

3차원 자세 추정 및 레이 캐스팅을 활용한 건설현장 보호장비 미착용 탐지 알고리즘

A PPE Non-Compliance Detection Algorithm for Construction Site Using 3D Pose Estimation and Ray-casting

○전 채 희*
Jeon, Chae-Hee

김 도 현**
Kim, Dohyun

이 상 훈***
Lee, Sanghoon

Abstract

The construction industry poses a high risk of accidents. To ensure effective safety management, it is necessary to establish a three-dimensional system that integrates contextual information and spatial data across various stakeholders. This study proposes a helmet-wearing detection algorithm based on 3D human pose estimation. The proposed system combines 2D object detection, 3D pose estimation, and ray-casting to determine helmet-wearing status. A key distinction of this approach lies in its integration of 2D equipment detection and 3D pose estimation into a unified pipeline. The system was evaluated through two scenarios, and future work will extend it with BIM-based spatial data for an integrated 3D safety management system.

키워드 : 건설 안전, 개인보호구, 인공지능, 3차원 자세 추정, 레이캐스팅, 자동 관제

Keywords : Construction Safety, PPE, Artificial Intelligence (AI), 3D Pose Estimation, Ray casting, Automated Monitoring

1. 서론

1.1 연구의 배경 및 필요성

건설산업은 전체 업종 중 가장 높은 산업 재해율을 보이는 업종으로, 사망사고 수에는 매년 소폭의 감소가 있었으나 전체적으로는 큰 변화 없이 유사한 수준에 머물고 있다. 주요 재해 유형으로는 전도, 추락, 낙하·비래 등이 있다. 국토교통부에 따르면, 지난 5년(2020~2024년)간 건설 재해 사망자는 총 1211명, 부상자는 3만340명에 달한다. 사망사고 1211건 중 622건(51%)이 추락으로 인해 발생하였으며, 부상 사고 3만340건 중 7109건(23%)이 전도로 인해 발생하였다.

건설 현장의 재해사고 방지를 위해서는 추락 방지와 가시설 및 장비 관련 안전 대책을 마련하는 것이 중요하며, 작업자의 단순 과실이 중대재해로 이어지지 않도록 체계적인 건설 안전관리 시스템이 요구된다. 이를 위하여 최근 현장에서는 IoT 센서를 활용한 스마트 안전 장비와 AI 기반의 스마트 관제시스템이 활발히 도입되고 있으며, 특히

CCTV 영상 기반 장비 인식 및 작업자 출입 통제를 위한 AI 영상분석 기술들이 사용되고 있다.

인력 중심의 복잡한 건설 현장은 정적인 공간과 동적인 주체들의 움직임과 다양한 행동 패턴 간의 상호작용이 무수히 많이 존재한다. 그러나 기존 영상 분석 시스템들은 장비 탐지는 2차원, 작업자 자세 추정은 3차원으로 분리된 시스템에 의존하고 있다. 이는 작업자의 다양한 자세나 공종별 작업 맥락을 안전 장비 착용 여부와 종합적으로 이해하여 재해로 이어지는 상황을 방지하기에는 한계가 있다. 건설 현장의 더욱 정밀한 자동 관제 시스템 구축을 위해서는 3차원의 정적, 동적 객체들의 통합적 모니터링 뿐만 아니라 동적 객체들에 대한 고차원적 정보의 반영이 필요하다. 따라서 건설 현장의 다양한 객체 간의 맥락적인 정보와 공간 정보를 통합하여 관리할 수 있는 안전관리 시스템을 제안하고자 한다.

1.2 연구의 목적 및 범위

본 연구의 목적은 건설 현장을 공간적 맥락성을 강화하여 현장의 다중 개체를 통합적으로 모니터링하고, 체계적인 안전 계획을 수립·실행할 수 있는 AI 기반 3차원 건설 안전 시스템 구축의 가능성을 제시하는 데 있다. 연구 범위는 작업자의 건설 안전 장비 중 안전모를 대상으로 한정하였으며, 안전모 미착용 여부를 실시간으로 판별하는 3차원 자세 추정 기반 건설 현장 작업자 안전 장비 추적 알고리즘을 개발하였다.

