나, 이 방식은 시간 소모가 크고 인력에 따라 측정 오차가 발생한다는 구조적 한계를 갖는다. 특히 대규모 병력 단위로 신체 데이터를 일괄 수집해야 하는 군 조직 특성상, 보다 정확하고 빠르며 자동화된 계측 기술의 도입은 필수적인 요구사항으로 대두되었다.

이에 본 연구는 컴퓨터 비전 기술(CV)과 딥러닝 기반 신체 특징점 인식 기술을 결합하여, 정면 A-포즈 이미지와 측면 이미지 두 장만으로 11개 신체 치수를 자동 산출하는 시스템을 구축하였다. [표 3-1]은 본 연구가 설정한 개발 목표치를 정한 것이다. 이를 위해 본 연구에서는 배경 제거 모듈, 신체 랜드마크 추출 모듈, 신체 경계 기반 폭/높이 분석 모듈, 딥러닝 회귀(Regression) 모델 등으로 구성된 통합 아키텍처를 설계·실현하였다.

[표 3-1] 시스템 개발 목표 사양

항목	목표치
측정 시간	1인당 1분 이내
신체 특징점	정면 25포인트, 측면 14포인트, 총 39포인트
측정 이미지	2컷 (정면 A-포즈 전신, 측면 전신)
측정 오차	최대 1인치(2.54cm) 이내
측정 항목	11개 항목
출력 형식	엑셀 호환 파일(.csv)

3.2 연구 환경 및 구성

3.2.1 데이터 수집

딥러닝 기반 예측 모델의 학습을 위해서는 다양한 체형 분포를 반영한 표본 기반 신체 이미지 및 실측 데이터를 확보하는 과정이 필수적이다. 본 연구는 체계적이고 통제된 환경에서 학습 데이터를 확보하기 위해 전용 촬영 스튜디오의 환경에서 이미지 데이터를 수집하였다.

데이터 수집 대상은 만 18세~55세 남성 200명으로 제한하였으며 오류측정에 대해하기 위해 추가로 4명을 더 측정하여서 총 204명의 이미지 데이터를 확보하였다. 이는 군복 제작의 주요 수요층인 현역 장정·예비역 체형을 우선 반영하기 위함이며, 초기 모델 학습에서 체형 변동 요인을 최소화하여 모델 안정성을 확보하기 위한 설정이다. 촬영 복장은 신체 윤곽이 명확히 드러나도록 몸에 밀착되는 반팔 티셔츠와 숏타이츠를 착용하도록 통일하였다. 또한 녹색 크로마키 배경을 설치하여 배경 제거 단계의 정확도를 극대화하였다.

[그림 3-1] 데이터 수집 촬영 환경 (좌: 정면 A-포즈, 우: 측면)

촬영은 [그림 3-1]과 같이 정면 A-포즈(full-body) 와 측면 포즈의 두 컷으로 구성하였으며, 각 참여자로부터 실측한 11개 신체 치수(목둘레, 어깨너비, 가슴둘레, 허리둘레 등)를 함께 기록하여 학습용 라벨 데이터로 활용하였다.

3.2.2 데이터 통계 분석

수집된 204명 데이터에 대해 기술통계 분석을 수행하였다. 키는 158.3cm~191.0cm 사이에 분포하였고 평균 174.7cm(표준편차 5.87)로, 한국 성인 남성 체형 분포를 고르게 포함하고 있음을 확인하였다. 몸무게는 최소 49.6kg에서 최대 111.7kg까지 다양하게 분포하였으며 평균값은 73.9kg(표준편차 10.55)이다. 또한 11개 신체 치수(목둘레, 어깨너비, 가슴둘레 등)는 모두 충분한 개체 간 변동성을 갖고 있어, 회귀 모델 학습에 필요한 분산(variance) 확보가 가능함을 확인하였다. 이는 모델이 특정 체형에 과적합되는 것을 방지하고 일반화 성능을 높이는 데 중요한 요소이다.

[표 3-2] 이미지 데이터 수집 기준

구분	내용
참여 자격	만 18세 이상 55세 이하 남성
최종 수집 인원	204명
수집 항목	신체 치수 11개 항목 + 신체 촬영 사진
촬영 배경	녹색 크로마키 배경
촬영 복장	몸에 붙는 반팔 티셔츠 + 숏타이츠

[표 3-3]에 제시된 기술통계량은 본 연구에서 구축한 학습·검증 데이터셋의 체형 분포와 변동성(variability) 을 요약적으로 보여주는 지표이며, 이후 딥러닝 회귀 모델의 일반화 가능성(generalizability) 을 판단하는 중요한 근거가 된다. 먼저 키는 158.3cm에서 191.0cm 범위로 분포하며 평균 174.7cm, 표준편차 5.87cm로 나타났다. 이는 데이터셋이 한국 성인 남성의 전형적인 신장

분포를 중심으로 하면서도, 저신장·고신장 구간을 모두 포함하고 있어 신장 스펙트럼 전반을 고르게 포괄하고 있음을 시사한다. 몸무게 역시 49.6kg에서 111.7kg까지 넓게 퍼져 있으며, 평균 73.9kg, 표준편차 10.55kg으로 나타나, 마른 체형부터 비만 체형까지 다양한 체형 유형이 포함된 데이터셋임을 확인할 수 있다.

[표 3-3] 수집 데이터 기술 통계량 (N=204)

측정 항목	최소값	최대값	평균	표준편차
키 (cm)	158.3	191.0	174.7	5.87
몸무게 (kg)	49.6	111.7	73.9	10.55
목둘레 (cm)	34.0	45.0	39.2	1.83
어깨너비 (cm)	41.0	82.0	49.2	3.28
소매길이 (cm)	47.0	71.0	60.9	3.10
가슴둘레 (cm)	81.0	121.0	100.5	7.00
배둘레 (cm)	64.0	109.0	82.9	8.32
허리둘레 (cm)	68.0	106.0	84.0	6.71
엉덩이둘레 (cm)	84.0	119.0	99.2	6.07
허벅지둘레 (cm)	46.0	71.0	58.4	4.59

둘레 항목을 보면, 목둘레·어깨너비·소매길이 등 상체 관련 지표는 모두 비교적 안정적인 표준편차(1.83cm~3cm대)를 보이며, 이는 상체 상단부 치수의 개체 간 편차가 상대적으로 제한적인 편임을 의미한다. 반면 가슴둘레(표준편차 7.00cm), 배둘레(8.32cm), 허리둘레(6.71cm)는 변동성이 더 크게 나타나는데, 이는 복부 비만도, 근육량, 체형 유형(정상·복부비만·근육형 등) 에 따라 상체 중단부와 하단부의 둘레 치수가 상당히 달라질 수 있다는 점을 반영한다. 엉덩이둘레와 허벅지둘레 또한 각각 6.07cm, 4.59cm의 표준편차를 보여, 하체 부위 역시 일정 수준의 체형 다양성을 내포하고 있음을 알 수 있다.

종합적으로 볼 때, 본 데이터셋은 단일 체형군에 국한되지 않고, 신장·체중·상·하체 둘레에 걸쳐 충분한 분산(variance)과 범위(range)를 확보한 표본 집단으로 평가할 수 있다. 이는 신체 치수 예측 모델이 특정 체형에 과적합

(overfitting)되지 않고, 다양한 체형에 대해 유의미한 회귀 관계를 학습할 수 있는 기반을 제공한다. 동시에 가슴둘레·배둘레·허리둘레와 같이 변동성이 큰 항목은 모델 설계 및 특성 공학(feature engineering) 단계에서 보다 세밀한 접근(예: 체형 지수, 비율 변수 추가 등)이 요구되는 대상임을 시사한다. 이러한 기술통계 분석은 단순한 값 나열을 넘어, 본 연구에서 구축한 SBMS 모델이 실제 군복·의류 맞춤 환경에 적용될 때의 적용 범위와 한계를 사전에 진단하는 중요한 기초 자료로 기능한다.

3.3 AI 기반 신체 측정 시스템 개발

본 연구에서 개발한 시스템은 이미지 전처리, 신체 특징점 추출 및 측정, 딥러닝 회귀 예측의 세 단계 구조로 개발되었다. 각 단계는 서로 독립적인 기능을 수행하면서도 전체 파이프라인 내에서 긴밀하게 연동되도록 설계되었다.

3.3.1 시스템 아키텍처

본 시스템은 [표 3-4]와 같이 크게 세 가지 모듈로 구성된다. 첫 번째, 전처리 모듈(PreProcess)은 rembg(U^2 -Net 기반) 모델을 사용하여 입력 이미지에서 배경을 제거하여, 랜드마크 추출 및 폭 분석의 정확도를 높인다. 두 번째, 신체 측정 모듈(BodyMeasure)은 MediaPipe Pose의 33개 신체 랜드마크를 추출한 후, 신체 각 부위의 위치를 계산하여 해당 지점의 폭(width)·높이(height)를 픽셀 단위로 측정한다. 마지막으로 회귀 예측 모듈(BodySize)은 측정된 픽셀 정보를 입력값으로 하여 TensorFlow 기반 DNN 회귀 모델이 실제 신체 치수(cm)를 예측한다.

[표 3-4] 시스템 모듈 구성

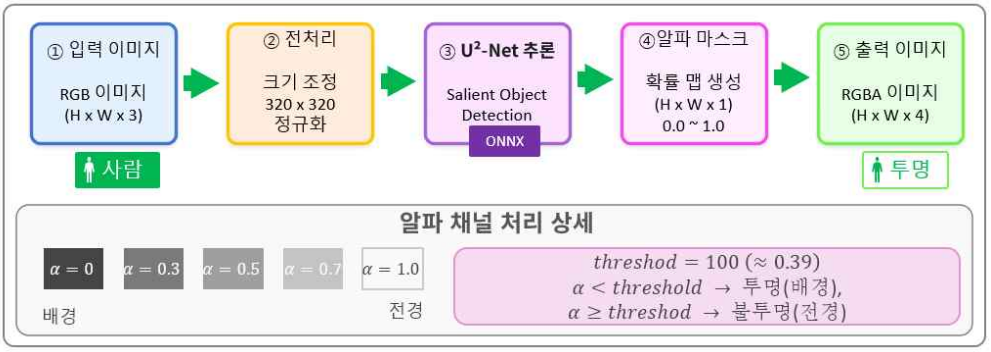
모듈명	파일명	주요 기능
전처리	preprocess.py	rembg + U² -Net 모델로 배경 제거
신체 측정	body_size.py	MediaPipe Pose로 33개 특징점 추출 및 측정
학습	train.py	TensorFlow DNN 회귀 모델 학습
예측	predict.py	학습된 모델로 신체 치수 예측

이와 같은 구조는 이미지 데이터를 통해 정보를 추출하고 최종적으로 값 변환이라는 계층적 처리 흐름을 통해 신체 치수 산출의 논리적 일관성과 확장성을 확보하게 된다.

