

딥러닝 객체 감지 기반의 지하공동구 작업자 낙상 검출 비교 연구

김정수

한국건설기술연구원 구조연구본부

Comparative Study on Deep-learning Object Detection Based Recognizing Worker Falls in Underground Utility Tunnels

Jeongsoo Kim

Dept. of Structural Engineering Research, Korea Institute of Civil Engineering and Building Technology

요약 지하공동구 작업자의 이상행동을 감지하기 위한 이미지 기반의 여러 딥러닝 모델이 제안되고 있다. 이들 모델의 다수는 딥러닝 추론 결과에 낙상을 판단하기 위한 규칙을 적용하고 있으나, 규칙 기반 모델은 카메라 각도 및 위치 등에 영향을 받아 낙상 검출 성능에 제한적일 수 있다. 본 논문은 지하공동구 작업자의 낙상 검출에 대한 규칙 및 학습 기반의 객체 감지 딥러닝 모델의 성능을 비교하였다. 규칙 기반의 낙상 판별 모델의 경우, COCO 데이터셋(Model A)과 지하공동구 데이터셋(Model B)으로 각각 학습된 YOLOv5 모델을 사용하며, 작업자 자세와 무관하게 경계상자의 종횡비를 활용하여 작업자의 낙상을 판별한다. 반면 낙상 학습 기반 모델(Model C)은 YOLOv5에 지하공동구 작업자의 낙상과 정상상을 구분할 수 있도록 학습하였다. 시험 결과는 Model C의 F1 지표가 0.867로 Model A 및 B 대비 각각 76%, 10% 높게 나타나, 낙상 자체를 학습 기반 모델이 종횡비 규칙 기반 모델 대비 지하공동구 작업자의 낙상 탐지에 효과적임을 보여주었다. 또한 규칙 기반 모델의 낙상 탐지 성능 개선을 위해 작업자 인식 성능 제고가 필요함을 알 수 있다. 본 논문의 결과는 지하공동구 작업자의 안전 모니터링을 위한 효율적인 방법 선택에 기여할 수 있을 것으로 기대된다.

Abstract Several image-based deep learning models have been proposed to detect abnormal behaviors of workers in underground utility tunnels. Many of these proposed models apply rules to the deep learning inference results to determine falls. On the other hand, rule-based models can be limited in their fall detection performance because of the influence of camera angles and positions. This paper compares the performance of rule-based and learning-based object detection deep learning models for detecting workers' falls in underground utility tunnels. For the rule-based fall detection model, this study used YOLOv5 models trained on the COCO dataset (Model A) and the underground utility tunnel dataset (Model B), respectively, and determined worker falls using the aspect ratio of the bounding box regardless of the worker's posture. In contrast, the learning-based model (Model C) for falls was trained on YOLOv5 to distinguish between falls and the standing postures of workers in underground utility tunnels. For the test data, the F1 score of Model C for detecting workers' falls in underground utility tunnels was 0.867, which was 76% and 10% higher than Models A and B, respectively. Hence, the learning-based model was more effective than the rule-based one. In addition, improving worker recognition performance was necessary to enhance the fall detection performance of rule-based models. The research findings are expected to contribute to the selection of efficient methods for the safety monitoring of underground utility workers.

Keywords : Fall Detection, Object Detection, Underground Utility Tunnel, YOLOv5, Bounding Box Ratio

본 논문은 2024년도 정부(과학기술정보통신부, 국토교통부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2020-0-00061, 디지털트윈 기반의 지하공동구 화재·재난 지원 통합플랫폼 기술개발)

*Corresponding Author : Jeongsoo Kim(Korea Institute of Civil Engineering and Building Technology)

email: jeongsookim@kict.re.kr

Received May 13, 2024

Revised June 12, 2024

Accepted August 2, 2024

Published August 31, 2024

1. 서론

지하공동구는 상·하수도, 전력·통신선, 가스 파이프 등 여러 사회기반 시설물을 공동 수용하는 지하 구조물이다. 내부 시설물 특성상 연장이 수 km에 이르며, 내부 시설물로의 접근을 위해 좁고 긴 통로가 미로처럼 연결되어 있다. 내부 시설물의 유지관리를 위해 관리 인원이 일 단위로 지하공동구를 한 차례 의무 순찰하며, 유지보수 등 필요시 외부 작업자의 출입이 발생한다. 또한 지하공동구 내부는 공기 흐름이 원활하지 않아 맨홀, 터널 등과 함께 산업안전보건법에 따라 밀폐공간으로 규정되고 있으며, 내부에 배치된 시설물의 유형에 따라 고온 질식, 유해가스 중독, 다습한 바닥 등으로 인한 낙상 유발 요인이 상존한다.

지하공동구 내부 순찰 및 작업 인원의 의식불명 등의 치명적인 상황이 발생한 경우, 사고 발생 지점까지의 접근성이 떨어지므로 신속한 사고 상황 인식 및 대응이 중요하다. 지하공동구는 내부의 이상 상황을 조기에 파악할 수 있도록 CCTV 및 센서를 통해 수집된 데이터가 관제센터로 송신되고 있으나, 관제 인력이 해당 데이터를 연속해 주시하는 것은 물리적으로 불가하다. 또한 공동구에 이상감지 목적으로 설치된 여러 물리적 센서도 잦은 오경보로 인해 사실상 무용지물이어서 사고 발생을 조기에 인식하기 어려운 환경이다.