* 서울시립대학교 도시과학대학 건축공학전공 학사과정

** 서울시립대학교 대학원 건축공학전공 석박사통합과정 수료

*** 서울시립대학교 도시과학대학 건축학부 교수, 공학박사

(Corresponding author : Department of Architectural Engineering, University of Seoul, sanghoon.lee@uos.ac.kr)

이 연구는 2025년도 (주)건축사사무소건원엔지니어링 연구비 지원에 의한 결과의 일부임. 과제번호:202507302002.

이를 위해 3차원 다중 인원 자세 추정 AI 모델을 활용하여 현장 작업자들의 자세를 3차원 공간상에 표현하였다. 특히 본 연구는 레이 캐스팅 기법을 통해 3차원 공간 내에 안전 규칙과 위험 임계치를 직접 적용하고, 이를 객체 검출 모델 및 헵가리안 알고리즘과 결합하여 개별 작업자의 안전모 착용 여부를 판별하는 시스템을 제안한다.

2. 선행연구 고찰

2.1 선행연구 분석

AI 기반 실시간 탐지 기술의 발전으로 작업자의 보호장비 착용 여부 탐지, 3차원 자세 인식, 협착 위험 구역 평가 등 다양한 건설 현장 안전사고 예방에 대한 연구가 활발히 진행되고 있다.

대표적으로는 2차원 영상을 CNN 기반 딥러닝 모델이 실시간 분석하여 작업자 안전 장비를 탐지하는 자동 관제 시스템 연구가 진행되었다. Lo et al.(2022)은 YOLO 계열 모델을 적용하여 안전모·조끼 실시간 탐지 시스템을 개발하였다. Zhang et al.(2025)은 YOLOv9, YoloPose, Strong SORT를 활용해 작업자 안전모 착용 상태 추적 알고리즘을 제안했으며, 40프레임 미만의 가림·누락 문제를 해결해 약 98% 이상의 정확도를 보였다.

작업자나 장비의 3차원 자세를 추정해 현장의 단일 카메라 영상으로 작업자 골격 안전, 협착 방지를 인식하는 연구 또한 다수 수행되었다. Chen et al.(2023)은 3차원 자세 추정 모델과 추적을 활용하여 작업자의 실시간 자세들을 위험도 지수로 환산하는 근골격계 위험 평가 알고리즘을 제안하였다. Wen et al.(2022)은 R-CNN 기반 딥러닝 모델을 적용하여 굴착기의 3차원 자세를 추정하여 장비 충돌 탐지 알고리즘을 제안하였다.

최근에는 3차원 기반 건설 현장의 다중 객체들에 대한 통합 안전관리를 실행하는 시스템에 관한 연구들이 제안되었다. Yan et al.(2025)는 정적요소(건축 구조물, 자재 등)와 동적 요소(작업자, 건설장비 등)의 실시간 상태를 탐지하여 3차원 가상 공간에 나타내는 모니터링 시스템을 개발하였다.

2.2 기존 연구의 한계점 및 본 연구의 차별성

딥러닝 AI 모델 기반 2차원 영상 분석 기술은 다중 인원을 탐지하는 상황이거나, 작업자라는 동적 객체의 자세, 각도 변화에 따라 가림·누락이 발생하여 탐지 성능이 저하되는 한계가 존재한다.

3차원의 BIM 기반 다중 객체 모니터링 시스템은 정적·동적 객체들의 통합 정보를 공간정보와 함께 결합하여 모니터링 하였다는 점에서 기여한 바가 크다. 하지만, 3차원에 표현되는 동적 객체들의 모델 형태와 움직임은 단순화된 형태로만 구현되어, 인간의 골격과 비정형적 움직임 패턴 등 정밀한 정보까지 활용할 수 없다는 한계가 있다.

