

YOLOv8-pose 기반의 실시간 스쿼트 자세 분석 및 교정 시스템 구현

정재성, 신은호, 전윤호*

국립한밭대학교

{json, eunho}@edu.hanbat.ac.kr, *yhjeon@hanbat.ac.kr

Real-time Squat Posture Analysis and Calibration System Based on YOLOv8-pose systems

Jaeseong Jeong, Eunho Shin, *Yunho Jeon

Hanbat National University

요약

본 논문은 인공지능 기반 실시간 자세 인식 기술을 활용하여 사용자의 스쿼트 동작을 정확하게 감지하고, 즉각적인 피드백을 제공하는 시스템을 제안한다. 본 시스템은 YOLO 포즈 추정 모델을 이용하여 어깨, 엉덩이, 무릎, 발목 등 주요 신체 키포인트를 검출하고, Intersection over Union(IoU) 기반 위치 검증 및 기준점 설정(Calibration)을 통해 사용자의 초기 자세를 보정한다. 또한, 어깨너비 변화를 이용한 각도 추정 방식을 적용하여 사용자의 회전 각도를 계산하고, 목표 각도와 비교하여 실시간 안내 메시지를 제공한다. 제안된 시스템은 상태 머신 기반 스쿼트 감지 로직을 통해 동작을 READY, DESCENDING, BOTTOM, ASCENDING, INVALID의 다섯 단계로 구분하며, 단계별 피드백을 시각적으로 제공한다.

1. 서론

최근 YOLOv8-pose[1]와 같은 실시간 포즈 추정 모델의 발전은 단일 2D 카메라만으로도 다양한 상호작용 시스템을 구현할 잠재력을 열어주었다. 그러나 스쿼트와 같이 복합적인 동작을 Raspberry Pi와 같은 엣지 디바이스 환경에서 정밀하게 분석하는 것은 모델의 추론 성능과 2D 데이터의 본질적인 한계로 여전히 어려운 과제로 남아있다.

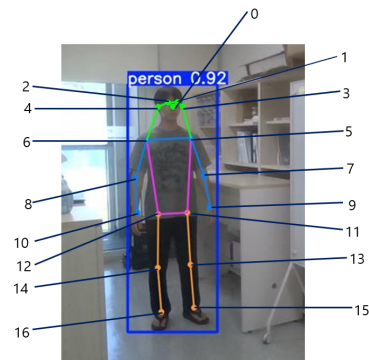
이러한 문제를 해결하기 위해, 2D 키포인트 데이터의 신뢰도를 높이기 위한 사용자 위치 보정 및 다층적 검증 알고리즘을 제안한다. 제안된 시스템은 의도된 운동 동작과 일상적인 움직임을 구분하는 필터링 매커니즘을 포함하여, 리소스가 제한된 환경에서도 안정적인 분석을 수행하도록 설계되었다.

이를 통해 저비용의 엣지 디바이스 상에서 딥러닝 모델을 활용한 실시간 자세 판정 시스템의 기술적 구현 가능성을 검증하고, 향후 정밀한 동작 분석 기술의 기반을 마련하고자 한다.

II. 관련 연구

2.1 Human Pose Estimation

사람의 신체 관절 위치를 추정하는 HPE(Human Pose Estimation)는 컴퓨터 비전 분야에서 꾸준히 연구됐다. DeepPose[2]는 최초로 딥러닝 기반 회귀방식을 포즈 추정에 적용하여 입력 이미지로부터 직접 사람의 신체 관절 위치를 나타내는 키포인트를 예측함으로써 딥러닝 접근법의 가능성을 입증하였다. 이후 객체 탐지와 포즈 추정을 결합한 YOLO 계열의 경량 모델들이 등장하였으며, 그중 YOLOv8-pose[1]는 사람을 바운딩 박스로 검출한 뒤 [그림 1]과 같이 어깨, 엉덩이, 무릎과 같은 17개의 신체 주요 키포인트를 직접 회귀하는 방식으로 높은 정확도와 실시간 성능을 동시에 확보하였다. 이러한 특성은 운동 자세 분석 및 교정과 같은 실시간 피드백이 중요한 응용 분야에 적합하며, 본 연구에서는 YOLOv8-pose[1]기반 포즈 추정 모델을 활용하여 스쿼트 동작을 효과적으로 인식할 수 있다.



[그림 1] 각 키포인트 위치

2.2 경량 딥러닝 모델

YOLOv8-pose와 같은 딥러닝 기반 인체 자세 추정 모델은 높은 정확도를 제공하지만, 많은 파라미터와 복잡한 연산이 필요하다. 이러한 특성은 고성능 장비에서는 문제가 되지 않지만, Raspberry Pi와 같은 저성능 엣지 디바이스에서는 실시간 추론이 어렵다는 한계가 있다.