사람의 의식불명 등 이상상황 감지에 대한 기술적 수요는 지하공동구 외 여러 분야에서 존재해 왔고, 최근 딤러닝 기법을 활용한 관련 연구도 빠르게 확산되고 있다. 요양원 내 고령자[1-3], 조선소 및 건설 현장 작업자[4-6] 등 낙상 가능성 및 위험성이 높은 분야를 중심으로 딤러닝 기술을 적용하는 많은 연구가 이뤄져 왔으나, 밝은 실내 환경 및 야외에서의 낙상을 감지하는 것에 초점이 맞춰져 있다. 반면 지하공동구는 내부 조명이 간헐적으로 배치되어 평균 조도가 30~40 lx 수준으로 낮고, 먼지 등에 의해 영상 품질이 상대적으로 낮다[7]. 또한 카메라 각도 및 관측 거리 등의 환경이 기존 타 산업 부문과 상이하고 설치 조건에 따라 최적화가 필요하다. 따라서 지하공동구 환경 기반의 이미지를 활용한 딤러닝 모델의 개발 및 평가가 요구된다.

딤러닝 모델을 활용한 낙상 판별 연구의 다수는 딤러닝 검출 결과로부터 얻은 기하학적 정보 또는 해당 정보의 시계열적 변화에 대해 사전에 정의한 낙상 규칙을 적용하는 접근 방식을 주로 사용한다. 이 방법은 사람에 대한 딤러닝 검출이 제대로 이뤄지고, 낙상을 판단하기 위

한 규칙이 적용 환경을 고려해 적절히 적용되는 것을 전제한다. 따라서 딤러닝 모델이 사람을 잘 인식할 수 있도록 많은 학습 데이터가 필요한데, 데이터 구축이 제한적인 경우 MS-COCO 등 방대한 데이터셋에 의해 사전 학습된 공개 모델을 활용하기도 한다[3,8-11]. 낙상 검출을 위해 키포인트 및 객체 감지 모델이 주로 적용되며, 특히 후자의 경우 검출 결과로 얻는 경계상자 중횡비를 활용해 낙상을 판별하거나[2,11], 낙상 자체를 학습하는 모델[4,12]로 구분할 수 있으며 관련된 여러 연구 결과가 제시되고 있다. 그러나 지하공동구 작업자의 낙상 감지 모델에 대한 연구는 부족한 실정이며, 특히 낙상 감지 목적의 여러 딤러닝 모델 간 비교 연구는 부재한 상황이다. 동일 대상에 대한 여러 낙상 감지 모델 간 비교를 통해, 모델 구축에 소요되는 시간·비용 및 기대성능 등을 고려한 의사결정을 지원할 수 있는 관련 연구가 필요하다.

본 연구는 지하공동구 작업자의 낙상 검출을 위한 여러 딤러닝 객체 감지 방법 가운데, 중횡비 규칙을 연계한 딤러닝 모델과 낙상 자체를 학습한 딤러닝 모델의 결과를 분석한다. 중횡비 규칙 연계 모델의 경우 MS-COCO 데이터셋으로 사전 학습된 YOLOv5 모델(Model A)과 지하공동구 영상을 추가로 학습한 YOLOv5 모델(Model B)의 작업자 검출 및 낙상 판별 결과를 비교한다. 또한 작업자의 정상 및 낙상을 학습한 모델(Model C)의 낙상 판별 성능을 함께 비교하였다.

2. 연구방법

2.1 YOLOv5

YOLOv5는 실시간 객체 감지 분야에서 널리 사용되는 딤러닝 모델 중 하나이다. 이미지를 그리드 단위로 분할해 객체 분류와 경계상자 위치 예측을 동시에 수행함으로써, R-CNN 등 다른 딤러닝 모델 대비 가볍고 추론 속도가 빠른 장점을 가지고 있고 다중 스케일 예측 및 CSPNet과 같은 알고리즘을 도입해 정확도도 높다[13]. 2023년 기준으로 YOLOv8까지 발표된 바 있으나, YOLOv5는 2020년 발표된 이후 많은 프로젝트를 통해 안정성 및 신뢰성이 확인되고 있어 여전히 여러 연구 및 실무 프로젝트에서 활용되고 있다.

YOLOv5는 하드웨어 성능에 맞는 모델을 선택할 수 있도록 5가지 크기의 신경망 구조를 제공한다. 본 논문은 지하공동구 내부가 일반환경 대비 어둡고 저조도 CCTV를 통해 수집된 이미지를 실시간으로 처리해야 함

을 감안하여 YOLOv5 medium 모델을 선택하여 모델 별 학습, 검증, 시험을 수행하였다.

2.2 데이터셋

2.2.1 MS-COCO

YOLOv5에서 기본 제공하는 사전학습 모델은 MS-COCO(2017)를 사용하며, 해당 데이터셋은 총 80 개 객체에 대한 이미지 및 라벨링 데이터를 제공한다. 여기서 사람이 포함된 훈련용 이미지는 약 64,000개로 전체 이미지 가운데 18.7 %를 차지한다[11]. 쓰러진 사람을 포함한 이미지를 일부 포함하고 있으나, 대부분 사람이 착석 또는 기립해 있거나 상반신 및 얼굴 등 신체 일부만 포함하는 이미지가 다수인 것으로 파악된다.

