

YOLO 기반 실시간 객체 탐지를 활용한 산업기계 비상정지 시스템 개발

임재빈, 박병국, 김완경
시큐어포인트

{jblim, bkpark, wkkim}@securepoint.co.kr

Development of an Emergency Stop System for Industrial Machinery Using YOLO Based Real-Time Object Detection

Jaebin Lim, Byeong Kuk Park, Wankyung Kim
AX Technology Research Center, SecurePoint

요약

최근 딥러닝 기반 객체 검출 기술의 발전으로 산업현장에서 YOLO(You Only Look Once)와 같은 인공지능 기반 알고리즘을 이용한 안전관리 시스템의 활용이 확대되고 있다. 그러나 대부분의 안전관리 시스템은 객체 인식을 통해 안전모 미착용 등 작업환경 기준 준수에 초점이 맞춰져 있다. 본 연구는 자체 구축 데이터셋을 통해 학습된 YOLO 알고리즘을 활용하여 객체 인식뿐만 아니라 객체의 움직임을 통해 비정상 행동을 탐지하고 이를 기반으로 산업기계를 정지시키는 시스템을 개발한다. 특히 산업재해 중 끼임 사고와 관련된 작업자의 비정상 행동을 실시간으로 탐지함으로써 끼임으로 인한 산업재해를 사전에 방지하고자 한다. 시스템은 영상 기반 객체 탐지와 행동 분석으로 정상 행동과 비정상 행동을 구분하고 위험 상황 발생 시 산업기계 제어 신호를 보내 비상정지 장치를 작동시킨다. 개발된 모델은 우수한 mAP(mean Average Precision)를 보여 주며 실시간 이상행동 탐지에서 또한 우수한 성능을 보여준다. 이를 통해 기존 사후 대응 중심 안전관리의 한계를 보완하고, 산업재해 예방을 위한 선제적 대응 체계를 제공할 수 있을 것으로 기대된다.

1. 서론

산업의 발전에 따라 현대 산업현장에서 기계화 및 자동화가 가속화되고 있으며 특히 산업기계는 대형화되고 있다. 또한 산업현장에서 발생하는 끼임, 떨어짐, 넘어짐 등의 산업재해는 여전히 산업기계와 높은 연관성을 보인다. 이러한 산업재해를 예방하기 위해 작업자는 개인보호구를 착용하고 있으며 2014년 산업재해 현황에 따르면 개인보호구를 착용할 경우 부상 발생률이 약 87% 이상 감소하는 것으로 보고되었다(KOSHA, 2016)[1]. 그러나 개인보호구는 사고 발생 시 피해 저감에는 효과적이지만 사고를 사전에 인지하고 능동적으로 대응하는 데에는 한계가 있다. 안전보건공단에 따르면 2025년 산업재해 사고 유형은 그림 1에서 보이는 것처럼 넘어짐 28,606건(19.4%), 떨어짐 13,277건(9.0%), 끼임 12,240건(8.3%), 절단·베임·찢림 11,418건(7.8%), 부딪힘 10,051건(6.8%)으로 보고되었다[2]. 산업기계의 끼임 사고는 8.3%로 매우 중대한 사고 유형으로 대부분 단순한 부상이 아닌 중상 또는 사망에 이르는 사고 유형으로 작업자의 생존과 직결되는 사고 유형이다.

최근 AI 발전으로 산업현장에서도 지능형 알고리즘을 활용하는 기술이 빠르게 발전하고 있지만, 주로 안전모 미착용 등 작업환

경 기준 준수 여부를 판단하는 데에 초점을 두고 있다[3, 4]. 이는 작업자의 기본적인 안전 확보를 위한 소극적인 방법으로 작업자의 이상행동 기반 위험 상황을 사전에 인지하는 등의 적극적인 AI 기술 활용이 요구된다. 따라서 최근에는 작업자의 행동 패턴을 기반으로 위험 상황을 능동적으로 인지하고 대응하기 위한 연구가 활발히 진행되고 있다.