이를 해결하기 위해 모델의 크기와 연산량을 줄이는 경량화 기법이 연구 [3]되고 있다. 대표적으로, 가중치의 정밀도를 낮추어 모델 크기와 속도를 개선하는 양자화, 불필요한 뉴런이나 연결을 제거해 구조를 단순화하는 가지치기(Pruning)가 있다.

본 연구에서는 이러한 경량화 기법이 적용된 YOLOv8-pose 모델을 활용하여, Raspberry Pi 환경에서도 실시간 자세 추정이 가능하도록 시스템을 구현하였다.

III. 본론

3.1 시스템 설계 및 구현

본 연구에서 제안하는 실시간 스쿼트 분석 시스템은 Raspberry Pi 환경에서 구현되었으며, 단일 2D 카메라와 YOLOv8-pose 모델을 기반으로 사용자의 스쿼트 수행 과정을 정량적으로 분석하고 그 유효성을 실시간으로 판별하는 것을 목표로 한다. 본 시스템에 사용된 실시간 포즈 추정 모델은 선행 연구[4]의 결과를 기반으로 최적화되었다. 시스템의 전체적인

구성은 1) 초기화 및 사용자 위치 보정, 2) 다층적 실시간 자세 검증, 3) 상태 머신 기반 동작 분석의 세 가지 핵심 모듈로 구성된다.

3.2 초기화 및 사용자 위치 보정

정확하고 일관된 분석을 위해 시스템은 시작 전, 사용자의 물리적 특성과 위치를 보정하는 초기화 단계를 거친다. 이는 개인별 신체 조건과 환경 차이를 표준화하여 분석의 신뢰도를 확보하기 위한 과정이다.

우선 IoU(Intersection over Union)를 이용해 사용자가 화면의 유효 영역 내에 있는지 확인하고, 스쿼트 분석에 필요한 주요 키포인트가 안정적으로 감지되는지 검증한다. 이후 2D 영상에서 무릎 각도를 원활히 측정할 수 있도록 약 45° 사선 뷰를 사용자가 설정하도록 보정이 이루어진다. 이 과정은 (1) 정면을 바라보게 하여 어깨너비를 측정하고 최댓값을 기준으로 저장한 뒤, (2) 현재 어깨너비와의 비율을 활용해 몸의 각도를 $\arccos$ 함수로 추정하며, (3) 추정된 각도를 목표 범위(40°~50°)에 맞추도록 안내하는 절차로 구성된다.

보정이 완료되면 스쿼트 깊이 측정을 위한 초기 엉덩이 높이, 각도 추정 및 발 너비 검증에 필요한 어깨너비 등 개인화된 기준점이 저장된다. 이 기준을 토대로 판정에 활용되는 임계값들이 설정되며, 구체적인 값과 조건은 [표 1]과 같다.

항목	기준값	정의 및 측정 방식
상위 무릎 각도	160°	직립 자세에서 하강 시작을 판단하는 기준 각도 엉덩이(11/12)-무릎(13/14)-발목(15/16) 키포인트의 내각
하위 무릎 각도	130°	유효한 스쿼트 깊이 판별 기준 각도 (상위 무릎 각도와 동일한 키포인트로 계산)
최소 엉덩이 하강 거리	40px	최소 운동 범위 보장을 위한 엉덩이 절대 하강 거리 초기 자세 대비 엉덩이(11, 12) 키포인트의 y좌표 변화량
좌우 무릎 각도 차이	20°	좌우 균형 평가를 위한 절대 허용 각도 편차 좌측 키포인트: 11-13-15, 우측 키포인트: 12-14-16
발 너비 최소 비율	0.5	안전성을 위한 어깨너비 대비 발 너비 최소 비율 발목(15-16)과 어깨(5-6) 키포인트 간 수평 거리
발 너비 최대 비율	1.5	과도한 자세 방지를 위한 어깨너비 대비 발 너비 최대 비율 (최소 비율과 동일 방식)

[표 1] 스쿼트 자세 분석을 위한 주요 임계값

3.3 다층적 실시간 자세 검증

시스템은 프레임마다 스쿼트 동작의 유효성을 우선순위에 따라 다층적으로 검증한다. 최우선으로, 사용자의 엉덩이 중심이 무릎 중심보다 낮아지는 ‘Sit Down’ 동작을 감지하여 즉시 무효로 하는 정확성 검증을 수행한다. 이 동작이 감지되면 시스템은 다른 모든 분석을 중단하고 무효 상태로 전환되어 사용자에게 완전한 자세 재설정을 요구한다.