2.2.2 지하공동구 낙상 데이터셋

지하공동구는 보안시설로 일반인의 출입이 제한되는 시설이다. 본 논문은 2021년부터 청주시설관리공단의 협조를 얻어 ○○공동구 내부에서 다양한 연출 영상을 획득하였고, Fig. 1에 나타낸 바와 같이 영상마다 1~3인의 작업자가 서로 다른 색상의 안전모 및 작업 조끼를 여러 조합으로 착용하도록 하였으며, 전력구 및 상수-통신구

일부 구간에서 작업자가 보행, 기립, 낙상을 반복하도록 하여 저조도 CCTV 및 휴대전화 카메라를 사용해 영상을 취득하였다(36개). 또한 실내 복도에서 작업자 낙상 상황을 연출하여 얻은 영상(3개)과 지하공동구 영상 촬영 준비 과정에서 얻은 이미지 60장을 데이터셋에 포함하였다.

동일 영상으로부터 추출된 이미지가 훈련, 검증, 시험 데이터에 혼재되면 모델의 성능 지표가 실제 검출 성능 대비 왜곡되게 평가될 수 있다. 또한 지하공동구 환경이라는 제한된 공간에 활용되는 딥러닝 모델인 점을 고려 하더라도 유사 이미지의 학습과 성능 평가로 인해 과적합된 모델이 도출될 수 있으므로, 데이터셋 구성 시 사실상 훈련에 사용된 이미지와 동일한 이미지가 검증 및 시험데이터에 사용되는 것을 최소화할 필요가 있다. 따라서 본 논문은 전체 39개 영상 가운데, 1명만 출현하는 영상은 모두 학습 목적으로 사전 분류하고, 영상마다 초당 1개 프레임을 추출해 작업자를 정상, 낙상으로 라벨링하였다. 2명 이상 출현하는 영상과 촬영 준비 과정에 대한 이미지는 검증 및 시험 목적으로 구분해 라벨링을 수행하였다. 위의 과정을 통해 구축된 본 논문의 이미지는 전체 16,056장이며, 전체 이미지에 대해 20,305개의 사람 객체가 포함된다. 정상과 낙상 객체 수는 각각 8,574개, 11,731개이며, 데이터셋의 객체별 객체 상자의 중첩비



Fig. 1. Sample images of workers in the underground utility tunnel

는 Fig. 2와 같이 0~10 크기로 분포된다. 구축한 데이터셋에 대해 사람 객체는 대부분 6 이하의 종횡비를 가지고, 낙상 객체의 경우 87%가 종횡비 2 이하였다.

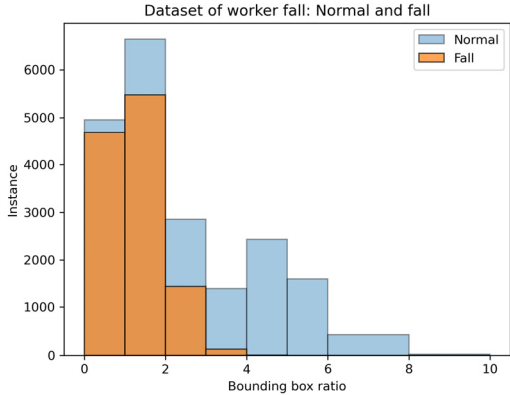


Fig. 2. Distribution of bounding boxes ratio for dataset of worker fall in underground utility tunnels

본 논문에서 사용한 지하공동구 작업자 데이터셋은 기본적으로 Model C의 학습을 위해 정상과 낙상으로 라벨링 되어 있다. 반면 Model B는 작업자의 정상 및 낙상 여부와 관계없이 작업자를 모두 단일한 객체로 인식하도록 학습이 필요하다. 따라서 Model C 학습 목적으로 구축된 데이터셋을 단일 클래스(작업자)로만 구성되도록 변환 작업을 수행해 Model B의 학습, 검증, 시험을 위한 데이터셋을 생성하고, 이를 학습에 활용하였다.

2.3 낙상 예측을 위한 딥러닝 모델

2.3.1 Model A

Model A는 YOLOv5를 COCO 데이터셋으로 사전학습해 제공하는 Medium 신경망 크기의 객체 감지 모델(YOLOv5m)이다. 80개 클래스 가운데 사람(작업자)만 추론하도록 설정하였으며, 추론 결과로부터 객체별 종횡비를 계산하여 낙상 여부를 판단하도록 후처리하였다. Fig. 2에 나타난 객체별 종횡비 분포를 고려하여 낙상 기준 종횡비를 2로 설정하였다.

2.3.2 Model B

Model B는 사전 학습된 YOLOv5m 모델에 지하공동구 데이터셋을 전이 학습하여 얻은 모델이다. 단, 해당 데이터셋은 정상과 낙상에 대한 구분이 없고 작업자를 단일한 객체로 인지하도록 변환되어 있으며, Model A와 마찬가지로 낙상 판별 종횡비를 2로 정의하였다.

2.3.3 Model C

Model C는 작업자를 정상과 낙상으로 구분한 데이터셋을 사용해 YOLOv5m 모델을 전이 학습한 모델이다. 작업자의 자세를 학습하므로, 종횡비 정보가 필요하지 않으며 추론 결과 자체로 낙상을 판별할 수 있다.

