

<http://dx.doi.org/10.17703/JCCT.2021.7.3.577>

JCCT 2021-8-69

딥러닝 기반 영상처리 기법 및 표준 운동 프로그램을 활용한 비대면 온라인 홈트레이닝 어플리케이션 연구

Non-face-to-face online home training application study using deep learning-based image processing technique and standard exercise program

신윤지*, 이현주**, 김준희**, 권다영**, 이선애**,
추윤진**, 박지혜**, 정자현**, 이형석**, 김준호***

Youn-ji Shin*, **Hyun-ju Lee****, **Jun-hee Kim****, **Da-young Kwon****, **Seon-ae Lee****,
Yun-jin Choo**, **Ji-hye Park****, **Ja-hyun Jung*****, **Hyoung-suk Lee*****, **Joon-ho Kim******

요약 최근 AR, VR 및 스마트 디바이스 기술의 발전에 따라 피트니스 산업에서도 비대면 환경을 기반으로 한 서비스 수요가 증가하고 있다. 비대면 온라인 홈트레이닝 서비스는 기존의 오프라인 서비스에 비해 시간과 장소의 제약이 없다는 장점이 있으나 운동 기구의 부재 및 사용자의 정확한 운동 자세 유지여부, 운동량의 측정이 어려운 단점이 존재한다. 본 연구에서는 이러한 단점을 보완할 수 있는 표준 운동 프로그램을 개발하고 딥러닝 기반 신체 자세 추정 영상처리를 통하여 새로운 비대면 홈트레이닝 어플리케이션 알고리즘을 제안한다. 본 연구의 알고리즘 기반 어플리케이션을 활용한다면 표준 운동 프로그램 영상의 트레이너를 사용자가 직접 보고 따라하면서 사용자 스스로 자세를 교정하며 정확한 운동이 가능하다. 나아가 본 연구의 알고리즘을 용도에 맞게 커스터마이징 한다면 공연, 영화, 동아리 활동, 컨퍼런스 분야로의 적용도 가능할 것이다.

주요어 : 딥러닝, 물체인식, 이미지 처리, 분류, 컴퓨터비전, 자세 추정

Abstract Recently, with the development of AR, VR, and smart device technologies, the demand for services based on non-face-to-face environments is also increasing in the fitness industry. The non-face-to-face online home training service has the advantage of not being limited by time and place compared to the existing offline service. However, there are disadvantages including the absence of exercise equipment, difficulty in measuring the amount of exercise and checking whether the user maintains an accurate exercise posture or not. In this study, we develop a standard exercise program that can compensate for these shortcomings and propose a new non-face-to-face home training application by using a deep learning-based body posture estimation image processing algorithm. This application allows the user to directly watch and follow the trainer of the standard exercise program video, correct the user's own posture, and perform an accurate exercise. Furthermore, if the results of this study are customized according to their purpose, it will be possible to apply them to performances, films, club activities, and conferences

Key words : Deep-learning, Object-detection, Image processing, Classification, Computer vision, Pose Estimation

*정회원, 한양대학교 일반대학원 (제1저자)

**정회원, 동서대학교 레저스포츠학과 (참여저자)

***정회원, (주)제이콥시스템 (참여저자)

****정회원, 동서대학교 전기정보제어학과 (교신저자)

접수일: 2021년 7월 30일, 수정완료일: 2021년 8월 5일

게재확정일: 2021년 8월 8일

Received: July 30, 2021 / Revised: August 5, 2021

Accepted: August 8, 2021

*Corresponding Author: heglerkim@du.ac.kr

Dept. of Electrical Information Control, Dong Seoul Univ, Korea

I. 서론

세계적으로 COVID-19 집단 감염 방지를 위한 사회적 거리두기로 피트니스 산업에서도 언택트 소비가 확산되고 있다. 이는 정해진 장소와 시간의 제약이 없다는 장점까지 더하여 비대면 디지털 피트니스 서비스에 대한 관심이 폭발적으로 증가하고 있다. 그러나 스트리밍 영상을 통해 사용자가 운동을 단순히 따라하는 비대면 디지털 피트니스 서비스는 직접 코칭을 받을 수 있는 오프라인 피트니스에 비해 운동을 따라하는 방식에 있어서 사용자의 정확한 자세 여부와 운동량의 측정이 어렵다는 단점이 존재한다. 또한 이러한 치명적인 단점은 사용자에게 운동에 대한 동기부여의 부족함으로 작용하여 운동의 효과를 충족시키기 위한 지속성을 갖기 힘들다. 최근 비대면 시대를 맞이하여 메타버스(Metaverse)가 전세계 트렌드로 주목받고 있으며 홈트레이닝 분야 또한 메타버스 시대에 맞춰 트레이닝 환경을 구축해야 한다.

