

<http://dx.doi.org/10.17703/JCCT.2024.10.6.649>

JCCT 2024-11-79

이미지 딥러닝을 이용한 건설현장 안전모 미착용 인식에 대한 연구

A Study on Recognition of not Wearing Safety Helmets at Construction Sites using Image Deep Learning

유동연*, 오태근**

Dongyeon, Yoo*, Taekeun Oh**

요약 건설현장에 개인보호 미착용에 따른 사고가 자주 발생한다. 산업재해 통계에 따르면 건설업 근로자의 주요 사망원인 중 하나가 안전모 미착용이고 현장에서는 안전모 착용을 위해 다양한 노력을 하고 있으나 이를 잘 지키지 못하는 것이 현실이다. 보호구 착용 관리를 효율적으로 하는 방법 중에는 CCTV 기반 이미지 딥러닝 알고리즘을 활용하는 것이 있으며 건설현장 이미지 데이터를 기반으로 Convolutional Neural Network(CNN) 알고리즘을 적용하여 안전모의 착용 및 미착용 인식을 통해 작업자의 안전모 착용 준수를 확인할 수 있는 다양한 방법들이 제안되고 있다. 이는 건설현장 근로자의 안전모 착용 준수 여부를 확인하고 안전 수칙 위반 근로자를 식별하여 안전관리에 도움을 줄 수 있다. 이에 본 연구에서 우리는 YouOnly Look Once(YOLO)를 활용하여 근로자의 안전모 미착용 여부를 효과적으로 판단할 수 있는 모델 구축 방법론을 제시하였다. CIS 공개 데이터 셋을 적용하여 안전모 미착용 근로자 검출을 위한 최적의 변수와 데이터 비율을 선정하고 다양한 학습과 검증을 통해 도출된 분석결과를 통해 얻은 안전모 착용 및 미착용 데이터 분포에 따른 판별 정확도 및 오차 분석을 통해 현장 활용성을 높이고자 하였다.

주요어 : 안전모, 객체인식, 이미지 딥러닝

Abstract The most common accidents at construction sites are caused by not wearing personal protective equipment. According to industrial accidents statistics, one of the main death causes is not wearing safety helmets, and although various efforts have been made to have workers wear safety helmets. One of the ways to manage the wearing of safety helmet is to use CCTV-based image deep learning algorithm, so various methods have been proposed to confirm workers' compliance with wearing safety helmets by the CNN. This can help with safety management by checking whether workers are complying with wearing safety helmets and identifying workers who violate safety rules. In this study, we proposed a model construction methodology that can effectively determine whether workers are not wearing safety helmets by utilizing YOLOv9. With CIS data set, we selected the optimal variables and data ratio for detecting workers not wearing safety helmets, and we aimed to increase field usability by analyzing the discrimination accuracy and error according to the distribution of data on wearing or not wearing safety helmets derived through various learning and verification.

Key words : Safety helmet, Object detection, Image deep learning

*정회원, 인천대학교 산업경영공학과 학사과정 (제1저자)

**정회원, 인천대학교 안전공학과 교수 (교신저자)

접수일: 2024년 8월 27일, 수정완료일: 2024년 9월 25일

게재확정일: 2024년 11월 5일

Received: August 27, 2024 / Revised: September 25, 2024

Accepted: November 5, 2024

**Corresponding Author: thoh@inu.ac.kr

Dept. of safety engineering, Incheon National Univ, Korea

I. 서 론

국내 건설현장 3곳 중 1곳은 안전모, 안전벨트 등 개인보호구를 제대로 착용하지 않고 있다는 정부 조사 결과가 있으며 건설업 사망사고 주요 원인 중 하나가 보호구 미착용으로 발표되고 있다[1]. 건설현장 근로자 보호구 착용 여부를 효율적으로 관리할 방법의 하나는 현장 CCTV 영상을 실시간으로 분석하는 AI 기술 기반 안전관리이다. 이에, 현장 관리감독자에 의한 수동적 안전관리가 아닌 빠른 속도로 발전하고 있는 이미지 딥러닝 기반 이미지 및 동영상 프레임 인식 기술에 의한 안전관리 방식들이 개발되고 있다[2-4]. 본 연구에서는 건설현장 이미지 데이터를 기반으로 CNN 딥러닝 알고리즘을 적용하여 안전모의 착용 및 미착용 인식을 통해 작업자의 안전모 착용 준수를 확인할 수 있는 실용적 방법론을 제안한다. 이는 건설현장 근로자의 안전모 착용 준수 여부를 확인하고 안전 수칙 위반 근로자를 식별하여 안전관리에 도움을 줄 수 있다[5-6].