3. 딥러닝 모델 학습

3.1 학습환경 및 데이터구성

Model B와 C는 지하공동구 작업자 이미지에 대한 전이 학습이 요구된다. 본 연구는 Windows 10 운영체제를 탑재한 컴퓨터에 GPU 가속이 가능한 딥러닝 학습환경을 구축하였다. 여기서 CPU 및 GPU 사양은 각각 Intel(R) Core(TM) i9-10940X @3.30GHz, NVIDIA GeForce RTX 4090이다.

YOLOv5m 모델의 학습 시 배치크기는 GPU VRAM (24GB×2EA) 사양을 고려해 32, 64, 128로 설정하였으며, 에포크는 50으로 설정하였다. 신뢰도 임계값(Confidence Threshold)과 중첩 영역 임계값(Intersection of Union, IoU)은 각각 0.25과 0.45로 설정하였으며, 학습률, 모멘텀 계수, 감쇠 계수 등 나머지 주요 매개변수는 YOLOv5m에서 제공하는 기본 설정을 사용하였다.

Model B와 C의 학습, 검증, 시험에 사용되는 이미지는 각각 12,111장, 1,957장, 1,958장이며, Model A는 별도의 모델 학습 과정이 수반되지 않으므로, 1,958장의 이미지가 시험데이터로써 활용된다.

3.2 모델별 학습결과

Model B와 C를 지하공동구 데이터셋에 대해 학습하고, 주요 결과를 Fig. 3과 Table 1에 나타내었다.

Fig. 3은 학습 진전에 따른 검증 데이터에 대한 손실함수 및 mAP 변화를 나타낸 것이다. 두 지표 모두 일정 이상 학습이 반복됨에 따라 모델 학습에 배치크기가 미치는 영향은 미미하였고, 10 Epochs가 경과함에 따라 손실함수 및 mAP 모두 수렴하는 경향을 보였다.

Table 1은 검증 데이터에 대한 모델의 손실함수 최솟값 및 최대 mAP를 정리한 것이다. 배치크기에 차이로 인한 차이는 미미하였으나, 모델별 mAP 간 차이는 0.05 수준으로 Model B가 더 우수하였다. Model C는 정상과 낙상을 구분해 학습을 진행하므로, 상대적으로 클래스별 학습 개체 수가 적고, 단순히 팔다리, 얼굴 등의 유

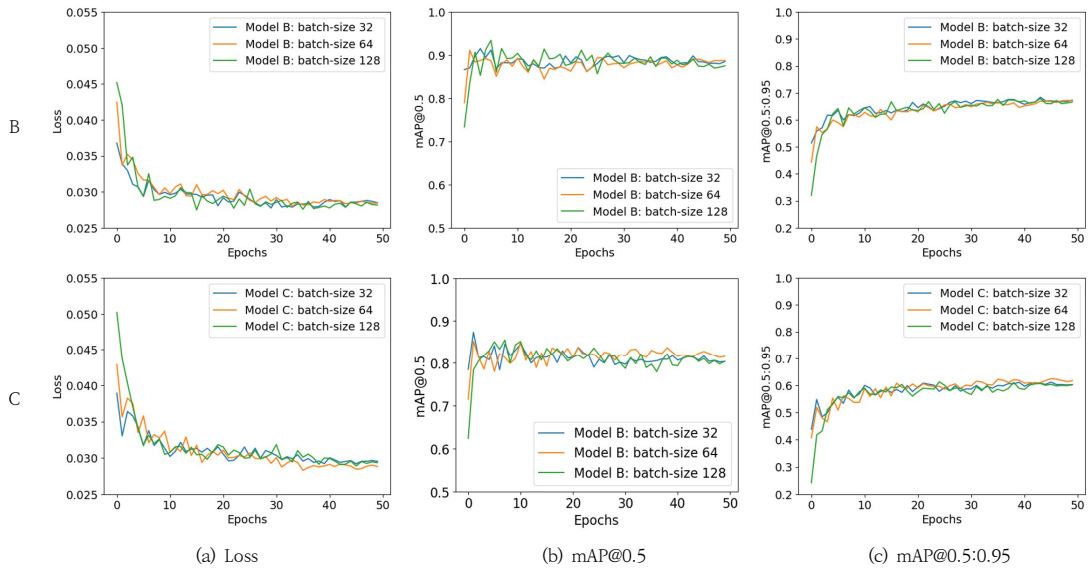


Fig. 3. Loss and mAP curves for validation data

무만으로 낙상과 정상을 구분할 수 없다. 반면 Model B는 작업자의 상태 구분 없이 학습이 이뤄지므로 학습 객체 수가 Model C 대비 많고, 얼굴, 팔·다리 등 사람을 특징할 수 있는 부분이 이미지에 포함되면 작업자를 작업자로 인식할 수 있다. 이러한 이유로 Model B의 mAP가 Model C보다 더 높게 나타난 것으로 추정되나, Table 1에서 Model B의 성능 지표는 낙상 인식에 대한 것이 아니라 작업자 인식에 대한 것이므로, 아직 모델별 낙상 검출 성능의 우위를 비교할 수 없으며 관련 내용은 4장에서 다룬다.