인공지능 기반 이상행동 탐지 기술은 영상 데이터를 기반으로 하며 객체의 위치와 클래스를 인식하는 객체 탐지 기술이 핵심 요소로 활용된다. 해당 기술은 대규모 연산을 요구하는 특성으로 인해 초기에는 고성능 컴퓨팅 환경에서 구현되었으나 최근에는 연산 최적화 및 경량화 기법의 발전으로 모바일 및 엣지 디바이스에서도 실시간 분석과 처리가 가능하다[5, 6, 7]. 이에 따라 현장 적용 가능성이 확대되고 있으며 시스템의 성능과 효율성을 검증하기 위한 연구가 활발히 이루어지고 있다. 객체 탐지 기술에서는 CNN(Convolutional Neural Network)을 기반으로 한 접근 방식이 기존 알고리즘[8, 9] 대비 우수한 성능을 보이고 있다[10]. 객체 탐지 기술의 초기 연구에서는 Two-Stage 구조가 주류를 이루었으며 대표적으로 R-CNN[11], Fast R-CNN[12], Faster R-CNN[13] 등이 있다. Two-Stage 모델은 Region Proposal을 생성한 후 각 영역에 대해 객체를 분류하는 방식으로 높은 정확도를 보

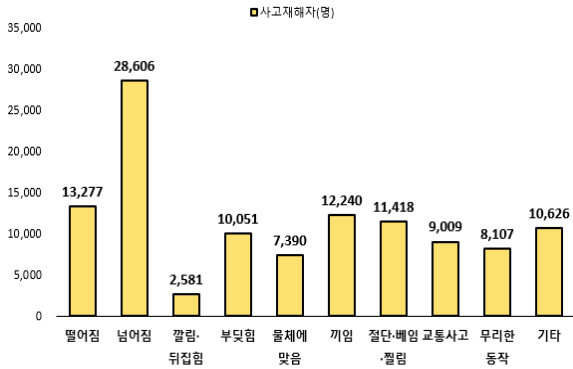


그림 1. 재해유형별 사고재해자[2]

이지만, 연산 비용이 커 실시간 처리에는 어려움이 있다. 반면 SSD[14]와 YOLO[15]와 같은 One-Stage 구조는 Region Proposal 단계 없이 전체 이미지에서 객체의 위치와 클래스를 동시에 예측함으로써 처리 속도를 크게 개선하였다. 이러한 특성으로 인해 One-Stage 기반 모델은 산업현장에서 요구되는 실시간 이상행동 탐지 시스템에 적합하며, 실제로도 산업 안전 관리 분야에서도 효과적으로 활용되고 있다.

따라서 본 연구에서는 실시간 이상행동 탐지에 효과적인 YOLO 모델을 활용하여 작업자의 정상 및 비정상 행동을 구분하고, 비정상 행동이 감지될 경우 산업기계에 제어 신호를 전달하여 비상정지 장치를 작동시키는 시스템을 개발한다. 해당 시스템은 객체 인식뿐만 아니라 이상행동을 실시간으로 탐지함으로써 끼임 사고를 사전에 예방하는 것을 목표로 한다. 이를 통해 기존의 사후 대응 중심 안전관리의 한계를 보완하고, 산업재해 예방을 위한 선제적 대응 체계를 구축할 수 있을 것으로 기대된다.

2. 연구방법

2.1 YOLOv5

본 연구에서는 산업현장에서 작업자의 이상행동을 실시간으로 탐지하기 위해 YOLOv5[16] 기반의 객체 탐지 모델을 사용한다. YOLOv5는 단일 신경망 구조를 기반으로 입력 영상을 한 번의 연산으로 객체의 위치와 클래스를 동시에 예측하는 One-Stage Detector로 높은 탐지 속도와 정확도를 동시에 확보할 수 있어 실시간 응용에 적합하다. YOLOv5는 모델 크기에 따라 YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x와 같이 다양한 모델을 제공한다. 각 모델은 정확도, 연산량, 처리 속도 측면에서 서로 다른 특성을 보인다. 그림 2 은 YOLOv5 모델별 성능비교 결과를 나타내며 모델 크기가 증가할수록 정확도는 향상되는 반면 처리 속도는 감소하는 경향

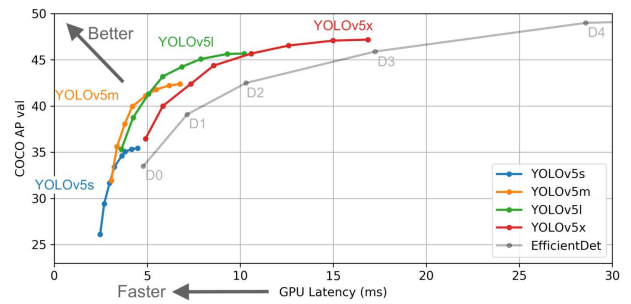


그림 2. YOLOv5 모델별 성능 비교[17]

표 1. YOLOv5 모델별 파라미터 및 추론 속도 비교

Model	Size(pixels)	Params(M)	Speed(ms)
YOLOv5s	640	7.2	6.4
YOLOv5m		21.2	8.2
YOLOv5l		46.5	10.1
YOLOv5x		86.7	12.1

을 보인다. 표 1은 YOLOv5 모델별 파라미터 수 및 추론 속도를 비교한 결과로 모델 크기가 증가할수록 파라미터 수와 연산 복잡도가 증가하고 이에 따른 추론 속도 저하가 발생함을 확인할 수 있다.