따라서 본 연구는 COVID-19로 인한 급격한 환경 변화에 신속하게 대응할 수 있는 새로운 비대면 홈트레이닝 환경을 제안할 뿐만 아니라 체계적인 운동프로그램 개발과 운동 시 소모되는 칼로리 계산 및 평가시스템 개발, 그리고 4차 산업혁명에 떠오르는 AI 알고리즘을 적용하여 표준모델과 사용자의 운동 자세를 비교하는 맞춤형 운동 프로그램을 제안하고자 한다. 표준 운동 프로그램 개발 및 표준 운동 프로그램에 AI 알고리즘을 적용하여 운동 표준 모델을 따라하는 사용자의 운동 상태를 실시간 체크함으로써 사용자의 정확한 자세와 운동 시 칼로리 소모량을 정확하게 예측가능하다.

II. 본론

2.1 표준 운동 프로그램 개발

운동프로그램을 개발함에 있어서 운동프로그램을 수행하는 참여자에게 정확한 자세를 알려주기 위해서는 각 운동에 맞는 표준 운동 모델이 필요하다.

첫 번째로, 표준 운동 모델을 구축하기 위해 학생들과 모션 캡처를 통해 정확한 자세를 측정하고, 이를 동영상 파일로 추출하여 지속적인 반복을 통해 운동효과를 검증하여 5종목 58가지의 운동프로그램을 개발하였다. 모션 캡처를 통한 표준 운동 모델 구축 과정은 그림 1과 같다.



그림 1 모션캡처를 통한 표준 운동 모델 구축
 Figure 1 Establishment of standard motion model through motion capture

두 번째로, 각 운동 별로 정확한 칼로리 소모량 계산이 필요한데 본 연구에서는 METs(Metabolism) 단위를 활용하였다. METs는 운동강도를 나타내는 표시법의 하나로, 안정상태를 유지하는데 필요한 산소량의 1단위이다. 운동생리학이나 운동처방에서는 1METs를 기준으로 각 활동이 안정시보다 산소요구량이 몇 배가 많은지 나타내는 지표로 활용하며 칼로리 계산에도 요긴하게 사용되고 있다. 예를 들어 80kg의 사용자가 METs 3.0 강도로 운동을 했다고 가정하면 $3.0 \text{ METs} = 3.0 \text{ ml} \times (3.5 \times 80 \times 60) = 50,400 \text{ ml}$, $50,400 \text{ ml} \times 5 \text{ kcal} = 252 \text{ kcal}$ 을 소모하게 되는 것이다. 표 1은 대표적인 운동종목의 METs 별 분류표이다.

표 1. 대표적인 운동종목의 METs 별 분류
 Table 1. Classification of representative sports by METs

METs	종목
3.0	볼링, 윈드서핑, 필라테스, 요가
3.5	체조, 계단오르내리기, 골프, 일반적인 걷기
4.0	자전거타기(천천히), 탁구, 배구, 출퇴근길 도보, 아쿠아로빅
5.0	야구, 스노클링, 사냥
5.5	헬스클럽(일반적), 사교댄스, 승마
6.0	저항운동, 조깅과 걷기 조합, 테니스(복식), 하이킹, 스키, 수영, 복싱(샌드백)
7.0	배드민턴시합, 농구, 라켓볼, 스케이팅, 축구, 스쿠버다이빙, 조깅, 스키, 테니스(일반)
8.0	등산, 달리기(8km/h), 미식축구, 암벽등반, 테니스(단식), 아이스하키
9.0	복싱(스파링)
10.0	줄넘기(보통), 장거리달리기, 킥복싱
12.0	줄넘기(빠르게), 스쿼시, 핸드볼, 자전거타기(30km/h)
15.0	계단달리기
16.0	달리기(16km/h)

본 연구에서 개발한 운동프로그램은 METs 3.0~5.9 사이, 즉 중간 정도의 활동을 기준으로 개발하였다. 단, 운동자의 신체적 특성(근육이 많은 타입, 지방이 많은

타입 등), 운동 경험도, 시간(운동 중 휴식시간 등) 등 변수요인이 있으므로 단순 비교하는 것은 무리가 있을을 인지하고 있어야 한다.

2.2 운동자세 인식을 위한 딥러닝 알고리즘

비대면 환경에서 정확한 운동자세 여부를 판별하기 위하여 웹캠 또는 모바일 카메라로 촬영한 사용자의 영상정보의 분석이 필요하다. 본 연구에서는 MoveNet이라는 딥러닝 기반의 Pose-Estimation 알고리즘을 활용하여 영상 내 사용자의 신체 17개 부위를 키포인트로 감지하고 스켈레톤 형태로 추론해낸다.