Table 1에 나타난 바와 같이 배치크기 변화가 모델의 손실함수 값 및 성능 지표에 미치는 영향이 미미하고, 모델의 과적합 경향도 확인되지 않는다(Fig. 3(a)). 본 논문은 배치크기 64로 학습된 모델을 선택해 모델별 낙상 검출 성능을 비교 분석하였다.

Table 1. Best metric of models from YOLOv5 for validation data (mAP: Mean average precision)

Model	Batch size	Box loss	mAP @0.5	mAP @0.5:0.95
A	-			
B	32	0.0278	0.9152	0.6839
	64	0.0280	0.9114	0.6724
	128	0.0275	0.9342	0.6791
C	32	0.0292	0.8732	0.6114
	64	0.0283	0.8514	0.6239
	128	0.0289	0.8545	0.6122

4. 모델별 낙상 검출 성능 비교

시험데이터 1,958장 이미지 가운데 8장은 지하공동구 작업자를 포함하지 않으며, 1,950장에는 3,757개의 작업자 개체가 포함되어 있다. 모델의 작업자 및 낙상 검출 성능 평가에서 모든 모델의 신뢰도 임계값과 IoU는 각각 0.25과 0.45가 적용되었고, 성능 지표는 개별 작업자에 대한 낙상 감지 결과를 기준으로 산출되었다.

4.1 작업자 인식 성능

본 논문은 Model C와 달리 Model A와 B의 작업자 낙상 검출 성능이 작업자에 대한 인식 성능에 지배적인 영향을 받게 되므로, 낙상 검출 성능뿐만 아니라 지하공동구 작업자 인식 성능을 함께 비교 분석하였다. 한편 Model C는 낙상과 정상을 학습한 모델이므로, 지하공동구 작업자 인식 성능 평가 시 낙상과 정상을 예측한 결과를 모두 작업자 인식 결과로 간주하여 지표를 산출하였다.

Table 2는 시험데이터에 대한 각 모델의 지하공동구 작업자 인식 성능 지표를 정리한 것이다. 모든 모델에 대해 작업자 인식 정밀도는 1.0에 근접하게 나타났으나(P), Model A는 타 모델 대비 재현율 및 정확도(R, A)가 상대적으로 낮았다.

Model A의 재현율과 정확도가 낮게 평가된 원인은 Model A의 낙상 작업자 인식 성능이 낮기 때문으로 조사되었다. 시험데이터(3,757개)에 대한 Model A의 추론

결과 중 인식에 실패한 작업자 개체는 1,301개였으며, 1,231개는 낙상한 작업자에 대한 것이었다. 위와 같은 결과는 비록 Model A가 COCO 데이터에 포함된 다량의 사람 이미지로 사전 학습되었으나, 학습에 사용된 대부분의 이미지가 지하공동구 및 유사 환경과 무관하고 낙상한 사람을 거의 포함하지 않고 있기 때문으로 추정된다. 또한 상수관 등 지하공동구 내부 시설물이 작업자의 신체를 가리거나, 작업자의 낙상 방향 및 자세로 인해 사람을 사람으로 구분하기 용이한 얼굴, 팔·다리 등의 신체 부위를 가린 것도 작업자 인식 저하의 원인으로 분석된다.

반면 Model B와 C는 지하공동구 내부에서 다양한 자세를 취하고 있는 작업자 이미지를 학습에 사용하여 정상뿐만 아니라 낙상 작업자에 대해서도 미탐이 적었고 (Model B: 509개, Model C: 497개), 오탐은 거의 없었다 (Model B: 2개, Model C: 3개). 그 결과 정밀도뿐만 아니라 타 성능 지표에 대해서도 Model A 대비 높은 성능을 가지고 있음을 확인할 수 있다.

Table 2. Metric of models A and B for test data to detect workers (P: Precision, R: Recall, F1: Harmonic mean of the P and R, A: Accuracy)

Model	P	R	F1	A
A	0.999	0.653	0.790	0.653
B	0.999	0.864	0.927	0.864
C	0.999	0.867	0.929	0.867

추가로, Model B와 C의 학습 데이터는 작업자의 자세를 구분하여 라벨링이 이뤄졌다는 점에서 모델 간 성능 차이가 예상되었으나 Table 2와 같이 성능 지표별 차이가 0.001 단위에서 나타나고 있어, 자세 구분 라벨링이 작업자 인식 성능 개선에 미치는 영향은 미미한 것으로 판단된다.

4.2 작업자 낙상 검출 성능

Fig. 4는 동일 이미지에 대한 모델의 추론 결과 일부를 나타낸 것이다. 그림에 정리된 각 이미지는 모델의 추론 결과와 라벨링 정보를 이미지 좌측 상단에 표출하며, 감지된 작업자에 대한 경계상자 및 상태 정보를 함께 나타낸다.