개발된 시스템은 산업기계 주변의 위험 상황을 실시간으로 탐지하고 즉각적인 비상정지를 수행해야 하므로 높은 정확도뿐만 아니라 낮은 지연 시간이 필수적이다. 이러한 요구사항을 고려하여 본 연구에서는 비교적 적은 파라미터 수와 빠른 추론 속도를 가지면서도 우수한 탐지 성능을 보여주는 YOLOv5s 모델을 사용한다. 비상정지 시스템에 적용하기 위해 사전 학습된 COCO 데이터셋 기반 가중치를 활용하고, 자체 구축 데이터셋으로 Fine-Tuning을 진행하였다. 또한 산업 환경에 맞게 이상행동(끼임) 클래스를 재정의하여 모델 학습을 진행하였다. 최종적으로 학습된 모델은 입력 영상으로부터 작업자의 위치 및 행동 정보를 실시간으로 추출하며 해당 결과는 비상정지 시스템의 입력으로 활용한다.

2.2 데이터셋 구성

본 연구에서는 산업현장에서 발생할 수 있는 상황을 분류하기 위해 안전(Normal), 주의(Caution), 경고(Warning), 긴급(Emergency)의 4개 클래스로 데이터셋을 구성하였다. 데이터는 실제 산업 환경과 유사한 환경을 구축하여 직접 촬영을 통해 수집하였으며, 각 클래스별로 100장의 이미지를 확보해 총 400장의 데이터셋을 구성하였다. 수집된 데이터는 모델의 학습 및 성능 평가를 위해 전체 데이터를 학습용 데이터와 테스트용 데이터로 7:3의 비율로 분할하였다.

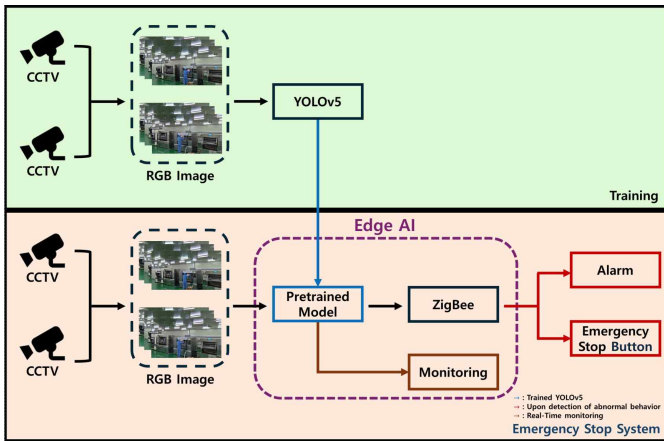


그림 3. YOLO 기반 산업기계 비정상지 시스템 구성도

표 2. Jetson AGX Xavier 모듈 사양

Jetson AGX Xavier 32GB Module	
CPU	8-core ARM 64-bit
GPU	512-core NVIDIA Volta
Memory	32GB LPDDR4x
Storage	32GB eMMC 5.1
Power	10~30W

3. 실험 및 결과 분석

3.1 개발 절차

제한하는 이상행동 탐지 및 비정상지 시스템을 구축하기 위해 그림 3에 나타난 구성도를 기반으로 시스템을 개발한다. **첫 번째** 모든 학습 과정은 AMD Ryzen 9 5900X 12-Core CPU와 Window 10, GPU는 NVIDIA GeForce RTX 3090, Python 3.10버전, Pytorch 2.0.1에서 진행하였다. 학습이 완료된 YOLOv5s 모델 가중치를 Edge 디바이스인 Jetson AGX Xavier에 탑재하며, 구체적인 하드웨어 사양은 표 2에서 확인할 수 있다. **두 번째** Jetson AGX Xavier로 입력되는 CCTV의 실시간 RGB 영상 데이터는 전처리 과정을 거쳐 사전에 학습된 YOLOv5s 모델로 전달된다. 모델은 입력된 프레임 내 작업자의 행동을 분석하여 안전, 주의, 경고, 긴급의 4단계 등급으로 분류한다. 이때 시스템의 신뢰성을 확보하기 위해 특정 신뢰도 임계값을 설정하여, 모델이 해당 클래스로 분류한 기준을 초과하는 경우에만 최종 상태로 확정하도록 로직을 설계하였다. 이를 통해 산업현장의 노이즈나 일시적인 오탐지로 인해 비정상지 장치가 불필요하게 작동하는 것을 방지하고 판정의 정확도를 높였다. **세 번째** 추론 결과가 '긴급'으로 분류될 경우 ZigBee 통신을 통해 제어 신호를 전송하며 이때 ZigBee 모듈은 XBee ZigBee TH(S2C) 을 사용한다. 이를 통해 하드웨어 비정상지 장치를 구동하여 기계를 강제 중단하도록 구현하였다.