3.3.2 rembg 라이브러리 및 U² -Net 배경 제거

본 rembg는 딥러닝 기반 이미지 배경 제거를 위한 Python 오픈소스 라이브러리이다. 내부적으로 U² -Net(U-Square-Net) 모델을 사용하여 Salient Object Detection(현저한 객체 탐지)을 수행한다. rembg는 ONNX Runtime을 통해 CPU 및 GPU 환경 모두에서 실행 가능하며, 다양한 사전 학습 모델을 제공한다.

U² -Net은 Qin et al.(2020)이 Pattern Recognition 저널에 발표한 Two-Level Nested U-Structure 기반의 딥러닝 아키텍처이다. 기존 SOD 모델들이 ImageNet 사전학습 백본(VGG, ResNet 등)에 의존하는 것과 달리, U² -Net은 새롭게 제안된 RSU(ReSidual U-block) 블록만으로 구성되어 처음부터 학습이 가능하다. RSU 블록은 내부에 소형 U-Net 구조를 가지며, 이를 통해 다양한 스케일의 특징을 효율적으로 추출한다.

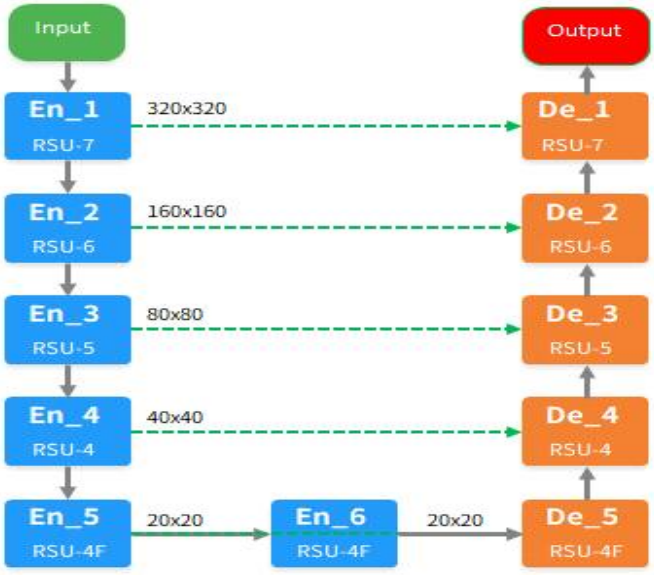


[그림 3-2] rembg 배경 제거 처리 파이프라인 Flow
(Qin et al., 2020; 연구자 재구성, 2025)

[그림 3-2]는 rembg 배경 제거 처리 파이프라인 Flow에 대한 설명이다. rembg 배경 제거 파이프라인은 입력된 RGB 이미지를 대상으로 전처리 → 특징 추출 → Salient Object Detection(SOD) → 마스크 생성 → 알파 합성(Alpha Compositing) 의 순차적 처리 구조로 구성된다.

먼저 입력 이미지는 ONNX Runtime 환경에서 최적화된 U² -Net 모델에 전달되며, 모델의 RSU(ReSidual U-block) 기반 다계층 특징 추출기(feature extractor)는 다양한 공간적 스케일(spatial scales)의 정보를 동시에 포착하여 인물 윤곽(silhouette)과 배경 간의 경계 정보를 고해상도로 분리한다. 이후 네트워크의 디코더(decoder) 단계에서는 물체와 배경의 확률적 분포를 나타내는 saliency map이 생성되며, 해당 맵을 기반으로 이진 마스크(binary segmentation mask)가 구성된다.

rembg는 이 마스크를 알파 채널과 결합하여 불투명한 인물 영역과 투명한 배1경 영역을 구분하는 RGBA 포맷의 출력 이미지를 생성한다. 마지막으로 후처리 과정에서는 사용자가 정의한 threshold 값 이하의 픽셀들을 완전 투명값(α=0)으로 재조정하여, 잔여 배경 노이즈를 최소화하고 신체 경계 검출에 최적화된 정제된 이미지를 확보한다. 이러한 파이프라인은 인물 분리 작업을 자동화할 뿐만 아니라, 후속 신체 특징점 추출 및 폭 측정 단계의 정확도를 실질적으로 높이는 기반 기능을 수행한다.



[그림 3-3] U² -Net 아키텍처 구조 (Qin et al., 2020; 연구자 재구성, 2025)

[그림 3-3]과 같이 U² -Net 아키텍처는 기존 U-Net 구조를 확장한 Two-level Nested U-Structure를 기반으로 하여, 다중 스케일 특징(multi-scale features)을 더욱 정교하게 추출할 수 있도록 설계된 Salient Object Detection(SOD) 모델이다. 핵심 구성 요소인 RSU(ReSidual U-block) 는 소형 U-Net을 블록 단위로 내장하고 있어, 각 단계에서 저수준 경계 정보와 고수준 의미 정보를 동시에 보존하며, 잔차 연결(residual connection)을 통해 깊은 네트워크에서도 안정적인 학습을 가능하게 한다.

RSU 블록은 컨볼루션과 풀링을 거쳐 점진적으로 수용영역(receptive field)을 확장한 뒤, 업샘플링과 스킵 연결(skip connection)을 통해 다양한 해상도의 정보를 결합함으로써, 복잡한 배경에서도 대상 객체의 윤곽(silhouette) 추출 성능을 극대화한다. 또한 U² -Net은 여러 해상도에서 생성된 예측 결과에 deep supervision을 적용하여 스케일별 특징 학습을 강화하며, 최종적으로 다중 스케일 saliency map을 융합(fusion)하여 고해상도 출력 마스크를 생성한다.

이러한 구조적 특성 덕분에 U² -Net은 경량 모델임에도 높은 정확도를 유지하며, 특히 배경 제거, 인물 분리와 같은 작업에서 세밀한 경계 복원과 노

이즈 억제 성능을 동시에 확보할 수 있는 효율적인 아키텍처로 평가된다.

[표 3-5] rembg 라이브러리 지원 모델
(rembg GitHub Repository; Qin et al., 2020; 연구자 재구성, 2025)

모델명	크기	용도 및 특징
u2net	176.3 MB	일반용 (기본값), 높은 정확도
u2netp	4.7 MB	경량화 버전, 모바일/엣지 디바이스 적합
u2net_human_seg	176.3 MB	인물 분리 전용, 사람 탐지에 최적화
u2net_cloth_seg	176.3 MB	의류 분리, 상의/하의/전신 분류
silueta	43 MB	u2net 압축 버전, 실루엣 추출
isnet-general-use	~170 MB	범용 고정밀 모델

rembg 라이브러리는 다양한 이미지 분리(Salient Object Segmentation) 요구에 대응하기 위해 여러 사전학습 모델(pretrained models)을 제공하며, 각 모델은 정확도, 연산 복잡도, 적용 도메인에 따라 차별화된 특성을 보인다. [표 3-5]는 이러한 모델별 크기와 용도에 대해 설명하고 있다. 기본 모델인 u2net(176.3MB) 은 가장 안정적이고 높은 정확도를 제공하는 범용 모델로, 인물·사물·배경이 혼재된 환경에서도 우수한 경계 추출 성능을 확보할 수 있어 일반적인 배경 제거 작업의 기본값(default choice)으로 사용된다. 이에 비해 u2netp(4.7MB) 은 파라미터 규모를 대폭 축소한 경량화(lightweight) 모델로, 모바일·엣지 컴퓨팅 환경에서 실시간 추론을 필요로 하는 애플리케이션에 적합하지만, 세밀한 경계 처리 능력은 기본 모델 대비 일정 부분 감소한다는 특징이 있다.

u2net_human_seg(176.3MB) 은 사람 인식(human parsing)에 특화된 버전으로, 신체 실루엣을 안정적으로 분리하도록 튜닝되어 의류 가상 착용(Fitting), 포즈 분석, 인체 스캔 등 사람 중심의 처리 작업에서 높은 정밀도를 보인다. 이와 유사한 맥락에서 u2net_cloth_seg(176.3MB) 은 의류 상·하의 및 전신 segmentation을 목적으로 설계되었으며, 패션 AI, 의류 분류, 전신 이미지 편집 등 특정 산업군에서 활용도가 높다. 반면 silueta(43MB) 모델은

u2net을 압축한 버전으로, 실루엣 추출을 중점적으로 수행하도록 설계된 효율형 모델로, 단순 배경 제거 또는 객체 외곽선 기반 처리에 적합한 균형 잡힌 성능을 제공한다. 그리고 isnet-general-use(~170MB) 은 U² -Net 계열을 개선한 고정밀(high-accuracy) 범용 모델로, 세밀한 경계선 복원 능력과 노이즈 억제 성능이 강화된 것이 특징이다. 전문 분야에서는 특히 복잡한 배경, 조명 편차, 인물·객체 간 경계가 모호한 환경에서 기존 모델 대비 높은 안정성을 보인다는 평가를 받는다.

종합적으로 볼 때, rembg 모델군은 정확도-속도-도메인 특이성(domain specificity) 간 다양한 트레이드오프 옵션을 제공함으로써, 연구 목적과 시스템 요구조건에 따라 최적의 모델을 선택할 수 있는 유연한 구조를 갖추고 있다. 이는 AI 기반 배경 제거를 활용하는 다양한 산업 및 연구 분야에서 rembg가 폭넓게 적용될 수 있는 기술적 기반을 형성한다.

```
from rembg import remove, new_session
import cv2

class PreProcess:
    def clear_background(self, image):
        # U²-Net 모델로 배경 제거 (u2net 세션 사용)
        image = remove(image, session=new_session("u2net"))
        threshold = 100  # 알파 채널 임계값 (0-255)
        image_copy = image.copy()
        # 알파값 < threshold인 픽셀을 완전 투명으로 설정
        image_copy[image_copy[:,:,3] < threshold] = 0
        return image_copy  # RGBA 이미지 반환
```