본 연구는 이와 같은 한계를 극복하고자 3차원 자세 추정을 기반으로 레이 캐스팅 알고리즘을 활용하여, 2차원 상에서 해결하였던 장비 착용의 안전 규칙을 3차원 상에서 동일하게 구현하는 알고리즘을 제안한다. 기존 연구에서는 바운딩 박스 사이 IoU를 헬멧 착용 여부 판단에 활용하였으나, 본 연구는 작업자의 머리와 헬멧 중심점 사이의 3차원상 절대거리를 기반으로 착용 여부를 판단하였다. 이는 작업자의 3차원 자세 정보와 깊이 정보까지 안전 장비 착용 여부 판단 규칙에 활용하여 정밀성을 높였다는 차별성을 가진다. 깊이 정보의 활용은 특히 다중 인물 탐지 상황에서 탐지 정확도를 향상시키고, BIM과 같은 공간정보와의 통합 가능성을 제시한다.

3. 연구 방법

3.1 시스템 개요

본 연구에서 제안하는 시스템의 워크플로우는 그림1과 같다. 시스템은 크게 세 단계로 구성된다. 먼저 YOLOv11 모델을 사용하여 사람 및 안전모를 검출한다. 다음으로 Moon et al.(2019)의 3차원 자세 추정 모델 (RootNet·PoseNet)을 적용하여 다중 인원의 3차원 깊이 정보와 골격을 추정한다. 마지막으로 레이 캐스팅과 헵가리안 알고리즘을 결합하여 각 인물들의 헬멧 착용 여부를 중복, 교차 배경 없이 판별한다.

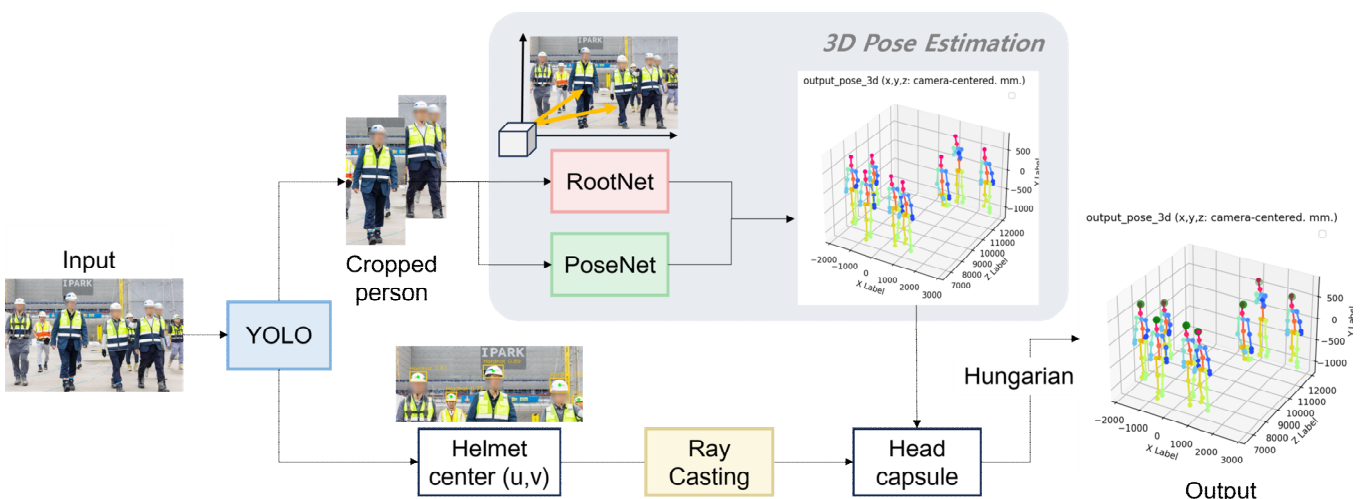


그림1. 시스템 워크플로우

3.2 시스템 개발

3.2.1 2차원 객체 검출 모듈

시스템의 첫 단계에서는 YOLO를 활용해 사람·안전모 탐지를 수행한다. 본 연구에서는 건설 현장 PPE 탐지가 가능하도록 하기 위해 공개된 Construction-Site-Safety-PPE-Detection 데이터셋(Snehl, 2022)을 활용하여 YOLOv11 모델을 파인튜닝 하였다. YOLO에 의해 검출된 각 사람의 바운딩 박스는 개별적으로 추출된 후 3차원 자세 추정 모델의 입력값으로 사용된다. 헬멧 바운딩 박스의 중심점은 그림2와 같이 (u,v)로 정의되어 헬멧 미착용 탐지 알고리즘에 활용된다.