Fig. 4(a)에서 확인할 수 있듯이, 서 있는 작업자에 대

한 상태 추론은 대부분의 이미지에 대해 규칙 및 자세 추론 모델 모두 양호한 결과를 나타냈다. Model A의 경우, Fig. 4(b)와 같이 낙상한 작업자를 미인식하여 최종적으로 낙상 판별을 수행하지 못한 경우가 자주 확인되었으며, 이러한 경향은 작업자의 낙상 판별에 대한 모델의 성능 지표를 정리한 Table 3에서도 확인할 수 있다. 모델 간 정밀도 차이보다 재현율의 차이가 크게 나타났고, 이는 작업자를 미인식한 것에 주로 기인한다. 한편 Model A에서 Fig. 4(c)와 같이 작업자의 낙상이 통로를 가로질러(횡방향) 발생한 이미지에 대해 추론이 올바르게 이뤄지는 것을 확인할 수 있고, Fig. 4(b) 및 Fig. 4(d)와 같이 통로방향 낙상에 대해서는 그렇지 않은 경향을 보였다. 이는 횡방향 낙상이 통로방향 낙상에 비해 작업자의 얼굴, 팔·다리 등의 신체 부위가 명확히 드러나기 용이해 작업자 인식에 유리하기 때문이다. 또한 가로 방향으로 긴 경계상자가 생성되기 때문에 판단된다. 추가로 Model A의 경우, Fig. 4(e)와 같이 상수관을 사람으로 오인식하는 경우가 발생하였고, 이는 상수관을 다리 등의 신체 일부로 오인한 것으로 추정된다.

Model B는 지하공동구 작업자 이미지 학습을 통해 작업자 인식 성능이 개선되었기 때문에, 전반적으로 작업자 미인식으로 인한 오탐이 줄어 재현율이 크게 개선되었다. 그 결과 F1 값이 0.786으로 Model A 대비 낙상 검출 성능이 개선되었으나(Table 3), 중형비 기준을 적용하기 때문에 작업자의 낙상 방향 및 카메라 거리 및 각도 조건에 따른 오탐 결과도 확인되었다. 특히 Fig. 4(d)와 같이 통로 방향으로 작업자가 쓰러진 경우, 작업자에 대한 경계상자 중형비가 기준값을 초과하여 최종적으로 정상으로 판별하는 결과가 종종 확인되었다.

Table 3. Metric of models A and B for test data to detect a fallen worker (*Weighted metric, P: Precision, R: Recall, F1: Harmonic mean of the P and R, A: Accuracy)

Model	P	R	F1	A	
A	0.763	0.364	0.493	0.571	
B	0.841	0.737	0.786	0.770	
C	Normal	0.895	0.830	0.862	0.887
	Fall	0.919	0.820	0.867	0.856
	All*	0.909	0.824	0.865	0.869