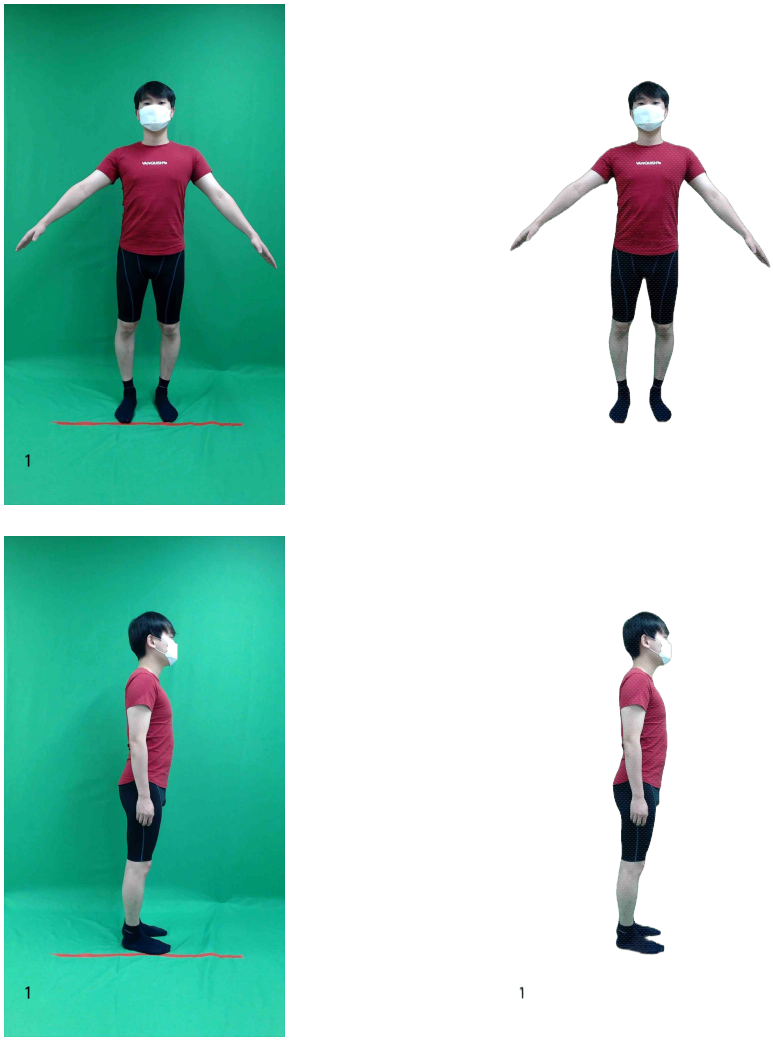
[그림 3-4] rembg를 이용한 배경 제거 모듈 (preprocess.py)

[그림 3-4]은 rembg를 이용한 배경 제거 모듈 (preprocess.py) 의 실제 코드로서 본 연구에서 구현한 PreProcess 모듈은 신체 치수 측정 파이프라인의 전처리 단계에서 핵심적인 기능을 수행하며, 입력 이미지로부터 배경 요소를 제거하여 후속 랜드마크 추출 및 신체 경계 탐색의 정확도를 극대화하기 위해 설계되었다. 해당 모듈은 rembg 라이브러리가 제공하는 U² -Net 기반

Salient Object Segmentation 모델을 호출하여 인물 영역과 배경 영역을 분리한 후, RGBA 형식으로 변환된 출력 이미지의 알파 채널(α)을 기반으로 신체 실루엣을 정제(refinement)하는 절차를 포함한다.

코드 상에서 remove() 함수는 지정된 세션(new_session("u2net"))을 사용하여 입력 이미지를 U^2 -Net 모델에 전달하고, 모델은 이미지 내의 사람 영역을 고정밀로 추출한 뒤 투명 배경이 적용된 RGBA 이미지로 반환한다. 이러한 방식은 기존 RGB 기반 마스킹보다 경계선(edge boundary)의 보존도가 높아, 신체 윤곽 추출 시 발생할 수 있는 외곽 잡음(noise) 및 배경 잔여물(residual artifacts)을 효과적으로 제거한다. 이후 후처리 단계에서는 연구자가 정의한 임계값(threshold = 100)을 기준으로 알파 채널 값이 일정 수준 미만인 픽셀을 완전 투명값(0)으로 설정하여, 실루엣 주변의 반투명한 경계 영역을 제거하고 신체 윤곽선을 더욱 명확하게 정제한다.

이와 같은 전처리 설계는 신체 경계 기반 폭 측정 및 랜드마크 기반 위치 추정 과정에서 발생할 수 있는 배경 간섭(background interference)을 최소화함으로써, 전체 신체 치수 예측 모델의 정확도 향상에 기여한다. 특히 녹색 크로마키가 아닌 일반 실내 배경에서도 안정적으로 작동할 수 있어, 본 연구의 시스템이 실제 환경에서 활용될 가능성을 고려했을 때 강인한 전처리 기반(robust preprocessing foundation) 으로 기능하는 점에서 의미가 크다.



[그림 3-5] rembg + U^2 -Net 모델을 이용한 배경 제거 처리 결과

3.3.3 MediaPipe Pose 랜드마크

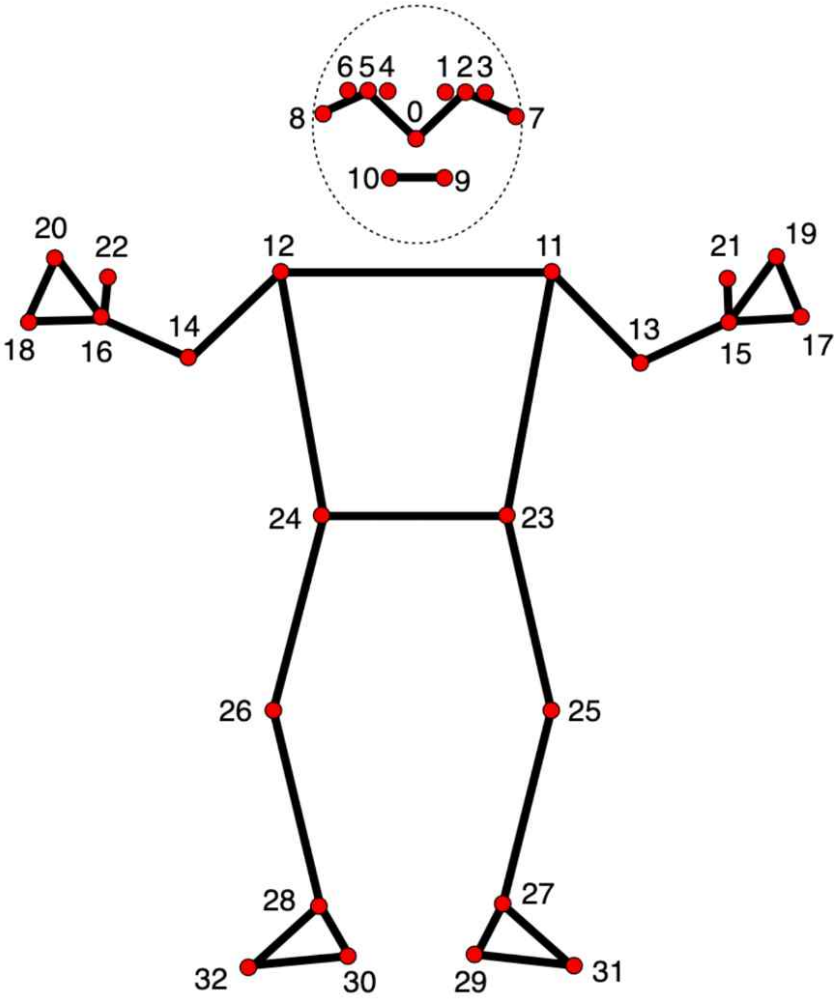
본 연구에서는 인체 자세 및 신체 치수의 정밀한 추정을 위해 MediaPipe Pose에서 정의하는 33개 전신 랜드마크를 채택하였다. 기존 COCO(17개), OpenPose(18~25개) 기반 포인트 구성은 상·하지의 주요 관절 위치 추출에는 적합하나, 발·손의 말단 포인트 및 얼굴 방향과 같은 세부 정보를 충분히 반영하기에는 한계가 있다. 반면 33개 랜드마크 구성은 얼굴, 손가락 기점, 발

뒤꿈치 및 발 앞부분까지 포함함으로써, 보행 분석, 균형 평가, 체형 비대칭성 분석 등과 같이 정밀도를 요구하는 응용에 보다 적합한 해상도를 제공한다. 또한 33개 포인트는 모바일·엣지 환경에서도 실시간 처리가 가능한 수준의 연산 복잡도를 유지하면서, 다양한 응용 분야(피트니스 코칭, 재활, 스포츠 분석, 신체 치수 측정)를 하나의 공통 스켈레톤 표현으로 통합할 수 있다는 점에서 연구·실무 양 측면의 균형적인 선택으로 판단된다.

MediaPipe Pose는 Google에서 개발한 실시간 신체 자세 추정 솔루션으로, 단일 RGB 이미지에서 33개의 3차원 신체 랜드마크를 추출한다. BlazePose 아키텍처를 기반으로 하며, 모바일 기기에서도 실시간 처리가 가능할 만큼 경량화되어 있다. 각 랜드마크는 x, y, z 좌표와 visibility(가시성) 값을 포함하며, 이미지 좌표계로 정규화되어 출력된다.

배경 제거 후, MediaPipe Pose를 활용하여 [그림 3-6]과 같이 MediaPipe Pose 33개 신체 랜드마크 위치를 추출하였다. 기존 COCO(17개), OpenPose(18~25개)에 비해 보다 세밀한 말단부(손가락·발끝) 및 얼굴 정보까지 포함함으로써, 신체 비율 계산 및 위치 추정 오류를 현저히 감소시킬 수 있다.

MediaPipe Pose는 BlazePose 아키텍처를 기반으로 경량화되어 있으며, 단일 RGB 이미지에서 x·y·z 및 visibility 값을 동시에 산출한다. 이를 통해 신체 각 부위의 상대적 위치와 신체 중심축을 보다 정밀하게 파악할 수 있었으며, 정면·측면 두 장의 이미지를 조합하여 11개 치수 측정 위치를 일관되게 계산할 수 있었다.



[그림 3-6] MediaPipe Pose 33개 신체 랜드마크 위치(Google MediaPipe Site, 2025; 연구자 재구성, 2025)

[표 3-6] 신체 랜드마크 위치

ID	그룹	한글명	영문명
0	얼굴/머리	코	NOSE
1		왼쪽 눈 안쪽	LEFT_EYE_INNER
2		왼쪽 눈	LEFT_EYE
3		왼쪽 눈 바깥쪽	LEFT_EYE_OUTER
4		오른쪽 눈 안쪽	RIGHT_EYE_INNER
5		오른쪽 눈	RIGHT_EYE
6		오른쪽 눈 바깥쪽	RIGHT_EYE_OUTER
7		왼쪽 귀	LEFT_EAR
8		오른쪽 귀	RIGHT_EAR
9		입 왼쪽	MOUTH_LEFT
10		입 오른쪽	MOUTH_RIGHT
11	상체 관절	왼쪽 어깨	LEFT_SHOULDER
12		오른쪽 어깨	RIGHT_SHOULDER
13		왼쪽 팔꿈치	LEFT_ELBOW
14		오른쪽 팔꿈치	RIGHT_ELBOW
15		왼쪽 손목	LEFT_WRIST
16		오른쪽 손목	RIGHT_WRIST
17	손가락 포인트	왼쪽 새끼손가락	LEFT_PINKY
18		오른쪽 새끼손가락	RIGHT_PINKY
19		왼쪽 검지손가락	LEFT_INDEX
20		오른쪽 검지손가락	RIGHT_INDEX
21		왼쪽 엄지손가락	LEFT_THUMB
22		오른쪽 엄지손가락	RIGHT_THUMB
23	하체 관절	왼쪽 엉덩이	LEFT_HIP
24		오른쪽 엉덩이	RIGHT_HIP
25		왼쪽 무릎	LEFT_KNEE
26		오른쪽 무릎	RIGHT_KNEE
27		왼쪽 발목	LEFT_ANKLE
28		오른쪽 발목	RIGHT_ANKLE
29	발 포인트	왼쪽 발뒤꿈치	LEFT_HEEL
30		오른쪽 발뒤꿈치	RIGHT_HEEL
31		왼쪽 발끝	LEFT_FOOT_INDEX
32		오른쪽 발끝	오른쪽 발끝

3.3.4 신체 경계 기반 폭·높이 측정

랜드마크를 기준으로 특정 y좌표를 계산한 뒤, 배경 제거된 이미지의 알파 채널(alpha channel) 을 탐색하여 신체의 좌·우 경계를 찾는 방식으로 폭(width)을 측정하였다. 이는 현존하는 많은 연구에서 사용하는 외부 bounding box 기반 방식보다 훨씬 정밀하며, 신체 굴곡·체형 특성에 따른 실루엣 변화를 정확히 반영할 수 있다는 강점이 있다. 가령 가슴둘레 측정의 경우, 어깨~엉덩이 구간의 수직 길이를 기준으로 상단에서 약 80% 지점의 y좌표에서 폭을 측정함으로써, 체형 편차로 인해 발생하는 위치 불일치 문제를 최소화하였다.