그림2. 바운딩 박스와 헬멧의 중심점

3.2.2 3차원 자세 추정 모듈

3차원 자세 추정 단계는 인체 관절의 3차원 좌표와 개별 사람의 이미지를 입력으로 하여 3차원 자세 추정 모델 RootNet과 PoseNet을 활용해 구현하였다. RootNet은 카메라 중심 좌표계로부터 각 사람 골반 중심점의 2차원 좌표(X, Y)와 절대깊이 좌표(Z)값을 추정한다. PoseNet은 잘라낸 사람 이미지로부터 3차원의 상대 깊이 값을 가진 관절 좌표를 추정한다. 3차원 자세를 추정하는 PoseNet은 한 사람의 관절 좌표 간의 상대 깊이만 추출하기 때문에, 다중 인원을 탐지할 경우에는 각 사람의 카메라로부터 절대 깊이를 추정하는 RootNet이 필요하다. 각각의 모델이 추정한 절대 깊이, 관절 상대 깊이 좌표를 통합하여 다중 인원의 3차원 자세를 추정하게 된다.

3.2.3 레이 캐스팅을 활용한 헬멧 미착용 탐지 모듈

시스템의 마지막 단계는 레이 캐스팅 기반 매칭을 통해 작업자의 헬멧 착용 여부를 판정한다. 본 연구에서는 헬멧의 착용 여부를 3차원적으로 판단하기 위해, 사람 머리의 부피를 근사하는 헤드 캡슐과 각 헬멧의 중심을 향하는 레이(ray)의 충돌/근접 관계를 사용하였다.

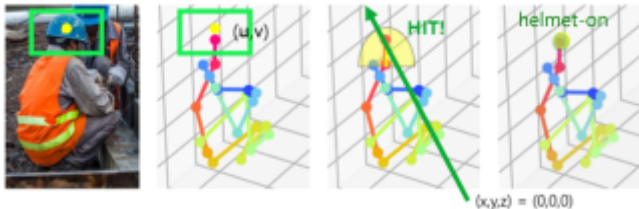


그림3. 레이 캐스팅을 활용한 헬멧 미착용 여부 탐지

그림3과 같이 카메라 좌표계의 원점 (0,0,0) 에서 YOLO로 얻은 헬멧 중심 (u,v)를 통과하는 방향으로 레이를 생성한다. 동시에 PoseNet의 3차원 자세 추정에서 목-머리 조인트를 잇는 선분을 축으로 반지름 $r=50\text{mm}$ 을 갖는 헤

드 캡슐을 생성한다. 레이와 캡슐 축의 최근접 거리 $\min_dist$ 를 계산하여, $\min_dist \leq r$ 이면 헬멧의 중심을 향해 진행하는 레이와 머리의 중심축과 충분히 가깝게 지나므로 착용(hit)으로 간주한다. 3차원 기반 헬멧 판정은 3차원 공간에서 서로 다른 절대 깊이를 가진 다중 인원에게 대해서도 원근 정보를 충실히 반영할 수 있다. 이는 기존의 2차원 바운딩 박스 IoU 기반 매칭의 한계점을 보완할 수 있다.

3.2.4 헝가리안 알고리즘의 일대일 매칭

헬멧을 향해 진행하는 레이의 시작점이 원점이므로, 모든 헬멧 레이와 깊이가 작은(카메라에 가까운) 사람에게 우선으로 배정되어 헬멧이 특정 인원들에게 중복 배정되는 문제가 발생하였다. 이를 해결하기 위해 모든 헬멧-사람 쌍에 대한 비용행렬(거리·깊이 차)을 구성하고 헝가리안 최소합 할당을 적용함으로써 헬멧의 중복 및 오배정을 방지하였다.