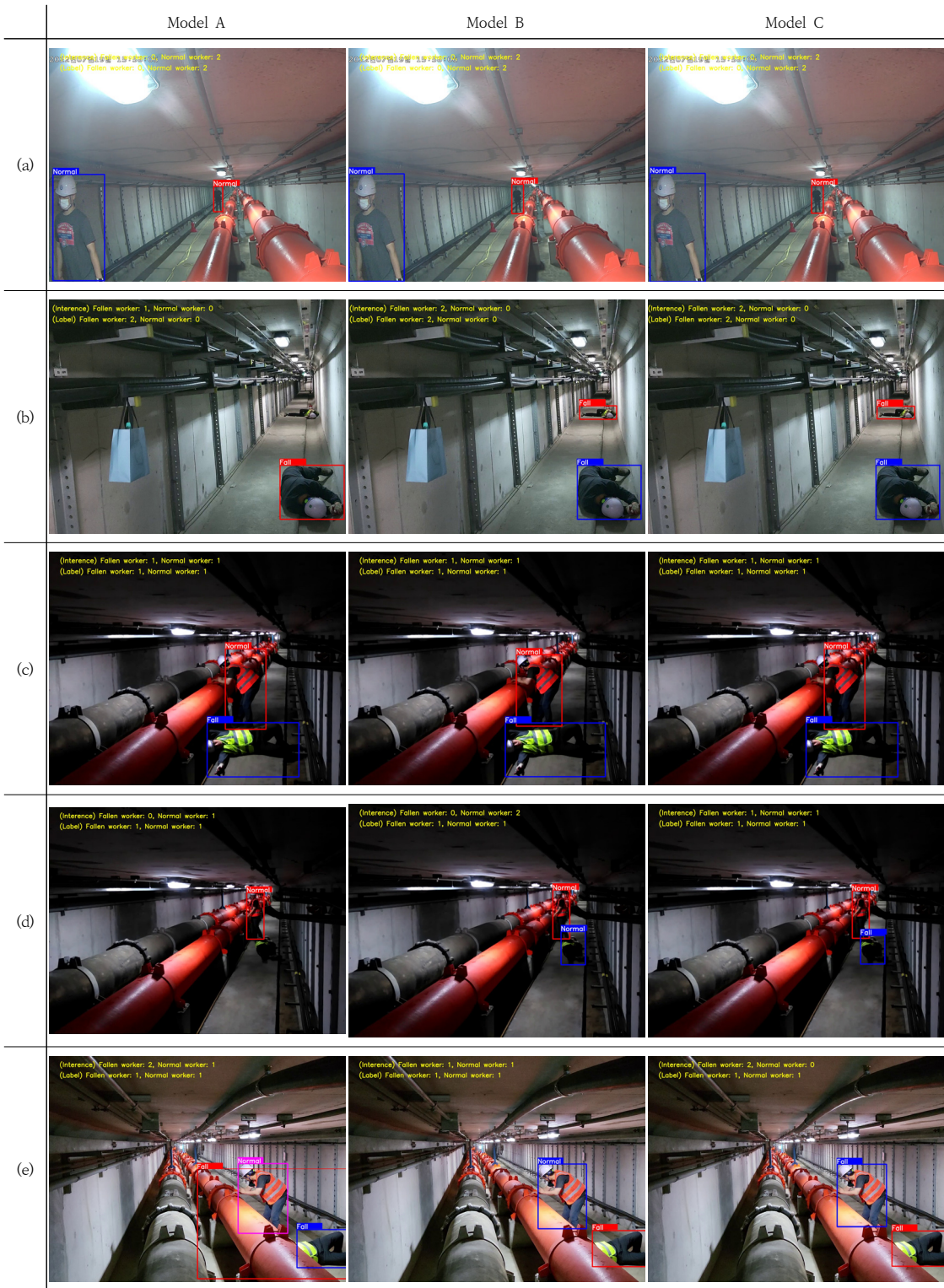


Fig. 4. Comparison with interference sample images from worker abnormal detection models in the utility tunnels

작업자의 자세를 직접 학습한 Model C는 F1 지표가 0.867로 성능이 가장 높았으며, Fig. 4와 같이 타 모델 대비 낙상 작업자를 잘 구분해 내었으나 Fig. 4(e)와 같이 오탐한 결과를 포함하였다. 해당 결과는 작업자를 인식하였으나 정상과 낙상을 잘못 판별한 경우로 Table 3의 정밀도 및 재현율이 작업자 인식(Table 2)에 비해 낮아진 원인이기도 하다. 작업자의 자세를 직접 학습하기 때문에, 일부 자세는 낙상했을 때와 유사해 이를 오인한 것으로 판단된다.

4.3 종횡비 규칙 모델 적용성 검토

시험데이터에 대한 모델별 성능을 조사한 결과, Model C의 낙상 감지 성능이 가장 우수하였다. 종횡비 규칙 모델은 자세 자체를 학습한 모델에 비해 성능이 낮았으나, 데이터 구축 노력을 줄일 수 있다는 점에서 장점이 있어 지하공동구 작업자 낙상에 대한 종횡비 규칙 모델의 적용 범위를 검토하였다.

Model B도 데이터셋 구축이 요구되나 Model C에 비해 객체 구분 및 낙상-기립 전이 과정에서 나타나는 모호한 자세에 대한 라벨링 노력을 줄일 수 있다는 점에서 장점이 있다. 특히 Table 2에 정리한 바와 같이 작업자 인식에 대한 성능 차이는 미미하였기 때문에 낙상 감지 모델을 적용하는 현장의 카메라 조건 및 작업자 낙상의 범주가 제한적이라고 한다면, 경계상자 종횡비를 활용한 기존 규칙 기반 낙상 모델도 활용성이 있다고 판단된다. 참고로 전체 데이터에서 낙상 작업자에 대한 경계상자 종횡비가 2.0인 경우가 87%였고 Model B의 성능 지표 감소가 Model C 대비 87%였으므로, 경계상자 종횡비 규칙을 유효하게 적용할 수 있는 지하공동구 조건이라면 두 모델 간의 성능 차이는 미미할 것으로 예상된다.

반면 Model A와 같은 공개 데이터셋을 활용한 접근 방법은 별도의 모델 학습이 불필요한 장점이 있으나, 종횡비 규칙 적용에 따른 불가피한 성능 감소뿐만 아니라 낙상 작업자를 충분히 포함하지 못한 데이터셋으로 모델이 학습된 경우 작업자 미인식 문제로 인해 낙상 검출 자체가 불가능한 경우가 많아 실제 적용은 어렵다고 판단된다.

5. 결론

본 논문은 지하공동구 작업자의 낙상 검출에 대한 딥러닝 객체 감지 모델에 기반한 종횡비 규칙 모델과 낙상 자세 학습 모델을 사용하여 낙상 작업자 검출 성능을 비

교 분석하였다. 이를 통해 다음과 같은 결론을 얻었다.

딥러닝 종횡비 규칙 모델의 낙상 검출 성능은 작업자 인식 성능에 의존적이었다. 지하공동구 작업자 인식 성능이 떨어지는 MS-COCO 데이터셋 사전학습모델에 종횡비 규칙을 적용한 모델은 F1 지표 0.493을 보였으나, 본 연구의 지하공동구 환경 낙상 작업자 이미지 학습 및 기존 종횡비 2.0가 적용된 모델은 F1 지표 0.786을 보였다. 한편 낙상 자체를 학습 모델이 종횡비 규칙을 적용한 모델 대비 우수한 낙상 인식 성능을 나타내었다. 낙상 자세를 학습한 모델의 경우 F1 지표 0.867로 종횡비 규칙 모델에 비해 낙상 검출 성능이 개선됨을 확인하였다. 따라서 종횡비 규칙을 적용한 모델은 자세 학습 모델 대비 데이터 구축에 필요한 노력을 저감할 수 있으나, 지하공동구 현장의 카메라 위치 및 화각, 작업자 낙상 지점 등의 다양한 변수가 존재해 낙상에 대한 오검출 가능성이 높으므로 제한적으로 적용하는 것이 필요하다.