[표 3-7] 본 연구에서 활용하는 주요 MediaPipe 랜드마크
(Bazarevsky et al., 2020; 연구자 재구성, 2025)

측정 항목	랜드마크 번호	랜드마크명	용도
목 위치	0	NOSE	목 y좌표 기준점
어깨 너비	11, 12	SHOULDER	어깨 좌우 끝점
소매 길이	11→13→15	SHOULDER→WRIST	팔 길이 계산
가슴/허리	11, 12, 23, 24	SHOULDER, HIP	상체 비율 계산
엉덩이	23, 24	HIP	엉덩이 y좌표
허벅지	23, 24, 27, 28	HIP, ANKLE	다리 위치 비율

본 연구에서는 MediaPipe Pose에서 정의하는 총 33개의 전신 랜드마크 중, 신체 치수 산출에 직접적으로 기여하는 핵심 관절점들을 선별적으로 활용하였다. [표 3-7]은 주요 MediaPipe 랜드마크 해설로서 이러한 랜드마크들은 피사체의 정면 및 측면 이미지에서 정규화된 2D 좌표(x, y)와 가시성 값(visibility) 으로 제공되며, 이를 기반으로 신체 부위의 위치 계산, 폭(width) 탐색, 신체 비율(body proportion) 모델링 등 다양한 치수 산출 절차에 적용된다. 특히 MediaPipe Pose의 랜드마크 구조는 단순한 위치 좌표를 넘어, 상·

하체의 기하학적 관계를 파악할 수 있는 근거를 제공하므로, 2D 이미지 기반 신체 측정의 정확도를 좌우하는 핵심 입력 요소라 할 수 있다.

[표 3-8] 부위별 측정 위치 계산 비율

측정 부위	기준점 (랜드마크)	위치 비율
가슴 (chest)	SHOULDER(12) ~ HIP(24)	상단에서 80%
배 (belly)	SHOULDER(12) ~ HIP(24)	상단에서 30%
허리 (waist)	SHOULDER(12) ~ HIP(24)	상단에서 18%
엉덩이 (hip)	HIP(23, 24)	엉덩이 y좌표
허벅지 (thigh)	HIP(24) ~ ANKLE(28)	상단에서 20%

3.3.4.1 목 위치의 기준점 활용

MediaPipe Pose에서 제공하는 NOSE(코 위치)는 33개 랜드마크 중 가장 안정적으로 검출되는 상체 기준점 중 하나로, 얼굴 방향에 크게 영향을 받지 않으며 정면·측면 사진 모두에서 높은 가시성을 확보한다.

본 연구에서는 NOSE의 y좌표를 목둘레 측정 높이를 추정하기 위한 기준점(vertical reference point) 으로 활용하였다. 이는 실제 인체 구조에서 코–목–어깨가 비교적 일정한 수직적 거리 비율을 가지는 점을 고려한 것으로, 상체 상단부의 기준점을 정규화하지 않을 경우 개개인의 체형에 따라 목둘레 측정 기준점이 상하로 흔들리는 문제를 방지하기 위해 적용하였다.

따라서 NOSE는 직접적으로 폭 측정을 수행하는 랜드마크는 아니지만, 세로 위치(y-axis) 정렬 기준점으로 활용됨으로써 이후의 경계 탐색 알고리즘의 일관성을 높이는 중요한 역할을 수행한다. [그림 3-6]은 목 위치 측정시 코와 어깨 사이의 비율 및 목둘레 측정 높이에 대한 추정 수식을 나타낸다. 코(NOSE)를 상체 상단 기준점으로 사용하고, 코–어깨 사이의 비율 λ_{neck}를 적용하여 목둘레 측정 높이 y_{neck}을 추정하였다.

여기서 λ_{neck}는 실측 데이터 기반으로 경험적으로 설정되는 상수(예: 0.3~0.5 범위)이다.

$$y_{\text{shoulder,avg}} = \frac{y_{11} + y_{12}}{2}$$

$$y_{\text{neck}} = y_0 + \lambda_{\text{neck}} \cdot (y_{\text{shoulder,avg}} - y_0)$$

3.3.4.2 어깨 너비 기준점 활용

어깨너비는 신체 폭 측정 항목 중 가장 대표적이며, 이미지 기반 신체 측정에서 비교적 높은 정확도로 산출 가능한 항목이다.

본 연구에서는 ID 11(왼쪽 어깨)과 ID 12(오른쪽 어깨) 의 x좌표 차이를 이용하여 어깨 수평 기준선(horizontal baseline)을 설정하였다. 우선 어깨 좌우 끝점을 연결한 직선은 상체 상단부의 기준 폭이며, 이후 가슴·배·허리 비율을 산출할 때 사용되는 상체 구간 계산의 기준선 역할도 수행한다.

또한 정면 이미지에서는 어깨 랜드마크의 깊이(z-axis) 변화가 제한적이므로, z좌표의 왜곡 영향을 최소화할 수 있는 장점이 있다. 따라서 어깨 랜드마크는 단순 폭 계산을 넘어 전신 비율 정규화(normalization) 모델의 기준 축(axis) 으로 표현되고 있다.

어깨너비의 픽셀 단위 폭은 좌우 어깨 랜드마크의 x좌표 차이로 정의할 수 있다.

$$W^{\text{px}}_{\text{shoulder}} = |x_{11} - x_{12}|$$

실제 길이(cm)로의 변환은 DNN 회귀모델 $f$shoulder(·)에 의해 이루어진다 .

$$L^{\text{cm}}_{\text{shoulder}} = f_{\text{shoulder}}\left(W^{\text{px}}_{\text{shoulder,front}}, W^{\text{px}}_{\text{shoulder,side}}\right)$$

3.3.4.3 소매길이 계산

팔의 길이는 다음 세 랜드마크를 연속적으로 연결하여 계산하였다. 어깨(11 또는 12), 팔꿈치(13 또는 14), 손목(15 또는 16), 팔의 길이는 단순 직선거리로 측정되는 것이 아니라, 상완(upper arm) + 하완(forearm) 두 구간으로 구성되므로 다음 두 거리의 합으로 산출한다.

오른쪽 팔 기준(왼팔은 동일 구조)으로, 어깨–팔꿈치–손목 구간의 길이 합으로 정의한다.

$$L^{\text{px}}_{\text{upper-arm}} = d(P_{12}, P_{14})$$

$$L^{\text{px}}_{\text{forearm}} = d(P_{14}, P_{16})$$

$$L^{\text{px}}_{\text{arm}} = L^{\text{px}}_{\text{upper-arm}} + L^{\text{px}}_{\text{forearm}}$$

이를 회귀모델 $f_{\text{sleeve}}(\cdot)$로 변환하여 소매길이(cm)를 추정하면 다음과 같다.

$$L^{\text{cm}}_{\text{sleeve}} = f_{\text{sleeve}}\left(L^{\text{px}}_{\text{arm,front}}, L^{\text{px}}_{\text{arm,side}}\right)$$

이는 인간의 관절이 굽혀질 수 있는 특성을 고려한 것으로, 단일 직선거리(shortest path)만을 사용할 경우 팔이 굽혀진 자세에서 큰 오차가 발생하기 때문이다. 수식적 접근은 사진 촬영 시 자세의 변형을 완전히 배제할 수 없는 현실적 환경에서 길이 측정의 기하학적 안정성을 확보하는 중요한 요소이다.

3.3.4.4 가슴·배·허리위치 산정

가슴둘레와 허리둘레는 인체의 대표적인 둘레 항목이지만, 2D 이미지에서는 실제 원형/타원형 단면을 직접 측정할 수 없기 때문에 높이 비율 기반 가상 단면 위치 추정 방법을 사용하였다. 본 연구에서는 어깨(11,12), 엉덩이

좌/우(23,24)의 네 지점을 사용하였다. 어깨–엉덩이 구간을 상체의 주요 세로축으로 간주하고, 해당 구간의 비율 값을 이용해 다음과 같은 방식으로 y좌표를 결정한다.

가슴, 배, 허리는 어깨–엉덩이 구간의 비율로 y좌표를 정의하며 일반형 수식은 다음과 같다.

$$y_{\text{level}} = y_{\text{shoulder,avg}} + r \cdot (y_{\text{hip,avg}} - y_{\text{shoulder,avg}})$$

r은 각 항목별 y좌표 비율은 가슴 : $r$chest = 0.80, 배 : $r$belly = 0.30, 허리 : $r$waist = 0.18 으로 정의할 수 있으며 다음과 같이 표현할 수 있다.

$$y_{\text{chest}} = y_{\text{shoulder,avg}} + 0.80 \cdot (y_{\text{hip,avg}} - y_{\text{shoulder,avg}})$$

즉, 랜드마크 자체가 가슴 위치를 직접 제공하는 것이 아니라, 상체 해부학적 비율 모델(anatomical proportional model) 을 적용하여 가상의 측정선을 생성하는 방식이다. 이렇게 도출된 y좌표에서 좌우 경계를 탐색한 폭(width)은, 회귀 모델에 입력되어 실제 둘레(cm)로 변환된다.

이는 2D 이미지의 구조적 제한을 극복하기 위한 필수적 접근 방법으로, 타 연구에서도 유사한 비율 기반 치수 추정 기법이 사용되고 있다.

3.3.4.5 가슴·배·허리위치 폭 측정

엉덩이 위치는 MediaPipe Pose가 제공하는 두 랜드마크인 LEFT_HIP(23), RIGHT_HIP(24)를 직접적으로 인체의 관골(hip bone) 위치를 반영하므로, 별도의 비율 계산 과정 없이 해당 y좌표를 그대로 사용한다. 이 부위는 가슴·허리와 달리 체형 변화가 적고, 각 랜드마크의 탐지 신뢰도(visibility)가 높은 편이기 때문에 직접적 측정값을 사용할 수 있다. 이는 하체 측정의 기준 축을 명확히 제공하여 허벅지둘레·사지 비율 계산에 활용된다.