4. 성능 평가 및 결과 분석

본 연구에서는 실제 건설 현장의 데이터를 사용하여 다음의 두 가지 시나리오 실험을 통해 제안된 시스템의 성능을 평가하였다. 첫 번째로 다수의 인원이 혼재한 상황에서 사람 간 가림에 대하여 헬멧 판별 성능을 검증하기 위하여 실험을 진행하였다. 실험 결과(그림 4), 다수 인원의 바운딩 박스가 중첩되고 밀집된 환경에서도 개별 인원의 깊이가 추정이 적절히 이루어졌으며, 헬멧 탐지와 인원별 매칭이 효과적으로 수행되었다.

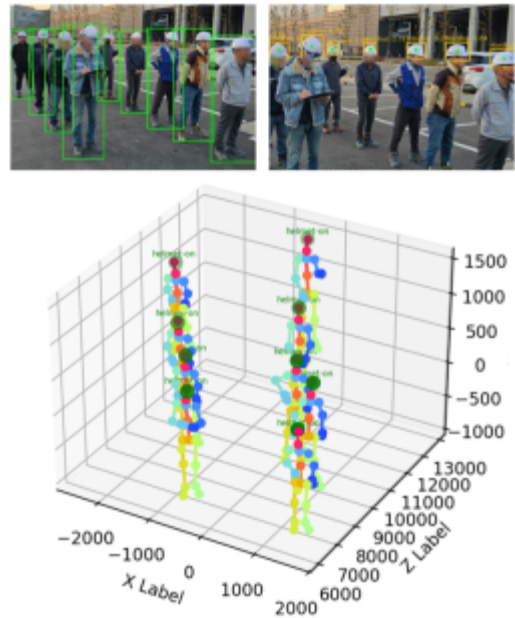


그림4. 다중 인원 혼재 시나리오의 결과값

두 번째는 자체에 의한 가려짐 상황이며, 작업자의 신체가 자체에 의해 일부 가려지거나, 인원 간 바운딩 박스가 겹쳐 있는 시나리오로 실험을 진행하였다. 실험 결과(그림 5), 신체가 일부 가려진 상황에서도 3차원 자세 추

정이 성공적으로 이루어졌으며, 헬멧 착용 판단이 머리-목 골격 정보에 기반하므로 가려짐에 의한 성능 저하 없이 안정적인 결과를 얻었다.

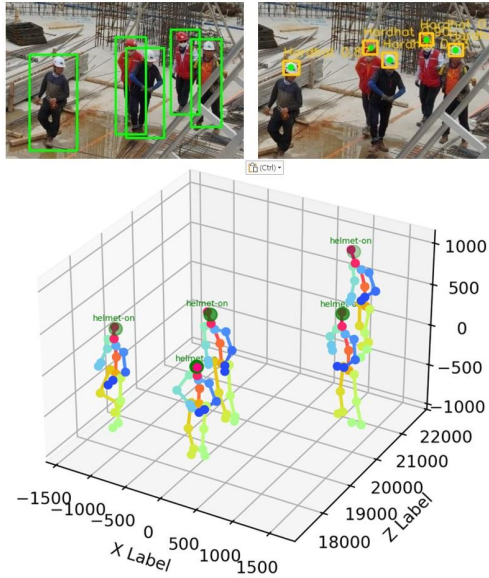


그림5. 자재에 의한 가려짐 시나리오의 결과값

제안된 시스템은 다중 인원 밀집 환경과 부분적으로 대상 객체가 가려진 상황에서도 안정적으로 3차원 자세를 추정하고 헬멧 착용 여부를 판별하여, 기존 2차원 기반 시스템의 한계를 보완한다. 2차원 이미지상 육안으로 확인되는 깊이와 3차원 공간상에서 추정된 깊이 값 사이에 미세한 오차가 관찰되었으나, 헬멧 착용 판별의 핵심 성능에는 영향을 주지 않아 실무적 적용 가능성을 확인하였다.