이에 본 연구에서는 그림 1과 같이 건설현장 근로자의 안전모 착용 및 미착용 데이터셋에 대해 YOLOv9 모델을 적용하여 안전모 착용 여부를 분석하고, 적절한 데이터 구성과 정확도 특성을 분석함으로써 향후 건설현장에서 보호구 미착용을 정확하게 확인할 수 있는 통제관리시스템 구축에 있어 기초자료를 제공하고자 한다.

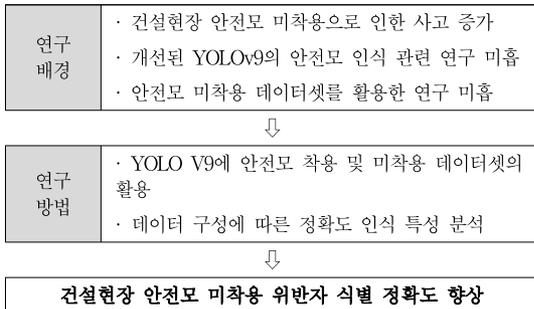


그림 1. 연구 개요
Figure 1. Structure of research

II. 이론 고찰

전 세계적으로 반도체와 메모리의 빠른 속도의 발달을 통해 빅데이터에 대한 그래픽처리장치(GPU)를 이용

한 딥러닝이 가능해졌고 이를 활용한 객체 인식 기술의 정확도는 높아지고 있다. 딥러닝이라 불리는 심층신경망은 객체 검출에 이미지 처리에 적합한 CNN이 대표적이다. 초기의 딥러닝 모델은 이미지 구역 기반의 R-CNN(Region-based CNN)으로 시작되었으나 실시간 동영상을 적합한 속도와 정확도로 처리하기 위해서 YOLO로 개선되었다[7-8]. 차이점은 R-CNN은 전체 이미지를 유지하면서 관심 영역을 나누지만, YOLO는 CNN으로 처리하기 전에 이미지를 구분하여 주어진 그리드로 분할하고, 그리드 내 객체탐지의 정확도를 계산한 신뢰도를 이용한다. 이런 과정을 거쳐 가장 높은 신뢰도를 가지는 경계 상자를 검출하는 것이다.

과거 연구에서는 주로 안전모 착용 데이터셋을 기반으로 하여 안전모 착용이 인식되지 않으면 미착용으로 판단하였으나 그럴 때 일반모자나 머리 스타일에 따라 안전모로 오인식할 수 있다. 따라서 오인식률을 낮추고 미착용 인식을 구체화하기 위해서는 안전모 미착용 데이터 셋을 필수적으로 사용하여야 하고 정확도를 높이기 위해서는 최소한의 데이터 수를 확보해야 한다.

건설현장의 환경은 복잡하고 많은 객체와 배경을 가지고 있어 객체 인식에 있어서 많은 오류가 발생한다. 따라서 본 연구에서는 그림 2와 같이 그라데이션 정보(programmable gradient information, PGD)와 일반 효율적 계층 집합 네트워크(generalized efficient layer aggregation network, GELAN)를 도입하여 실시간 객체 감지에 상당한 향상을 이룬 YOLOv9을 통해 안전모 착용 여부를 탐지하고자 한다. 이전의 YOLO 모델들에서는 입력치과 목표치의 맵핑과정에서 입력치가 특징 추출과 공간 변형을 거치고 나면 막대한 양의 정보가 손실되는 것을 방지하지 못했다. 이러한 손실 때문에 훈련 모델이 부정확한 예측을 하게 된다. 그러나 YOLOv9에서는 PCI와 GLEAN를 통해 예측하는 데 충분한 양의 정보를 획득할 수 있으며, 새로운 경량 네트워크 구조를 통해 훨씬 좋은 결과를 출력할 수 있도록 한다[9-10]. 이러한 특징들을 이용하여 건설 현장 내에서의 복잡하고 많은 객체가 모여있음에도 안전모 착용 및 미착용을 보다 잘 탐지할 수 있다.

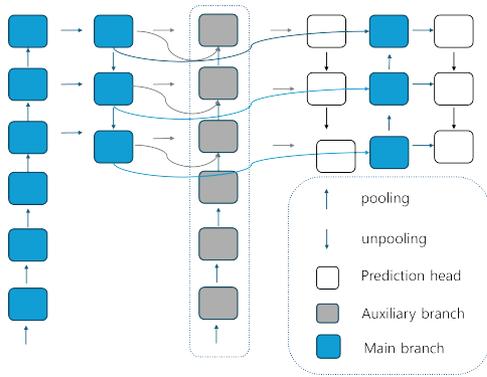


그림 2. YOLOv9 PGI 방법
 Figure 2. YOLOv9 PGI method

III. 연구 방법

3.1. 실험 및 컴퓨터 환경

건설현장 근로자 안전모 미착용 감지를 위해 사용된 YOLOv9 모델 알고리즘의 성능 평가에 사용된 컴퓨터 환경과 객체 인식에 대해 사용된 데이터 및 소프트웨어의 정보는 표 1과 같으며 운영체제(operating system, OS)는 유닉스를 기반으로 개발된 우분투를 사용하고 중앙처리장치(central processing unit, CPU)는 AMD 기반으로 그래픽 처리장치(graphic processing unit, GPU)는 지포스 RTX 4090을 4개 장착하여 최적의 성능을 유지하고자 하였다.