각 부위의 y좌표에서 알파 채널을 사용해 좌우 경계를 탐색하면 임계값은 $T\alpha$ = 100 (0 – 255)이며, 해당 수평선에서 가장 왼쪽/오른쪽 인체 경계 좌표 값은 다음과 같다.

$$x_{\text{left}}(y) = \min\{x \mid A(x,y) \geq T_\alpha\}$$

$$x_{\text{right}}(y) = \max\{x \mid A(x,y) \geq T_\alpha\}$$

따라서 폭을 계산하면 다음과 같다.

$$W(y)^{\text{px}} = x_{\text{right}}(y) - x_{\text{left}}(y)$$

이 값을 기준으로 각 부위에 적용하면 다음과 같다.

$$W_{\text{chest}}^{\text{px}} = W(y_{\text{chest}})^{\text{px}}, \quad W_{\text{belly}}^{\text{px}} = W(y_{\text{belly}})^{\text{px}}, \quad W_{\text{waist}}^{\text{px}} = W(y_{\text{waist}})^{\text{px}}$$

실제 둘레(cm)는 각 항목별 회귀모델 $fk(\cdot)$에 의해 추정된다.

$$C_{\text{chest}}^{\text{cm}} = f_{\text{chest}}\left(W_{\text{chest,front}}^{\text{px}}, W_{\text{chest,side}}^{\text{px}}\right)$$

$$C_{\text{belly}}^{\text{cm}} = f_{\text{belly}}\left(W_{\text{belly,front}}^{\text{px}}, W_{\text{belly,side}}^{\text{px}}\right)$$

$$C_{\text{waist}}^{\text{cm}} = f_{\text{waist}}\left(W_{\text{waist,front}}^{\text{px}}, W_{\text{waist,side}}^{\text{px}}\right)$$

3.3.4.6 엉덩이둘레 측정

엉덩이 위치는 LEFT_HIP(23), RIGHT_HIP(24) 두 랜드마크가 직접적으로 인체의 관골(hip bone) 위치를 반영하므로, 별도의 비율 계산 과정 없이 해당

y좌표를 그대로 사용한다.

이 부위는 가슴·허리와 달리 체형 변화가 적고, 각 랜드마크의 탐지 신뢰도(visibility)가 높은 편이기 때문에 직접적 측정값을 사용할 수 있다. 이는 하체 측정의 기준 축을 명확히 제공하여 허벅지둘레·사지 비율 계산에 활용된다.

엉덩이 높이는 엉덩이 랜드마크 평균 y좌표로 직접 정의할 수 있다.

$$y_{\text{hip}} = \frac{y_{23} + y_{24}}{2}$$

해당 높이에서의 폭 계산식은 다음과 같다.

$$W_{\text{hip}}^{\text{px}} = W(y_{\text{hip}})^{\text{px}}$$

따라서, 실제 엉덩이둘레의 계산식은 다음과 같다.

$$C_{\text{hip}}^{\text{cm}} = f_{\text{hip}}\left(W_{\text{hip,front}}^{\text{px}}, W_{\text{hip,side}}^{\text{px}}\right)$$

3.3.4.7 허벅지둘레 측정

허벅지둘레는 HIP (23,24), ANKLE (27,28) 과 같이 네 개 랜드마크를 기반으로 계산한다. 허벅지의 실제 단면 위치는 사람마다 상당한 차이가 있으므로, 연구에서는 HIP과 ANKLE 구간의 세로축 상에서 상단에서 약 20% 지점을 허벅지 단면 위치로 설정하였다.

이는 운동형, 비만형 등 다양한 체형을 고려한 경험적·통계적 비율이며, 기존 인체 공학 연구에서 허벅지 근위부(thigh proximal region)의 평균적 위치와도 부합한다. 해당 지점에서 신체 경계를 탐색하여 폭을 계산하고, DNN 회귀 모델을 통해 실제 둘레로 변환한다.

허벅지 측정 높이는 HIP – ANKLE 구간 상단에서 20% 지점으로 정의한다.

$$y_{\text{thigh}} = y_{\text{hip}} + 0.20 \cdot (y_{\text{ankle,avg}} - y_{\text{hip}})$$

$$y_{\text{ankle,avg}} = \frac{y_{27} + y_{28}}{2}$$

해당 높이에서의 폭 계산식은 다음과 같다.

$$W_{\text{thigh}}^{\text{px}} = W(y_{\text{thigh}})^{\text{px}}$$

따라서, 실제 허벅지둘레의 계산식은 다음과 같다.

$$C_{\text{thigh}}^{\text{cm}} = f_{\text{thigh}}\left(W_{\text{thigh,front}}^{\text{px}}, W_{\text{thigh,side}}^{\text{px}}\right)$$

3.3.4.8 총장·상기장·하기장 측정 일반식

어깨부터 발목까지의 길이 산정식은 다음과 같다.

$$L_{\text{total}}^{\text{px}} = d(\bar{P}_{\text{shoulder}}, \bar{P}_{\text{ankle}})$$

이를 DNN 기반 회귀모델로 변환하면 다음과 같다.

$$L_{\text{total}}^{\text{cm}} = f_{\text{total}}\left(L_{\text{total,front}}^{\text{px}}, L_{\text{total,side}}^{\text{px}}\right)$$

3.3.5 신체 특징점 추출 및 측정

신체 측정 모듈은 MediaPipe Pose 모델에서 추출된 랜드마크를 기준으로 목, 가슴, 배, 허리, 엉덩이, 허벅지 등의 측정 위치를 계산한다. 각 측정 위치에서 배경이 제거된 PNG 이미지의 알파 채널을 탐색하여 신체의 좌측과 우측 경계를 찾고, 해당 부위의 폭을 픽셀 단위로 측정한다.

```
def find_coordinate(self, img, x, y):
    x1 = x2 = x
    threshold = 100  # 알파 채널 임계값 (rembg 출력과 동일)
    # 좌측 경계 탐색 (알파값이 threshold 미만이 될 때까지)
    while img[y, x1][3] >= threshold:  # [3]은 알파 채널
        x1 -= 1
    # 우측 경계 탐색
    while img[y, x2][3] >= threshold:
        x2 += 1
    return [[x1, y], [x2, y]]  # 좌측, 우측 경계 좌표
```

[그림 3-7] 알파 채널 기반 신체 경계 탐색 함수 (body_size.py)

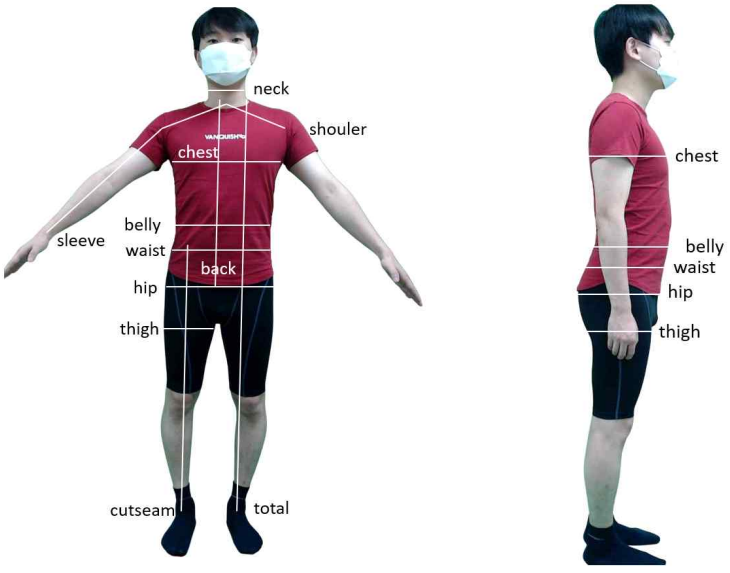
[그림 3-7]은 배경이 제거된 RGBA 이미지에서 특정 신체 단면의 좌·우 경계를 탐색하여 폭을 산출하기 위한 핵심 알고리즘이다. 입력으로는 알파 채널을 포함한 이미지와 기준 좌표 (x,y)(x,y)가 주어지며, 해당 y좌표를 따라 수평 방향으로 경계 탐색을 수행한다. 초기 탐색 위치는 기준 x좌표로 설정되며, 알파 채널 값이 임계값(threshold=100) 이상인 영역을 신체 내부로 간주한다. 좌측 탐색은 알파 값이 임계값 미만이 될 때까지 x좌표를 감소시키며, 우측 탐색은 동일한 조건에서 x좌표를 증가시키는 방식으로 수행된다. 이 과정은 배경과 인체 실루엣의 경계가 알파 채널 값의 급격한 변화로 표현된다는 전제에 기반한다. 최종적으로 반환되는 좌·우 경계 좌표는 해당 신체 부위의 픽셀 단위 폭 계산에 직접 활용되며, 후속 회귀 모델의 입력 특성으로 사용된다.

3.3.6 신체 측정 위치 추출

정면 및 측면 이미지에서의 신체 측정 위치 추출 결과는 본 연구에서 제안한 랜드마크 기반 위치 추정 알고리즘이 안정적으로 동작함을 시각적으로 검증하는 핵심 근거를 제공한다. 정면 이미지에서는 MediaPipe Pose를 통해 추출된 어깨, 엉덩이, 무릎, 발목 등의 주요 랜드마크를 기준으로 상체 및 하체

의 세로 기준축이 정확히 정렬되었으며, 이를 기반으로 가슴, 배, 허리, 엉덩이, 허벅지 측정 위치가 인체 해부학적 구조와 일관되게 배치됨을 확인할 수 있었다. 특히 어깨–엉덩이 구간의 비율 모델을 적용한 가슴·허리 측정선은 체형 차이에도 불구하고 상대적으로 안정적인 위치를 유지하였다.

측면 이미지에서는 정면에서 관측하기 어려운 체형 두께 정보가 반영되어, 가슴과 복부의 전후 폭(depth)에 해당하는 픽셀 정보가 효과적으로 추출되었다. 이는 동일한 정면 폭을 가지더라도 실제 둘레가 달라질 수 있는 체형 차이를 보완하는 중요한 요소로 작용한다. 또한 배경 제거 처리 후 알파 채널 기반 경계 탐색이 정면·측면 모두에서 명확한 신체 실루엣을 형성하여, 측정 위치에서의 좌우 경계 검출이 일관되게 수행되었음을 확인하였다. 이러한 결과는 본 연구의 신체 측정 위치 추출 방식이 2D 이미지 기반 환경에서도 신뢰성 있는 입력 데이터를 생성할 수 있음을 보여주며, 이후 딥러닝 회귀 모델을 통한 실제 치수 예측의 타당성을 뒷받침한다.