5. 결론

본 연구에서는 건설 현장의 다중 객체 정보를 통합한 3차원 기반 건설 안전관리 시스템의 가능성을 검토하고, 그 첫 사례로서 작업자의 헬멧 착용 여부를 3차원 공간상에서 탐지하는 방법론을 제안하였다.

제안된 시스템은 다음의 세 단계로 구성된다: (1) YOLO를 이용한 사람과 헬멧의 검출. (2) 검출된 각 사람 영역을 기반으로 RootNet · PoseNet을 활용한 절대 깊이 좌표 및 3차원 관절 좌표를 추정. (3) 3차원 골격 기반으로 캡슐화된 머리 영역과 헬멧의 위치를 레이 캐스팅, 헝가리안 알고리즘으로 매칭. 실험 결과 본 시스템의 신뢰성있는 헬멧 착용 여부 탐지 성능을 확인하였다.

본 연구는 2차원 객체 탐지, 3차원 자세 추정, 레이 캐스팅 기법을 결합하여 안전 장비 착용 여부 판단 근거를 깊이와 자세 정보까지 확장하였다. 3차원 공간 정보에 기반한 다중 객체들의 통합 안전관리 시스템의 필요성과 가능성을 확인하였다는 기여점을 가진다. 하지만, 단안 카메라를 기반으로 3차원 공간을 추정하기 때문에, 절대 거리 단위를 고정하지 못하면 사람-카메라 간 거리 변화 등의 요인들에 의해 시간에 따라 추정된 스케일이 변화하는 스케일 드리프트 현상이 우려된다. 스케일의 누적 오차는 안

정적인 절대 거리 규칙 적용에 있어 한계점을 가진다.

향후 연구로는, 객체 추적 기술과의 접목을 통해 실시간 영상에서도 안정적으로 3차원 자세를 추정할 수 있도록 개선하고자 한다. 또한, 건설 중장비의 3차원 골격을 추출하고, 작업자와 장비 간의 절대 거리를 기반으로 협착 위험을 감지하는 등 3차원 기반 안전 관리 대상의 범위를 확장한다. 이때, BIM의 공간적 정보와 결합하여 스케일의 기준점이 될 기준 물체를 설정하거나, BIM 모델의 평면을 기준 지면으로 확보함으로써 스케일 드리프트 현상을 방지하고자 한다. 이러한 기술적 통합을 통해 다양한 유형 객체의 3차원 정밀 정보와 공간 정보를 종합 분석할 수 있는 통합 안전 관리 시스템 구축에 기여할 수 있을 것으로 기대된다.

참고문헌

1. Moon, G., Chang, J. Y., & Lee, K. M. (2019). Camera distance-aware top-down approach for 3D multi-person pose estimation from a single RGB image. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 2019, pp. 10132-10141.
2. Snehil Sanyal (2022). Construction-Site-Safety-PPE-Detection. <https://github.com/snehilsanyal/Construction-Site-Safety-PPE-Detection>
3. Lo, C.-C., Lin, K.-Y., & Hung, W.-H. (2023). Real-Time Detection of Safety Helmet and Vest of Workers for Construction Safety Using YOLO-Based Deep Learning Models. *Sustainability*, 15(1), 391.
4. Zhang, Y., Huang, S., Qin, J., Li, X., Zhang, Z., Fan, Q., & Tan, Q. (2025). Detection of helmet use among construction workers via helmet-head region matching and state tracking. *Automation in Construction*, 171, 105987.
5. Chen, Y., Zhang, W., Han, S., & Lin, J. (2023). Real-time ergonomic risk assessment in construction using a co-learning-powered 3D human pose estimation model. *Computer-Aided Civil and Infrastructure Engineering*, 38(2), 138-157.
6. Wen, C., Fukuda, T., & Sekimoto, Y. (2023). 3D excavator pose estimation using projection-based pose optimization for contact-driven hazard monitoring. *Automation in Construction*, 151, 104900.
7. Yan, H., Liu, C., Yang, X., & Feng, K. (2025). Real-time digital twin-driven 3D near-miss detection system at construction sites. *Journal of Construction Engineering and Management*, 151(4), 04025021.