MoveNet은 신체 17개의 키포인트를 감지하는 매우 빠르고 정확한 모델이다. 이 알고리즘은 Thunder, Lightning로 알려진 두 가지 모델로 Thunder는 높은 정확도를 위한 모델이고, Lightning은 빠른 속도처리를 위한 모델이다. 두 모델은 실시간(30FPS이상처리를 위해 개발되었으며, 라이브 피트니스, 건강 및 웰빙등의 분야에도 응용가능하다 판단되어 딥러닝 알고리즘 중 MoveNet을 활용하였다.

또한 이 알고리즘은 영상의 이전 프레임의 감지를 기반으로 지능형 객체 선택과 자르기로 영상처리 영역을 감소시켜 효율적인 컴퓨팅 리소스 사용은 물론 빠른 속도로 더 정확한 품질을 예측할 수 있었다.

MoveNet은 히트맵을 이용하여 인간의 신체부위에 키포인트를 정확하게 bottom-up 추정 모델로 전체적인 아키텍처는 두가지 구성요소로 Feature extractor와 Prediction heads로 구성되어 있다. 예측하는 방식은 One-stage Detector 방식의 CenterNet 알고리즘 기반으로 속도와 정확도를 향상시켰다.

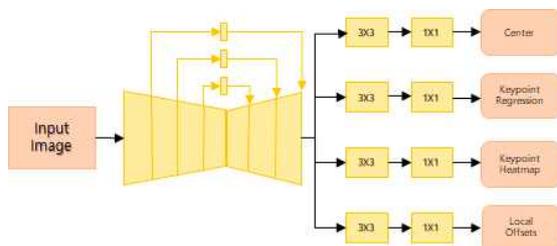


그림 2. MoveNet 아키텍처
 Figure 2. MoveNet architecture

또한 사람의 스켈레톤 모델 추정은 물론 운동이라는 특정 행동 인식의 정확성을 극대화하고자 Training Dataset은 COCO 데이터 세트와 Google 내부 데이터 세트인

Active 데이터 세트 총 두 가지의 데이터 세트를 사용하여 학습하였다. COCO 데이터 세트는 피트니스 및 댄스에 관한 어플리케이션이 부족하지만 다양한 장면 및 대용량 데이터라는 강점을 가지고 있으며, Active 데이터 세트는 COCO와 같은 표준 17개의 신체 키포인트를 가짐과 동시에 YouTube의 요가나 피트니스 등 다소 까다로운 특정 행동 데이터로 제작하였기 때문에 COCO의 부족한 점을 보완하기 위해 활용하였다.

2.3 표준 운동 프로그램 영상의 데이터 전처리

딥러닝 기반 영상처리 과정은 한정된 시간 내 많은 컴퓨팅 리소스를 요구하기에 사용자의 운동 영상만 실시간으로 처리하는 것을 목표로 하고자 했다. 따라서 표준 운동 프로그램을 기반으로 정확한 자세로 운동하는 트레이너의 영상과 사용자의 운동 영상에서 트레이너의 신체 부위별 좌표를 획득하고 이를 표준 운동 좌표 데이터로 활용하기 위한 데이터 전처리 과정이 필요하다. 이러한 데이터 전처리 과정의 알고리즘 순서도는 그림 3과 같다.