3.2 평가지표

개발된 시스템의 성능을 평가하기 위해 본 연구에서는 객체 탐지 모델의 정확도를 종합적으로 나타낼 수 있는 지표인 mAP를 활용하였다. mAP는 IoU 임계값에 따라 산출되는 지표로, mAP@0.5와 mAP@0.5:0.95는 대표적인 평가 기준이다. 각 지표 값이 높을수록 객체 탐지 정확도가 우수함을 의미한다. 또한, 모델의 실시간 처리 성능을 평가하기 위해 FPS를 추가적으로 측정하였으며 FPS가 높을수록 더 많은 프레임을 빠르게 처리할 수 있음을 의미한다.

3.3 정량적 및 정성적 실험 결과

본 논문에서는 YOLOv5 계열 모델(YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x)을 이용하여 커스텀 데이터셋으로 학습한 후 이상행동(끼임) 탐지 성능을 비교·평가하였다. 성능 평가는 mAP@0.5, mAP@0.5:0.95, FPS를 기준으로 수행하였다. 실험 결과 YOLOv5x는 가장 우수한 mAP 성능을 나타냈으나 약 4 FPS의 낮은 처리 속도로 인해 실시간 적용에는 한계가 있는 것으로 나타났다. 반면 YOLOv5s는 mAP 및 FPS 측면에서 균형 잡힌 성능을 보였으며 해당 결과는 표 3을 통해 확인할 수 있다. 추가적으로 YOLOv5s의 탐지 결과를 정성적으로 분석하였고, 그 결과는 그림 4에서 확인할 수 있다. YOLOv5s 모델은 4가지 클래스에 대해 전반적으로 높은 신뢰도로 객체를 탐지하며 정상 행동과 비정상 행동을 명확하게 구분하는 성능을 보여준다. 또한 실험 환경에서 수집된 데이터에 대해 일관된 탐지 결과를 나타내며 안정적인 동작 특성을 확인할 수 있다. 이러한 결과는 개발된 모델이 실제 산업현장 환경에서도 효과적으로 적용될 수 있을 것으로 기대된다.

4. 결론

본 연구에서는 산업현장에서 발생하는 끼임 사고와 같은 중대 산업재해를 사전 예방을 위해 YOLO 기반의 실시간 이상행동 탐지를 이용한 비정상지 시스템을 개발한다. 기존 산업 안전 관리 방식은 개인보호구 착용과 같은 사후 대응 중심에 머물러 있다. 이를 개선하기 위해, 작업자의 행동 패턴을 기반으로 위험 상황을 능동적으로 인지하고 즉각적으로 대응할 수 있는 시스템을 구축한다. 제안된 시스템은 객체 탐지 기반으로 작업자의 정상 및 비정상 행동을 실시간으로 분류하며 위험 상황이 감지될 경우 ZigBee 통신을 통해 산업기계에게 제어 신호를 전달하여 비정상지 장치를 작동시킨다. 이를 통해 단순한 모니터링 수준을 넘어 실제 산업 기계의 직접적인 제어까지 수행함으로써, 사고 발생 가능성을 근본적으로 차단할 수 있도록 설계하였다. 또한 엣지 디바이스 환경에서도

표 3. YOLOv5 모델별 정량적 실험 결과

Model	mAP@0.5	mAP@0.5:0.95	FPS
YOLOv5s	91.4	88.0	16
YOLOv5m	93.9	90.6	10
YOLOv5l	94.2	91.0	6
YOLOv5x	95.4	91.1	4