표 1. 실험 및 컴퓨터 환경
 Table 1. Experiment and computer environment

Division	Computer specification
OS (operating system)	Ubuntu 22.04 LTS
CPU (entral processing unit)	AMD Ryzen Threadripper PRO 5995WX
GPU (raphic processing unit)	NVIDIA GeForce RTX 4090
RAM (random access memory)	503GB
CUDA (computed unified device architecture)	11.8
CUDNN (CUDA deep neural network)	8.40
Image size	640
Software	python/pytorch
Model	YOLOv9C

3.2 데이터 셋 (Date set)

안전모 착용 여부 학습데이터는 건설현장 근로자의 보호구 착용 및 미착용 사진을 충분히 제공하고 있는 CIS-Dataset(<https://github.com/XZ-YAN/CIS-Dataset>)을 사용하였다. 데이터 셋은 helmet, no-helmet의 2개 class만 사용하였다. CIS에서의 helmet 이미지 데이터는 대략 6만여 장 정도이고 no-helmet 이미지는 대략 6천 장 정도이다. 이 중 가장 분별이 잘 되는 데이터를 helmet 3000장, no-helmet 3000장 선정하여 학습에 활용하였다. 데이터 셋은 학습 데이터를 기준으로 학습 : 검증 : 테스트 데이터 비율을 8:1:1로 하였다. 그림 3과 같이 클래스는 helmet과 no-helmet 2개로 라벨링 하였으며, 영상 데이터는 표 2와 같이 helmet과 no-helmet의 비율에 따라 데이터 셋을 6개로 구분하였으며 최소 500장 이상 학습을 통해 과적합을 방지하고자 하였다.

라벨링 할 때 Ground Truth의 영역은 안전모를 착용하지 않고 들고 있는 상황을 helmet으로 탐지하는 것을 방지하기 위해, 그림 3과 같이 사람 전체 영역에 대한 경계 상자로 설정하였다.

또한, 객체 검출에 대한 모델 알고리즘 성능 평가를 위해서 혼동행렬(Confusion Matrix), PR(Precision-Recall) 그래프, 정밀도(Precision), 재현율(Recall) 등의 정량적 지표를 이용하였다.

표 2. 실험에 적용한 데이터셋
 Table 2. Dataset applied to the experiment

Case	helmet	no-helmet	Ratio
1	3000	0	-
2	2500	500	5:1
3	2000	1000	2:1
4	1500	1500	1:1
5	1000	2000	1:2
6	0	3000	-



그림 3. 데이터셋에 대한 라벨링 결과
Figure 3. Labeling results for the dataset.

IV. 연구결과 및 분석

4.1 개인보호구 검출

안전모 착용 및 미착용 탐지를 위한 최적 성능의 모델을 생성하기 위해 YOLOv9c 모델에 대해 다양한 배치(batch)와 에포크(epoch)를 적용하여 최적의 값을 찾는 과정을 거쳤다. 그림 4는 배치 16, 에포크 100에 대한 PR 그래프이며 과적합 없이 정확한 예측을 보여주고 있다.

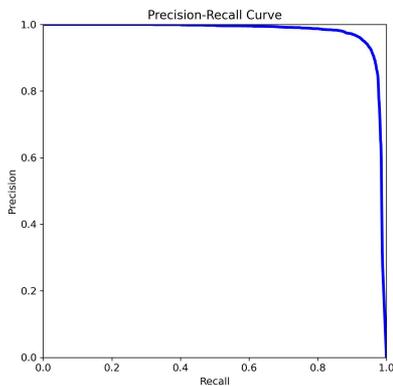


그림 4. 배치 16 및 에포크 100에 대한 PR 그래프
Figure 4. PR graph for batch 16 and epoch 100.

표 3은 Case 1~6에 대한 근로자의 안전모 착용 및 미착용을 검출한 결과를 보여주고 있으며 no-helmet 검출률은 0.76~0.88, helmet 검출률은 0.90~0.97, mAP는 0.910~0.976 이었다.