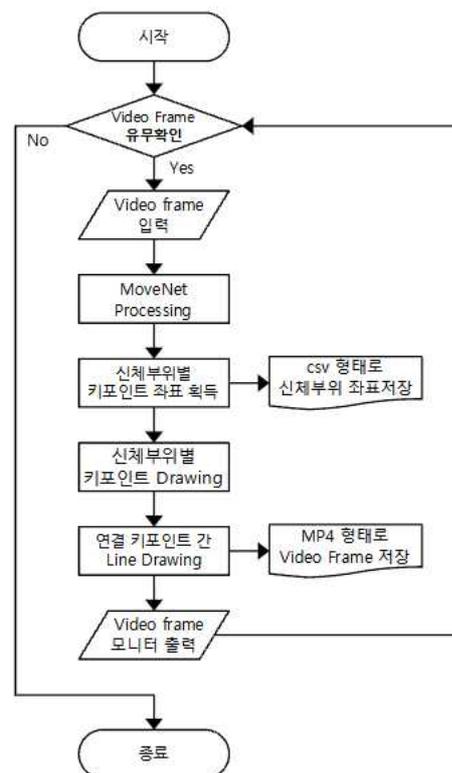


그림 3 전처리 알고리즘 순서도
 Figure 3 Preprocessing Algorithm Flowchart

표준 운동 프로그램 영상의 각 프레임을 MoveNet 알고리즘에 연속적으로 입력하여 트레이너의 17개 신체 부위의 좌표를 csv 파일 형식으로 저장한다. 그리고 각 신체 부위 키포인트를 연결하는 라인 드로잉으로 한 눈에 쉽게 알아볼 수 있도록 이미지화 한다. 마지막으로 키포인트 및 라인 드로잉을 완료한 이미지들을 MP4 파일 형식의 별도의 비디오 영상으로 저장하고 화면에 출력하도록 설계하였다. 표준 운동프로그램 영상에 키포인트와 라인 드로잉을 마친 이미지는 그림 4와 같다.



그림 4. 각 신체 부위별 키포인트, 라인드로잉 완료 영상
 Figure 4. Key point and line drawing completion video for each body part

2.4 사용자 운동 영상처리 및 운동자세 평가

사용자의 운동 영상은 카메라 앞에서 사용자가 표준 운동 프로그램의 영상을 보며 운동을 따라하는 과정이 실시간으로 출력되는 영상이다. 따라서 사용자 운동 영상처리는 표준 운동 프로그램 영상 데이터 전처리 과정과 동일하게 진행됨과 동시에 사용자의 운동 자세 정확도 판별과정이 추가되어야 한다.

사용자 운동 영상도 이전과 동일하게 각 프레임 별로 MoveNet 알고리즘에 입력하여 영상 내 사용자의 각 신체 부위 좌표 추정과 키포인트 및 라인 드로잉을 완료한다. 그리고 표준 운동 프로그램 영상 데이터 전처리 알고리즘에서 획득한 비디오 영상 프레임과 각 프레임 신체 부위별 좌표 데이터를 참조하여 참가자의 운동 타이밍과 정확한 자세를 판별하도록 한다.

판별과정에서 정확한 운동자세 여부는 사용자의 운동 영상에서 신체 부위 별 키포인트 좌표가 표준 운동 프로그램의 트레이너 신체 부위 별 키포인트 좌표의 반경 3px(pixel) 이내 여부로 설계하였다. 3px 오차범위 이내 인 경우는 정확한 운동자세로 판단하여 GOOD 메시지를 출력하고, 3px 이상 차이가 날 경우에는 부정확한 운동자세로 판단하여 BAD 메시지를 출력한다.

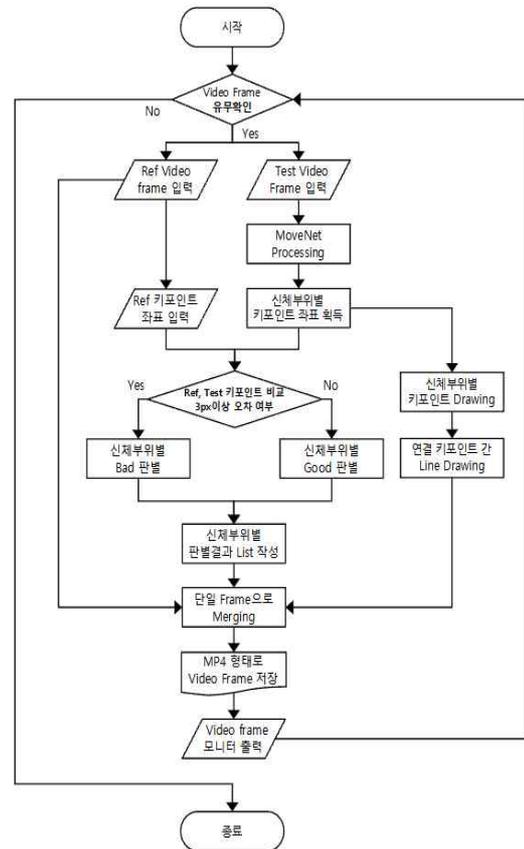


그림 5 참가자의 정확성 판별 알고리즘 순서도
 Figure 5 Participant Accuracy Determination Algorithm Flowchart

마지막으로 사용자 스스로 정확한 운동 여부를 확인하며 부정확한 부분이 어딘지 개선하기 위한 서비스로 표준 운동 프로그램의 트레이너 영상, 사용자의 운동 영상, 신체 부위별 정확한 운동 상태 판별 결과를 동시에 MP4 파일 형식으로 저장 및 모니터 출력을 하도록 설계하였다.

III. 실험 및 결과

설계한 알고리즘을 토대로 표준 운동 프로그램 영상의 데이터 전처리를 위한 용도와 실시간 사용자 운동 영상 분석용으로 두 가지 어플리케이션을 완성하였다.