[그림 3-8] 신체 측정 위치 추출 결과 (좌: 정면, 우: 측면)

3.3.7 딥러닝 회귀 모델

측정된 픽셀 값(폭, 높이)을 실제 치수(cm)로 변환하기 위해 TensorFlow 기반의 DNN(Deep Neural Network) 회귀 모델을 설계하였다. 각 측정 항목(목둘레, 어깨너비 등 11개)에 대해 개별 모델을 학습시켰다. 모델 구조는 입력층(2) → Dense(32) → Dense(64) → Dense(128) → Dense(256) × 5 → 출력층(1)으로 구성되며, 활성화 함수로 ReLU를 사용하였다.

```
model = tf.keras.Sequential([
    tf.keras.layers.Dense(32, activation='relu', input_shape=(2,)),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dense(128, activation='relu'),
    tf.keras.layers.Dense(256, activation='relu'),  # x5 반복
    tf.keras.layers.Dense(1)  # 출력층 (회귀)
])
model.compile(optimizer='adam', loss='mse', metrics=[MySum()])
model.fit(x_train, y_train, epochs=32, batch_size=16, validation_split=0.1)
```

[그림 3-9] DNN 회귀 모델 구조 (train.py)

[그림 3-9]의 딥러닝 회귀 모델은 신체 측정 과정에서 산출된 2차원 픽셀 특성값(정면·측면 폭 또는 폭·높이 조합)을 입력으로 받아 실제 신체 치수(cm)를 예측하도록 설계되었다. 모델은 TensorFlow의 Sequential 구조를 기반으로 하며, 입력층 이후 점진적으로 노드 수를 확장하는 다층 완전연결(Dense) 구조를 채택하여 비선형 회귀 관계를 효과적으로 학습한다. 각 은닉층에는 ReLU 활성화 함수를 적용하여 기울기 소실 문제를 완화하고 학습 안정성을 확보하였다. 출력층은 단일 뉴런으로 구성되어 연속형 값을 직접 예측하는 회귀 문제에 적합하도록 설계되었다. 학습 과정에서는 Adam 옵티마이저를 사용하여 빠른 수렴을 유도하고, 손실 함수로 평균제곱오차(MSE)를 적용하여 예측값과 실제값 간의 오차를 최소화하였다. 또한 전체 데이터의 10%를 검증용(validation)으로 분리하여 과적합을 방지하고 모델의 일반화 성능을 평가하였다.

Ⅳ. 연구 결과

4.1 딥러닝 회귀 모델 기반 측정 정확도 분석

본 연구에서 제안한 AI 기반 2D 이미지 신체 치수 측정 시스템의 성능을 검증하기 위해, 실측값과 시스템 예측값 간의 오차를 비교 분석하였다. 평가 대상은 길이 계열 항목(소매길이, 어깨너비, 총장, 상기장, 하기장)과 둘레 계열 항목(목둘레, 가슴둘레, 배둘레, 허리둘레, 엉덩이둘레, 허벅지둘레)으로 구분하여 수행하였다. 이는 각 항목이 갖는 기하학적 특성과 2D 이미지 기반 측정에서의 구조적 난이도가 상이하기 때문이다.

분석 결과, 길이 계열 항목의 경우 평균 약 99% 수준의 높은 측정 정확도를 나타냈다. 이는 소매길이, 어깨너비, 총장과 같은 항목이 MediaPipe Pose 랜드마크 간의 거리 계산을 통해 비교적 직접적으로 산출될 수 있으며, 신체의 회전이나 체형 변화에 따른 영향이 제한적이기 때문으로 해석된다. 특히 어깨–팔꿈치–손목으로 구성되는 소매길이는 관절 구조가 명확하고 랜드마크 검출 신뢰도가 높아, 딥러닝 회귀 모델이 픽셀 거리와 실제 치수 간의 비선형 관계를 안정적으로 학습할 수 있었다.

둘레 계열 항목 중 목둘레와 허벅지둘레 또한 약 98% 수준의 정확도를 기록하였다. 이들 부위는 상대적으로 단면 위치가 명확하고, 체형 변화에 따른 위치 편차가 비교적 작으며, MediaPipe Pose 랜드마크를 기준으로 한 비율 기반 단면 설정이 효과적으로 작동한 사례로 판단된다. 특히 허벅지둘레의 경우, 엉덩이–발목 구간을 기준으로 한 비율적 위치 설정과 알파 채널 기반 경계 탐색 방식이 실제 체형 변동성을 충분히 반영한 것으로 분석된다.

이러한 결과는 본 연구에서 설계한 DNN 회귀 모델 구조가 단순 선형 변환이 아닌, 픽셀 기반 특징값과 실제 신체 치수 간의 복잡한 비선형 관계를 효과적으로 모델링할 수 있음을 실증적으로 보여준다. 또한 정면·측면 이미지에서 추출된 두 개의 입력 특성을 활용한 모델 구조가 2D 이미지 기반 측정의 한계를 일정 수준 보완했음을 의미한다.

4.2 측정 항목의 오차 분석 및 구조적 한계

가슴둘레, 배둘레, 허리둘레 항목에서는 평균적으로 약 3cm 이상의 오차가 관측되었다. 이러한 오차는 단일 요인에 의해 발생한 것이 아니라, 2D 이미지 기반 신체 측정이 갖는 구조적·환경적 제약이 복합적으로 작용한 결과로 해석된다.

첫째, 촬영 시 착용한 의류의 두께와 주름은 신체 실루엣에 직접적인 영향을 미친다. 특히 가슴과 복부 영역은 의복의 밀착 정도에 따라 외곽선이 실제 신체 단면보다 크게 혹은 작게 표현될 수 있으며, 이는 알파 채널 기반 경계 탐색 과정에서 측정 폭의 변동성을 증가시키는 요인으로 작용하였다. 둘째, 호흡 상태에 따른 흉곽 및 복부의 확장·수축은 가슴둘레와 배둘레 측정에 본질적인 변동성을 유발한다. 이는 정적 이미지 기반 측정 방식이 갖는 근본적인 한계로, 단일 시점에서 촬영된 이미지로는 동적 생체 변화를 완전히 반영하기 어렵다.

셋째, 체형 유형(마른형, 근육형, 복부 비만형 등)에 따라 동일한 정면·측면 폭을 갖더라도 실제 단면의 형태는 원형, 타원형, 비대칭형 등으로 상이할 수 있다. 즉, 동일한 폭(width)을 갖는 경우에도 내접 타원의 둘레는 달라질 수 있으며, 이는 2D 이미지 기반 둘레 추정에서 발생하는 근본적인 기하학적 제약이다. 이러한 한계는 단순 알고리즘 보정만으로 해결되기 어렵고, 구조적으로는 3D 정보 또는 다각도 영상의 도입이 요구되는 영역이다.

다만 본 연구에서는 이러한 한계를 단순히 성능 저하 요인으로만 해석하지 않고, 향후 시스템 고도화를 위한 명확한 개선 방향으로 인식하였다. 예를 들어, 체형 분류(classification) 기반 회귀 모델 분기, 복부 돌출 정도를 반영하는 추가 특징 변수 도입, 또는 제한적인 다중 각도 촬영을 통한 준(準)-3D 형태 추정 방식 등이 현실적인 개선 방안으로 도출되었다.

4.3 연구 결과 및 시사점

본 연구는 AI 기반 2D 이미지 스캐닝 기술을 활용하여, 기존 수작업 중심

의 신체 치수 측정 방식을 자동화할 수 있는 실용적이고 구현 가능한 신체 치수 측정 시스템을 개발하는 것을 목표로 수행되었다. 이를 위해 총 204명의 성인 남성을 대상으로 신체 치수 실측 데이터와 정면·측면 전신 이미지를 수집하였으며, 배경 제거(rembg, U^2 -Net), 신체 랜드마크 추출(MediaPipe Pose), 알파 채널 기반 신체 경계 탐색, 그리고 TensorFlow 기반 DNN 회귀 모델을 단계적으로 결합한 통합 시스템을 구현하였다.

연구 결과, 길이 계열 항목에서 약 99%, 일부 둘레 항목에서 약 98%의 높은 측정 정확도를 달성함으로써, 2D 이미지 기반 신체 측정 방식이 실제 현장 적용이 가능한 수준의 정밀도를 확보할 수 있음을 실증적으로 확인하였다. 이는 군복 제작, 기능성 의류, 스포츠웨어, 맞춤형 보호장비 등 대량 인원에 대한 신체 치수 수집이 요구되는 산업 분야에서 본 시스템이 충분한 실용성을 가질 수 있음을 시사한다.

특히 본 연구는 단순한 알고리즘 구현을 넘어, 25년 이상 현장에서 축적된 실무 경험을 바탕으로 “이론적으로 가능한 모델”이 아닌 “현실적으로 작동하는 시스템”을 지향하였다. 촬영 환경, 복장 조건, 처리 시간, 시스템 복잡도 등 실제 적용 환경에서 발생할 수 있는 제약을 고려하여 설계되었다는 점에서, 본 연구의 결과는 학술적 기여뿐만 아니라 실무적 가치 또한 갖는다.

결론적으로, 본 연구는 2D 이미지 기반 신체 측정 기술의 가능성과 한계를 동시에 규명하였으며, 향후 3D 정보의 부분적 도입이나 체형 인식 기반 보정 모델을 통해 더욱 고도화된 신체 치수 측정 시스템으로 확장될 수 있는 기반을 마련하였다. 이는 AI 기반 스마트 피팅 및 자동화 계측 기술 연구의 실질적인 출발점으로서 의미 있는 기여를 한다고 판단된다.

Ⅴ. 결 론

5.1 결론 및 시사점

본 연구는 AI 기반 2D 이미지 분석 기술을 활용하여 의류 디자인 산업에 적용 가능한 신체 치수 자동 측정 시스템을 개발하고 그 정확도를 검증하였다. Google의 MediaPipe Pose 프레임워크와 라마누잔(Ramanujan) 타원 둘레 근사 공식을 결합하여, 정면과 측면 2장의 이미지만으로 목둘레, 어깨너비, 소매길이, 총장, 상기장, 가슴둘레, 배둘레, 허리둘레, 엉덩이둘레, 허벅지둘레, 하기장 등 11개 항목의 신체 치수를 추출하는 시스템을 구현하였다.

실험 결과, 주요 둘레 측정 항목에서 평균 ±1.5cm 이내의 오차를 달성하여 의류 산업에서 일반적으로 허용되는 오차 범위(±2cm) 이내의 정확도를 확보하였다. 특히 어깨너비 측정에서 MAE 1.0cm, 상관계수 0.95의 가장 높은 정확도를 보였으며, 가슴둘레, 허리둘레, 엉덩이둘레 등 핵심 둘레 항목에서도 상관계수 0.93~0.95 수준의 높은 정확도를 달성하였다. 이러한 결과는 고가의 3D 스캐닝 장비를 대체하여 저비용으로 산업적 요구 수준을 충족하는 신체 치수 측정이 가능함을 실증적으로 입증한 것이다.