이 무효 상태에서 벗어나기 위해서 사용자가 단순히 일어서는 것을 넘어 완전한 초기 직립 자세로 복귀해야만 한다. 시스템은 사용자의 무릎 각도와 엉덩이 높이가 초기 준비 상태의 기준을 다시 만족했을 때만 무효 상태를 해제하고 분석을 재개한다. 이처럼 명확한 회복 조건을 둬으로써 잘못된 동작이 다음 동작 분석에 영향을 미치는 것을 차단한다. 이 정확성 검증을 통과한 경우에만 앞서 3.2절에서 정의한 임계값들을 바탕으로 사용자의 기본자세를 순차적으로 검증한다.

3.4 상태 머신 기반 동작 분석

본 시스템은 스쿼트 동작의 흐름을 정밀하게 분석하기 위해 무효화 및 회복 메커니즘과 4단계 정상 동작 분석 로직이 결합한 이중 구조의 상태 머신을 구성하였다.

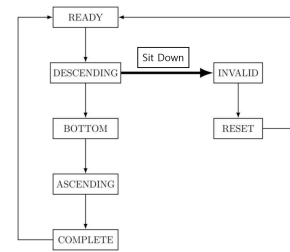
스쿼트 동작은 준비(READY), 하강(DESCENDING), 최저점(BOTTOM), 상승(ASCENDING)의 네 가지 상태로 정의된다. 각 상태는 사용자의 평균 무릎 각도와 개인화된 엉덩이 하강 깊이를 종합적으로 고려하였다.

READY 상태는 사용자가 스쿼트를 시작할 준비가 된 초기 직립 자세로, 무릎 각도가 일정 기준 이상일 때 유지된다. 이 상태에서는 안정적인 준비를 확인하기 위한 ‘Ready’ 피드백이 제공된다. 여기서 무릎 각도가 기준치 아래로 줄어들면 DESCENDING 상태로 넘어간다.

DESCENDING 상태는 하강 단계로, 이 구간에서는 무릎 각도와 엉덩이 하강 깊이를 함께 확인한다. 하강 깊이가 충분하면 ‘Good!’ 피드백과 함께 BOTTOM으로 전이되며 올바른 자세로 판단한다. 깊이가 부족할 경우 ‘Go deeper!’ 피드백을 제공하여 더 깊은 자세를 유도하며, 조기 상승이 발생하면 ‘Too early!’ 피드백과 함께 READY 상태로 복귀한다.

BOTTOM 상태는 동작 반전 지점으로, 무릎 각도가 다시 기준을 초과하기 시작하면 ASCENDING 상태로 전이된다. 이때는 ‘Up’ 피드백을 제공하여 사용자가 상승 동작 개시를 인지할 수 있도록 한다.

ASCENDING 상태에서는 무릎 각도와 엉덩이 높이가 초기 조건을 충족할 때 한 동작이 완전히 수행된 것으로 인정되며 이에 따라 카운트가 증가한다. 성공적인 수행 시 ‘Complete!’ 피드백과 시각적 체크 표시가 제공된다.



[그림 2] 시스템 흐름도

IV. 결 론

단일 카메라와 YOLO 기반의 실시간 포즈 추정 기술을 활용하여 사용자의 스쿼트 동작을 정밀하게 분석하고 교정하는 시스템을 설계·구현하였다. 제안된 시스템은 사용자의 신체 비율에 맞춰 개인화된 기준점을 설정하고, 다층적 검증 구조와 상태 머신을 통해 스쿼트 동작의 정확성, 균형 등을 종합적으로 평가한다. 특히 의도된 스쿼트 자세와 주저앉는 동작을 구분하는 동작 유효성 검증 로직을 도입하여 분석의 신뢰도를 높였다. 이러한 점에서 단순 횡수 측정을 넘어 수행 과정의 유효성을 정량적으로 판별하고 구체적인 피드백을 제공한다는 데 의의가 있다.

그러나 2D 영상 기반 분석에 따른 한계를 가진다. 예를 들어 긴 바지와 같은 의복 착용 시 무릎 검출 정확도가 저하되며, 3차원 동작을 2차원으로 투영하는 과정에서 발생하는 왜곡 때문에 관절 각도 계산에 오차가 발생한다. 이러한 한계로 판정 임계값을 완화해야 했으며, 이는 시스템의 정밀도를 제한하는 요인으로 작용한다.

참 고 문 헌

- [1] Ultralytics YOLOv8. (n.d.). Ultralytics.
<https://docs.ultralytics.com/ko/models/yolov8/>
- [2] A. Toshev and C. Szegedy, “DeepPose: Human pose estimation via deep neural networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1653-1660, 2014.
- [3] Mittal, P. Mittal, “A comprehensive survey of deep learning-based lightweight object detection models for edge devices,” Artificial Intelligence Review, vol. 57, article no. 242, Aug. 2024.
- [4] 진경은, 신은호, 전윤호. *Raspberry Pi 환경에서의 포즈 추정 모델을 활용한 박수 게임*. 2025년도 한국통신학회 동계종합학술발표회 논문집, 한국통신학회, 2025, 쪽 1344-1345.