표준 운동 프로그램 영상 데이터 전처리를 위한 어플리케이션은 본 연구에서 개발한 표준 운동 프로그램 5종목 58가지 운동 모두 좌표 데이터의 획득여부를 확인하였다. 실험결과 모든 운동에서 비교적 정확하게 표준 운동 프로그램 영상에서의 트레이너 신체부위 17개소

좌표 데이터를 얻을 수 있었다.

사용자 운동 영상 분석용 어플리케이션은 표준 운동 프로그램 영상과 사용자의 운동 영상 모두를 출력함과 동시에 운동 중 사용자의 각 신체 부위별 위치의 정확성 판별 여부를 메시지로 출력하도록 구성하였다. 실험은 최대한 동일한 자세로 운동을 따라한 경우와 실제로 트레이너의 운동을 따라하면서 실수하거나 타이밍이 맞지 않는 경우 총 2가지로 진행하였다.

첫 번째 실험결과로 최대한 동일한 자세로 운동을 따라한 경우, 각 신체 부위별 운동 위치나 타이밍이 정확하다고 할 수 있는 GOOD 메시지가 출력되었으며, 이는 그림 6에서 확인할 수 있다.



그림 6. 운동 자세나 타이밍이 최대한 동일한 경우
 Figure 6. When the exercise posture or timing is the same as possible

두 번째로 유사한 자세로 트레이너를 따라 운동은 하고 있지만 실제로 발생할 수 있는 실수를 하거나 운동 타이밍이 틀린 경우의 실험결과를 신체 부위별로 위치가 정확한 경우 GOOD 메시지를 출력하고 부정확한 신체 부위는 BAD 메시지를 출력하였다. 신체 부위 별 정확성 판별 여부는 그림 7에서 확인할 수 있다. 출력 신체 부위별로 정확한 위치 여부 메시지를 출력함으로써 최대한 정확한 운동을 유도할 수 있었다.



그림 7. 실수로 인하여 운동 자세나 타이밍이 다른 경우
 Figure 7. If the exercise posture or timing is different due to mistake

IV. 결 론

본 연구에서는 장기간 지속되고 있는 COVID-19로 인한 오프라인 피트니스 시장의 한계를 극복할 수 있는 비대면 디지털 피트니스 서비스의 기본 모델을 제안하

였다. 표준 운동 프로그램 개발로 정확한 칼로리 소모량 계산이 가능하며, 딥러닝 기반 운동자세 추정 알고리즘을 활용하여 사용자의 정확한 운동자세 여부 판별이 가능하다. 이로써 사용자의 정확한 운동량 분석으로 운동에 대한 동기부여를 제공하고 지속적인 운동을 가능하게 하며 시간과 장소의 한계를 극복하는 메타버스 시대에 적합한 피트니스 환경을 구축했다 할 수 있다.

또한 표준 운동 모델과 사용자의 운동 자세를 정확하게 비교하기에 홈 트레이닝 피트니스 분야 외에도 공연, 영화, 동아리 활동, 컨퍼런스 등 비대면 관련 활동에 활용할 수 있을 것이라 기대한다.

References

- [1] CAO, Zhe, et al. OpenPose: realltime multi-person 2D pose estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence*, 2019, 43.1: 172-186.
- [2] IM, Jin-mo, et al. Mixed reality health management model using smart phone. *The Journal of the Convergence on Culture Technology*, 2018, 4.2: 185-189.
- [3] LEE, Soowook. Deep structured learning: architectures and applications. *International Journal of Advanced Culture Technology*, 2018, 6.4: 262-265.
- [4] LIN, Tsung-Yi, et al. Microsoft coco: Common objects in context. In: *European conference on computer vision*. Springer, Cham, 2014. p. 740-755.
- [5] ZHOU, Xingyi; WANG, Dequan; KRÄHENBÜHL, Philipp. Objects as points. *arXiv preprint arXiv:1904.07850*, 2019.
- [6] LI, Ang, et al. The ava-kinetics localized human actions video dataset. *arXiv preprint arXiv:2005.00214*, 2020.
- [7] HUANG, Jonathan, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 7310-7311.
- [8] LIN, Tsung-Yi, et al. Feature pyramid networks for object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 2117-2125.
- [9] HOWARD, Andrew G., et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

※ 본 논문은 '콘텐츠진흥원 2020년 문화기술연구
개발' 사업을 지원받아 진행된 연구결과입니다.
사업명 : 비대면 휘트니스 에듀테인먼트 서비스
를 위한 5G/MR 기반 플랫폼 및 센싱기술 개발