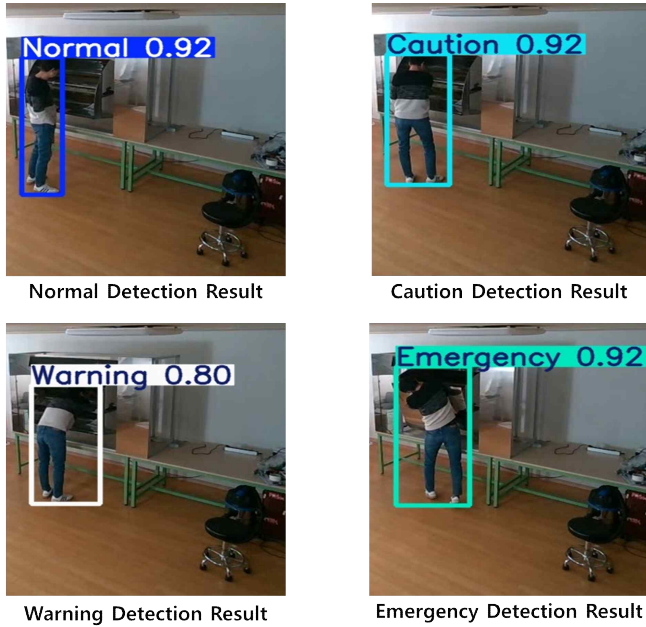


그림 4 YOLOv5s 정성적 결과

실시간 처리가 가능하도록 구현함으로써 실제 산업현장에 적용 가능한 수준의 성능과 효율성을 확보하였다. 결론적으로 본 연구에서 개발한 시스템은 산업기계 관련 위험 상황을 사전에 탐지하고 즉각적으로 대응할 수 있는 지능형 안전관리 방안을 제시한다. 향후 다양한 환경에서의 데이터 확보를 통해 시스템의 성능을 개선할 수 있을 것으로 기대된다.

Acknowledge

본 성과물은 중소벤처기업부에서 지원하는 2024년도 스마트 제조혁신 기술개발(R&D)(RS-2024-00448164)의 연구수행으로 인한 결과물임을 밝힙니다.

참고 문헌

[1] 길국호, “산업재해에서의 개인보호구 특성 분석”, 대한인간공학 회지, pp. 753-764, 12, 2017.
 [2] 안전보건공단, 「2025년 산업재해현황」, 2025.
 [3] 김대성, et al. “Faster R-CNN 학습데이터 구축과 모델을 이용한 안전모 탐지 연구”, J. Korean Soc. Hazard Mitig, pp. 119-126, 2020.
 [4] 박용석, et al. “스마트건설 현장에서 개인 보호장비 검출의 개선 방법”, 한국통신학회논문지, pp. 2202-2209, 2020.

[5] 윤현석, et al. “엣지 디바이스용 실시간 열화상 객체 검출을 위한 YOLOv5 기반 경량화 방법론.”, 방송공학회논문지, pp. 703-712, 2024.
 [6] 김철진, et al. “대용량 학습 데이터를 활용한 YOLO 기반 실시간 화재 감지 시스템 구축 연구”, 한국산학기술학회, pp. 10-17, 2025.
 [7] 양승민, et al. “AI Accelerator Implementation Technology for On-Device Deployment in Defense AI.”, 한국컴퓨터종합학술대회, pp. 877-879, 2025.
 [8] Felzenszwalb, Pedro F., et al. “Object Detection with Discriminatively Trained Part-Based Models.”, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), pp 1627-1645, 2010.
 [9] Dalal, Navneet, and Bill Triggs. “Histograms of Oriented Gradients for Human Detection.” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886-893, 2005.
 [10] Qaddour, Jihad. “Object detection performance: a comparative study.”, Research Square, 2023.
 [11] Girshick, Ross, et al. “Rich feature hierarchies for accurate object detection and semantic segmentation.”, In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 580-587, 2014.
 [12] Girshick, Ross. “Fast r-cnn”, In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 1440-1448, 2015.
 [13] Ren, Shaoqing, et al. “Faster r-cnn: Towards real-time object detection with region proposal networks.”, Advances in neural information processing systems (NeurIPS), pp. 91-99, 2015.
 [14] Liu, Wei, et al. “Ssd: Single shot multibox detector.”, European conference on computer vision (ECCV), pp. 21-37, 2016
 [15] Redmon, Joseph, et al. “You only look once: Unified, real-time object detection.”, Proceedings of the IEEE conference on computer vision and pattern recognition. (CVPR). pp. 779-788, 2016.
 [16] G. Jocher et al., “YOLOv5 by Ultralytics,” <https://github.com/ultralytics/YOLOv5>, 2020.
 [17] Lee, Yongwoo, et al. “Deep-learning-based object filtering according to altitude for improvement of obstacle recognition during autonomous flight.”, Remote Sensing, 2022.