본 논문은 지하공동구 환경에서 대해 비시계열 딥러닝 객체 감지 방법에 기초한 여러 낙상 검출 모델 간 성능을 비교함으로써, 낙상 자세를 학습한 모델이 종횡비 규칙 모델 대비 우수함을 밝혔고 지하공동구 작업자 인식 성능이 확보된다면 종횡비 규칙 모델로도 소정의 성능을 확보할 수 있음을 보였다는 점에 의의가 있다고 판단된다. 향후 시계열 딥러닝 기법을 활용한 타 낙상 검출 모델과의 성능을 비교함으로써, 낙상 검출 관련 모델 개발 및 선택에 필요한 기초 자료를 제공하고자 한다.

References

- [1] S. Choi, S. Youm, "A study on a fall detection monitoring system for falling elderly using open source hardware", *Multimedia Tools and Applications*, Vol.78, pp.28423-28434, Oct. 2019. DOI: <https://doi.org/10.1007/s11042-017-5452-9>
- [2] J. -T. Park, K. P. Han, Y. W. Park, "A Dangerous Situation Recognition System Using Human Behavior Analysis", *Journal of Korea Multimedia Society*, Vol.24, No.3, pp.345-354, Oct. 2021. DOI: <https://doi.org/10.9717/kmms.2020.24.3.345>
- [3] K. -W. Kang, S. -Y. Park, "The modified fall detection algorithm based on YOLO-KCF for elderly living alone care", *Journal of the Korea Institute of Convergence Signal Processing*, Vol.21, No.2, pp.86-91, June 2020.
- [4] K. -M. Park, S. -D. Kim, C. -O. Bae "A study on the detection of fallen workers in shipyard using deep learning", *Journal of the Korean Society of Marine Environment & Safety*, Vol.26, No.6, pp.601-605, Dec.

2020.
DOI: <https://doi.org/10.7837/kosomes.2020.26.6.601>
- [5] Y. Yoon, T. Oh, "A study on the improvement of construction site worker detection performance using YOLOv5 and OpenPose", *The Journal of the Convergence on Culture Technology*, Vol.8, No.5, pp.735-740, Sep. 2022.
DOI: <http://dx.doi.org/10.17703/JCCT.2022.8.5.735>
- [6] Y. -H Cho, H. J. Lee, "Implementation of an alarm system with AI image processing to detect whether a helmet is worn or not and a fall accident", *The Journal of Korea Institute of Convergence Signal Processing*, Vol.23, No.3, pp.150-159, Sep. 2022.
DOI: <http://dx.doi.org/10.23087/jkicsp.2022.23.3.006>
- [7] J. Kim, C. -W. Lee, S. -H. Park, J. H. Lee, C. -H. Hong, "Development of fire detection model for underground utility facilities using deep learning: training data supplement and bias optimization", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.21, No.12, pp.320-330, Dec. 2020.
DOI: <https://doi.org/10.5762/KAIS.2020.21.12.320>
- [8] J. M. Kim, K. B. Yoon, J. Y. Sim, S. Y. Park, Y. S. Sin, "Real-time fall detection using YOLOv3 algorithm", *Proceedings of 2020 conference on Communications and Information Sciences*, KICS, Yongpyeong, Korea, pp.1294-1295, Aug. 2020.
- [9] J. -H. Kim, J. -H. Choi, Y. -H. Park, A. Nasridinov "Abnormal situation detection on surveillance video using object detection and action recognition", *Journal of Korea Multimedia Society*, Vol.24, No.2, pp.186-198, Feb. 2021.
DOI: <https://doi.org/10.9717/kmms.2020.24.2.18>
- [10] J. -H. Song, J. -H. Bai, S. -H. Park, "Developing a school safety walking monitoring system applied YOLO and DeepSORT model", *Journal of Fisheries and Marine Sciences Education*, Vol.35, No.2, pp.301-314, June 2023.
DOI: <https://doi.org/10.13000/JFMSE.2023.4.35.2.301>
- [11] J. Kim, S. Park, C. -H. Hong, "A study on falling detection of workers in the underground utility tunnel using dual deep learning techniques", *Journal of The Korean Society of Disaster Information*, Vol.19, No.3, pp.498-509, Sep. 2023.
DOI: <https://doi.org/10.15683/kosdi.2023.9.30.498>
- [12] S. Youm, C. Kim, S. Choi, Y. -S. Kang, "Development of a methodology to predict and monitor emergency situations of the elderly based on object detection", *Multimedia Tools and Applications*, Vol.78, pp.5427-5444, March 2023.
DOI: <https://doi.org/10.1007/s11042-018-6660-7>
- [13] I. Yang, W. H. Jeon, "Development of performance improvement method of deep learning based automatic road dangerous object detection system", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.23, No.11, pp.8-18, Nov. 2022.
DOI: <https://doi.org/10.5762/KAIS.2022.23.11.8>

김 정 수(Jeongsoo Kim)

[정회원]



- 2016년 2월 : 연세대학교 일반대학원 토목환경공학과 (공학박사)
- 2016년 3월 ~ 2018년 9월 : 한국건설기술연구원 박사후연구원
- 2018년 10월 ~ 현재 : 한국건설기술연구원 전임연구원

<관심분야>

전산역학, 지하구조물, 딥러닝, 수치해석