본 연구의 기술적 의의는 다음과 같다. 첫째, 기존의 고가 3D 스캐너 대비 현저히 낮은 비용으로 유사한 수준의 측정 정확도를 달성하였다. 스마트폰 카메라 한 대와 오픈소스 소프트웨어만으로 전문 측정 장비 수준의 결과를 도출할 수 있다는 점은 기술의 민주화 측면에서 중요한 진전이다. 둘째, MediaPipe Pose를 활용함으로써 별도의 딥러닝 모델 학습 없이도 높은 정확도의 포즈 추정이 가능하였으며, 이는 시스템의 즉시 배포 가능성과 유지보수 용이성을 높인다. 셋째, 라마누잔 근사 공식을 통한 둘레 변환 방식은 계산 효율성과 정확도를 동시에 확보하여 실시간 처리가 가능한 경량 시스템 구현의 기반을 마련하였다.

산업적 활용 가능성 측면에서 본 시스템은 다양한 분야에 적용될 수 있다. 온라인 의류 쇼핑몰에서는 고객의 신체 치수를 원격으로 측정하여 맞춤형 사이즈 추천 서비스를 제공할 수 있으며, 이를 통해 사이즈 불일치로 인한 반품

률을 획기적으로 감소시킬 수 있다. 맞춤 정장, 웨딩드레스, 한복 등 고가 맞춤 의류 시장에서는 초기 치수 측정 단계의 비용과 시간을 절감할 수 있다. 피트니스 및 헬스케어 분야에서는 체형 변화 추적 서비스에 활용하여 개인 맞춤형 건강 관리 솔루션을 제공할 수 있다. 또한 정형외과, 재활의학, 스포츠 과학, 인체공학 등 인체 계측이 필요한 다양한 분야로의 확장 적용이 가능하다.

ESG(환경·사회·지배구조) 경영 관점에서 본 연구는 중요한 시사점을 제공한다. 환경적(E) 측면에서는 고가의 3D 스캐닝 장비 제조 및 폐기 과정에서 발생하는 탄소 배출을 줄이고, 기존 스마트폰을 활용함으로써 추가적인 전자 폐기물 발생을 억제할 수 있다. 사회적(S) 측면에서는 기술 접근성이 향상되어 중소기업, 지역 기반 산업, 개발도상국 등 자원 제약이 큰 조직에서도 고성능 계측 기술을 활용할 수 있게 된다. 특히 거동이 불편한 고령자나 장애인의 경우 원격 측정을 통해 맞춤 의류 서비스에 대한 접근성이 크게 향상될 수 있다. 지배구조(G) 측면에서는 데이터 기반 의사결정을 촉진하여 디지털 전환에 뒤처진 전통 산업군의 경쟁력 향상에 기여한다.

적정기술(Appropriate Technology) 관점에서 본 연구의 결과는 글로벌 저비용 솔루션으로서 의의를 갖는다. 전문 인력에 대한 의존도를 낮추고, 산업별 디지털 격차 해소에 기여할 수 있다. 특히 글로벌 경기 침체 상황에서 비용 효율적인 기술 솔루션에 대한 수요가 증가하고 있으며, 본 연구에서 제시한 AI 기반 2D 스캐닝 기술은 이러한 시장 수요에 부합하는 기술적 대안을 제공한다. 이는 국가 단위의 디지털 제조혁신, 공공 보건 데이터 구축, 개인 맞춤형 서비스 산업 등 다양한 영역에서 국가 산업 경쟁력 강화에 기여할 수 있을 것으로 기대된다.

5.2 연구의 한계점

본 연구에는 다음과 같은 한계점이 존재한다.

첫째, 촬영 환경에 대한 의존성이 있다. 입력 이미지의 해상도, 조명 조건, 배경 복잡도에 따라 MediaPipe Pose의 랜드마크 검출 정확도가 영향을 받으며, 이는 전체 측정 오차로 이어질 수 있다. 본 연구에서는 500lux 이상의 조명과 단순 배경 조건에서 최적의 결과를 얻었으나, 다양한 실사용 환경에서의 강건성 확보를 위해서는 추가적인 전처리 및 보정 알고리즘이 필요하다.

둘째, 측정 조건의 표준화 요구가 있다. 본 시스템은 정면 T-포즈와 측면 직립 자세, 밀착형 의류 착용을 전제로 설계되었다. 피측정자가 지시된 자세를 정확히 취하지 않거나 두꺼운 의복을 착용한 경우 측정 오차가 증가할 수 있으며, 이를 보완하기 위한 실시간 자세 가이드 기능이나 의복 보정 알고리즘의 개발이 향후 과제로 남아 있다.

셋째, 데이터셋의 다양성 제한이 있다. 본 연구의 실험 데이터셋은 204명의 성인 남성으로 구성되어 여성, 아동, 고령자 및 다양한 체형을 포함하지 못하였다. 향후 연구에서는 더 다양한 인구 집단을 포함한 데이터셋 구축을 통해 모델의 일반화 성능을 검증할 필요가 있다.

5.3 향후 연구방향

향후 연구에서는 다음과 같은 방향으로 시스템을 개선하고자 한다. 첫번째, 실시간 배경 제거 및 자세 가이드 기능을 탑재한 모바일 애플리케이션을 개발하여 일반 사용자의 접근성을 높인다. 카메라 프리뷰 화면에 가이드라인을 표시하고, 올바른 자세가 감지되었을 때 자동으로 촬영이 이루어지도록 하여 사용자 편의성과 측정 일관성을 동시에 확보한다.

두번째, 딥러닝 기반 의복 보정 알고리즘을 개발하여 다양한 착의 조건에서도 정확한 측정이 가능하도록 한다. 의복의 종류와 두께를 인식하고, 이에 따른 보정 계수를 자동으로 적용하는 방식을 연구한다.

세번째, 더 다양한 체형, 연령대, 성별, 인종을 포함한 대규모 데이터셋을 구축하여 모델의 일반화 성능을 향상시킨다. 특히 Size Korea 등 국가 인체치수 데이터베이스와의 연계를 통해 한국인 체형에 최적화된 모델 개발을 추진한다.

네번째, TensorFlow Lite, ONNX, Core ML 등으로의 모델 변환을 통해 다양한 엣지 디바이스에서의 실시간 온디바이스(On-Device) 처리를 구현한다. 이를 통해 네트워크 연결 없이도 측정이 가능하고, 개인 신체 데이터의 프라이버시를 보호할 수 있는 시스템을 구축한다.

다섯번째, 멀티뷰(Multi-View) 융합 기술을 도입하여 3장 이상의 이미지(정면, 측면, 후면 또는 45도 각도)를 활용한 360도 신체 형상 복원 기능을 개발한다. 이를 통해 단일 뷰에서 발생하는 깊이 모호성(Depth Ambiguity) 문제를 해결하고, 측정 정확도를 더욱 향상시킬 수 있을 것으로 기대된다.

본 연구에서 제시한 AI 기반 2D 이미지 스캐닝 신체 치수 측정 기술은 의류 산업의 디지털 전환을 가속화하고, ESG 경영과 적정기술 확산에 기여할 수 있는 잠재력을 갖추고 있다. 향후 지속적인 연구개발을 통해 기술적 한계를 극복하고, 더욱 다양한 산업 분야에서 활용될 수 있도록 시스템을 고도화해 나갈 것이다.

참 고 문 헌

1. 국내문헌

김희은 외. (2017). 한국인 인체치수조사 방법론 및 데이터: 사이즈코리아 프로젝트. International Journal of Industrial Ergonomics, 60, 25-32.

남윤자 외. (2004). Size Korea 3차원 인체측정방법 표준화. 패션정보와 기술, 1, 19-26.

백경자 외. (2008). 3차원 바디 스캐너를 이용한 남성 상반신 인체측정 - 직접측정과의 비교. 한국의류학회지, 32(1), 29-37.

한희경 외. (2010). Comparative analysis of 3D body scan measurements and manual measurements of size Korea adult females. International Journal of Industrial Ergonomics, 40(5), 530-540.

남윤자 외. (2001). 삼차원 인체 레이저 스캔 데이터를 이용한 남성 재킷 원형 설계방법에 관한 연구. 한국의류학회지, 25(5), 835-846.

김혜경. (2001). 3D Scanner를 이용한 여성용 기성복 재킷의 착의적합성에 관한 비교평가연구. 한국의류학회지, 25(10), 1707-1718.

유희천 외. (2008). 헬리콥터 조종실 설계를 위한 헬리콥터 조종사 인체측정 및 분석. 대한인간공학회지, 27(4), 37-44.

이원섭 외. (2013). An anthropometric analysis of Korean male helicopter pilots for helicopter cockpit design. Ergonomics, 56(5), 879-887.

한국전자통신연구원. (2018). 딥러닝 기반 객체 분류 및 검출 기술 분석 및 동향. 전자통신동향분석, 33(4), 33-44.

한국전자통신연구원. (2020). 딥러닝기반 입체 영상의 획득 및 처리 기술 동향. 전자통신동향분석, 35(5), 112-122.

이지용 외. (2024). 헬스케어 분야에서 활용 가능한 AI 기반 체형 3D 모델링 기술 개발. 한국전자통신학회 논문지, 19(3), 509-516.

한국컴퓨터그래픽스학회. (2022). K-SMPL: 한국인 체형 특성을 반영한 3D 인체 모델. 한국컴퓨터그래픽스학회 논문지, 28(4), 1-12.

서동애 외. (2012). 3D 인체 스캔 데이터를 활용한 체형별 인대모형 개발에 관한 연구. 한국콘텐츠학회 논문지, 12(9), 136-145.

조신현 외. (2012). 사이즈코리아 제5차, 6차 3D 인체형상데이터 비교를 통한 30대 여성 신체변화 연구. 한국패션디자인학회지, 12(1), 87-102.

김소영 외. (2020). Body shape classification of Korean middle-aged women using 3D anthropometry. Fashion and Textiles, 7, 34.

김윤상 외. (2019). 3D 스캔 데이터에 의한 성인 남성의 체간부 형태 유형화. 한국의상디자인학회지, 21(4), 165-179.

권미정 외. (2018). 인체의 전면 및 후면 3D 스캔 영상을 이용한 체형 분석 방법. 대한전자공학회 논문지, 55(8), 37-44.

윤승현. (2004). 3차원 인체치수측정과 형상변형 기술. 패션정보와 기술, 1(45), 54-61.

강여선. (2007). 성인 남자 의류 치수체계 개발을 위한 신체 치수 및 체형 분석 (제2보). 한국의류학회지, 31(2), 247-257.

정선영. (2008). 한국 성인 여성의 3차원 인체 형상 라이브러리 구축에 관한 연구. 서울대학교 박사학위논문.

최영림. (2008). 매스 커스터마이제이션을 위한 상반신 측면형태 분류와 패턴 변형. 서울대학교 석사학위논문.

박성준 외. (2014). 3D Hand Anthropometry of Korean Teenager's and Comparison with Manual Method. HCI International 2014, Communications in Computer and Information Science, 435, 485-490. Springer.

한국인공지능연구원 AI Hub. (2021). 한국인 신체 3D 스캐닝 데이터.

2. 국외문헌

Ashby, N., LaPorte, G. J., Richardson, D., Scioletti, M., Heymsfield, S. B., Shepherd, J. A., & Tinsley, G. M. (2023). Translating digital anthropometry measurements obtained from different 3D body image scanners. European Journal of Clinical Nutrition, 77(9), 872–880.

Bartol, K., Bojanić, D., Petković, T., & Pribanić, T. (2021). A review of body measurement using 3D scanning. IEEE Access, 9, 67281–67301.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). BlazePose: On-device real-time body pose tracking. CVPR Workshop on Computer Vision for Augmented and Virtual Reality.

Bogo, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., & Black, M. J. (2016). Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. Proceedings of the European Conference on Computer Vision (ECCV), 561–578.

Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2021). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(1), 172–186.

D'Apuzzo, N. (2007). 3D body scanning technology for fashion and apparel industry. Proceedings of SPIE, 6491, 64910O.

Farina, G. L., Orlandi, C., Cremascoli, R., Cè, E., Esposito, F., & Galvani, C. (2025). Advances in the estimation of body fat

percentage using an artificial intelligence 2D-photo method. npj Digital Medicine, 8, 18.

Farkas, L. G. (1996). Accuracy of anthropometric measurements: Past, present, and future. Cleft Palate-Craniofacial Journal, 33(1), 10-22.

Franke-Gromberg, C., Schüler, G., Hermanussen, M., & Scheffler, C. (2010). Digital 2D-photogrammetry and direct anthropometry - A comparing study on test accomplishment and measurement data. Anthropologischer Anzeiger, 68(1), 11-20.

García-Muro San José, F., González-Fernández, L., & García-López, J. (2024). Emerging roles of 3D body scanning in human-centric applications. Technologies, 13(4), 126.

Han, K., Kwon, H. J., Choi, T. H., Kim, J. H., & Son, D. (2010). Comparison of anthropometry with photogrammetry based on a standardized clinical photographic technique using a cephalostat and chair. Journal of Cranio-Maxillofacial Surgery, 38(2), 96-107.

Heymsfield, S. B., Bourgeois, B., Ng, B. K., Sommer, M. J., Li, X., & Shepherd, J. A. (2018). Digital anthropometry: A critical review. European Journal of Clinical Nutrition, 72(5), 680-687.

Heymsfield, S. B., & Shepherd, J. A. (2024). Body composition measurement from images. Annual Review of Nutrition, 44, 237-261.

Kanazawa, A., Black, M. J., Jacobs, D. W., & Malik, J. (2018). End-to-end recovery of human shape and pose. Proceedings of the IEEE Conference on Computer Vision and Pattern

Recognition (CVPR), 7122-7131.

Khan, F., Salahuddin, S., & Javidnia, H. (2020). Deep learning-based monocular depth estimation methods - A state-of-the-art review. Sensors, 20(8), 2272.

Leymarie, F. F., & de Aguiar, E. (2023). Designing a contactless, AI system to measure the human body using a single camera for the clothing and fashion industry. Scientific Reports, 13, 13876.

Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., & Black, M. J. (2015). SMPL: A skinned multi-person linear model. ACM Transactions on Graphics (Proc. SIGGRAPH Asia), 34(6), 248:1-248:16.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. arXiv preprint arXiv:1906.08172.

Majmudar, M. D., Chandra, S., Yakkala, K., Kennedy, S., Agrawal, A., Sippel, M., ... & Thakor, N. V. (2022). Smartphone camera based assessment of adiposity: A validation study. npj Digital Medicine, 5(1), 79.

Mitchell, H. L. (1992). History of photogrammetry in medicine. SPIE Proceedings, 1380, 62-68.

Ng, B. K., Hinton, B. J., Fan, B., Kanaya, A. M., & Shepherd, J. A. (2016). Clinical anthropometrics and body composition from 3D whole-body surface scans. European Journal of Clinical Nutrition, 70(11), 1265-1270.

Okimoto, M. L. R., & Klein, A. A. (2012). Photogrammetry procedures applied to anthropometry. Work, 41(Suppl 1), 4046-4052.

Pateraki, M., Fragkoulidou, V., & Stoltidou, A. (2006). Photogrammetry vs. Anthroposcopy. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(Part 5).

Rumbo-Rodríguez, L., Sánchez-SanSegundo, M., Ferrer-Cascales, R., Ruiz-Robledillo, N., Clement-Carbonell, V., & Zaragoza-Martí, A. (2021). Comparison of body scanner and manual anthropometric measurements of body shape: A systematic review. International Journal of Environmental Research and Public Health, 18(12), 6213.

Schonberger, J. L., & Frahm, J. M. (2016). Structure-from-motion revisited. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4104-4113.

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1, 519-528.

Smith, B., McCarthy, C., Dechenaud, M. E., Wong, M. C., Shepherd, J., & Heymsfield, S. B. (2022). Anthropometric evaluation of a 3D scanning mobile application. Obesity, 30(6), 1181-1188.

Staynor, J. M. D., Cronin, N. J., Cools, P., & Delahunt, E. (2022). Agreement of anthropometric and body composition measures predicted from 2D smartphone images and body impedance scales with criterion methods. Obesity Research & Clinical Practice, 16(1), 37-43.

Tinsley, G. M., Harty, P. S., Siedler, M. R., Stratton, M. T., & Rodriguez, C. (2024). Mobile phone applications for

3-dimensional scanning and digital anthropometry: A precision comparison with traditional scanners. European Journal of Clinical Nutrition, 78(5), 452-458.

Treleaven, P., & Wells, J. (2007). 3D body scanning and healthcare applications. Computer, 40(7), 28-34.

Ulijaszek, S. J., & Kerr, D. A. (1999). Anthropometric measurement error and the assessment of nutritional status. British Journal of Nutrition, 82(3), 165-177.

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. UN General Assembly.

Wang, J., Gallagher, D., Thornton, J. C., Yu, W., Horlick, M., & Pi-Sunyer, F. X. (2006). Validation of a 3-dimensional photonic scanner for the measurement of body volumes, dimensions, and percentage body fat. American Journal of Clinical Nutrition, 83(4), 809-816.

Weinberg, S. M., Scott, N. M., Neiswanger, K., Brandon, C. A., & Marazita, M. L. (2004). Digital three-dimensional photogrammetry: Evaluation of anthropometric precision and accuracy using a Genex 3D camera system. Cleft Palate-Craniofacial Journal, 41(5), 507-518.

Wells, J. C. K., Treleaven, P., & Cole, T. J. (2007). BMI compared with 3-dimensional body shape: The UK National Sizing Survey. American Journal of Clinical Nutrition, 85(2), 419-425.

Zhao, C., Sun, Q., Zhang, C., Tang, Y., & Qian, F. (2020). Monocular depth estimation based on deep learning: An overview. Science China Technological Sciences, 63(9), 1612-1627.

3. 정부기관 및 연구기관

국가기술표준원. (2004). 제5차 한국인인체치수조사사업 인체치수결과보고서.

국가기술표준원. (2010). 제6차 한국인인체치수조사사업 인체치수결과보고서.

국가기술표준원. (2021). 제8차 한국인 인체치수조사사업 결과보고서. 산업통상자원부.

한현숙. (2017). 의류산업용 3차원 인체모델 생성 및 변형방법 개발. 고려대학교 연구보고서.

ABSTRACT

Development of AI-Based Deep Learning Algorithm for Image Scanning: Focusing on the Apparel Design Industry

Kim, Jung-In

Major in Knowledge Service & Consulting

Dept. of Smart Convergence Consulting

Graduate School of Knowledge Service & Consulting Hansung University

Image scanning technology originated in the late 19th century with the advent of photogrammetry, initially relying on manual dimension extraction and measurement methods, with significant variability depending on the operator's proficiency. From the mid-20th century, the proliferation of optical equipment, photography, and raster-based scanning technologies led to the widespread adoption of 2D image-based measurement. The commercialization of digital sensors further expanded image scanning applications across various fields, including industry, healthcare, fashion, safety, and sports. While 2D scanning technology provided quantified data for human body measurement, limitations in

accuracy persisted due to variables such as image distortion, shooting angles, lighting, and resolution.

From the late 1990s, technologies including Structured Light, laser scanning, and Time-of-Flight sensors were introduced to overcome these limitations, enabling the acquisition of precise surface geometry and volumetric information. Bartol et al. (2021) reported through a systematic literature review on 3D scanning-based body measurement technology that its utilization has significantly increased in custom pattern production for the apparel industry, body shape analysis in orthopedic and rehabilitation medicine, and sports and healthcare applications. However, 3D scanners have structural limitations including high equipment costs, spatial constraints, the need for specialized personnel, and maintenance expenses, making adoption difficult for small and medium-sized enterprises (SMEs), small business owners, and startups.

In this context, the advancement of AI algorithm-based 2D body measurement technology has recently gained significant attention. Conventional 2D scanning methods offer advantages in cost and accessibility, as data collection is possible with just a smartphone or entry-level camera. However, limitations existed in achieving dimensional precision due to the inability to measure explicit depth. Recently, with the rapid development of various AI technologies including deep learning-based pose estimation, Monocular Depth Estimation, and parametric human model-based 3D reconstruction, body dimension estimation approaching 3D scanning accuracy has become possible from 2D images. In particular, SMPL (Skinned Multi-Person Linear Model) proposed by Loper et al. (2015) is a statistical human body model capable of parametrically representing body shape and pose based on 3D scan data from thousands of individuals, establishing itself as a core foundational technology for research on reconstructing 3D human body

shapes from single 2D images.

This study aims to verify the accuracy of AI-based 2D scanning body measurement technology based on these technological advancements and analyze the potential to replace 3D scanning technology. The study experimentally verifies accuracy variations according to shooting environment, resolution, camera specifications, and clothing conditions, while also proposing the socio-economic implications of 2D AI scanning from ESG management and appropriate technology perspectives. The ability to measure dimensions with low-cost equipment offers advantages in terms of reduced carbon emissions from equipment manufacturing and disposal (Environmental), enhanced technology accessibility for SMEs and developing countries (Social), and promotion of digital transformation (Governance).

This study empirically examines whether AI-based 2D scanning technology can replace the technical and economic limitations of 3D scanning in SME industrial settings. The research findings are expected to contribute to improving accessibility in the anthropometric industry, bridging technological gaps, and establishing low-cost digitalization strategies. Furthermore, the results are anticipated to be applicable across various domains including national-level digital manufacturing innovation, public health data construction, and personalized service industries.

Keywords: Deep Learning, 2D Body Measurement, Image Scanning, SMPL, Pose Estimation, Apparel Design Industry, Appropriate Technology