표 3. 케이스 별 모델 성능

Table 3. Model performance by Case

Case	helmet : no-helmet	precision	recall	mAp@0.5
1	-	0.941	0.99	0.976
2	5 : 1	0.955	0.96	0.910
3	2 : 1	0.925	0.98	0.937
4	1 : 1	0.938	0.97	0.924
5	1 : 2	0.939	0.97	0.937
6	-	0.946	0.97	0.961

케이스 별 정확도를 종합적으로 판단하기 위해 표 4과 같이 혼동 행렬의 TP, TN 값 위주로 판단하였고 FP, FN 값에 대해서도 비교·분석하였다. 전체적 결과는 표 5와 같이 Case 6을 제외하고 TP값이 92% 이상으로 유의미한 정확도를 나타내었다.

표 4. 혼동 행렬

Table 4 Confusion Matrix

	total	prediction	
		helmet	no-helmet
actual	helmet	TP	FN
	no-helmet	FP	TN

우선, Case 1과 같이 helmet 데이터만 사용한 경우에도 TP값이 0.97로 높은 정확도를 보인다는 점에서 YOLOv9 모델이 안전모 객체 인식을 잘하고 있다는 것을 알 수 있다. 그러나 현장 주요 관심사는 no-helmet을 정확하게 식별하여 안전 수칙 위반자를 구분하고 철저한 안전관리를 하는 것이다.

따라서 Case 1과 같이 helmet 데이터만을 사용할 경우, no-helmet을 인식하지 못하고 오인식이 발생할 가능성이 높다. 또한 Case 6과 같이 no-helmet 데이터만 이용한 때도 no-helmet의 탐지 정확도가 떨어지는 단점이 있다. 따라서 helmet과 no-helmet 데이터를 적절하게 조합하여 훈련을 시켜야 안전모 미착용 경우를 정확히 탐지하여 유용한 정보로 활용할 수 있다. 하지만, no-helmet 공개 데이터는 확보에 어려움이 따른다. 따라서 어느 정도까지 확보해야만 no-helmet를 비교적 정확하게 탐지할 수 있는지 검토가 필요하다.

표 5의 Case 3~5에서 보듯이 helmet, no-helmet 데이터 수가 1000장 이상인 경우 TP, TN 값이 85% 이상인 것을 확인할 수 있다. '안전모 미착용 감지'에 목적을 두는 경우 TN을 중점적으로 보고 FP, FN를 가능한

낮추면 좋으며 FN 보다는 FP를 낮출 수 있는 경우가 목적을 충족시키는 데 합리적이라고 판단된다.

정리하면 helmet 데이터만 이용할 때 no-helmet의 경우를 인식할 수 없고 no-helmet 데이터만 쓸 때도 helmet 경우를 인식하지 못하고 오히려 정확도도 떨어지므로 helmet 데이터 1,000장 이상, no-helmet 데이터 1,000장 이상 사용하여야만 비교적 정확한 인식률이 확보될 것으로 보인다.

표 5. Case 별 혼동행렬 값 비교

Table 5. Comparison of confusion matrix values by Case

Case	TP	TN	FP	FN
1	0.97	-	0.03	-
2	0.95	0.76	0.01	0.14
3	0.95	0.88	0.03	0.04
4	0.94	0.85	0.02	0.09
5	0.92	0.88	0.04	0.05
6	-	0.76	-	0.24

4.2 Case별 이미지 비교분석

Case별 결과의 정성적 비교·분석을 위해서 FP, FN에 해당하는 경우의 이미지 판독을 실시하였으며 대표적인 경우는 그림 5 및 6과 같다.

그림 5과 같이 helmet 데이터만 사용할 때 안전모 미착용의 다양한 경우를 탐지하기 어렵고 일반모자인 경우에도 안전모로 잡는 경우가 존재한다. 또한, 그림 6의 Case 2와 같이 no-helmet의 비율이 지나치게 작은 경우 helmet을 착용하지 않은 사람도 helmet이라고 판단하는 경우가 있다. 따라서 Case 3과 같이 no-helmet 데이터가 1000장 이상은 확보되어야 no-helmet을 어느 정도 예측할 수 있다. FP, FN 수치의 비율을 최대한 낮추기 위해서는 helmet 데이터 수를 증가시키면 된다. 본 사례 연구 결과를 건설현장에 활용한다면 효과적인 안전모 미착용 검출이 가능할 것이다.



그림 5. Case 1 : 일반모자를 helmet으로 인식
 Figure 5. Case 1: Recognizing a regular hat as a helmet



그림 6. Case 2 : no-helmet을 helmet으로 인식
 Figure 6. Case 2: no-helmet is recognized as a helmet

한편, Case 3인 경우에는 그림 7과 같이 helmet을 손에 들고 있는 경우에도 no-helmet으로 정확하게 예측이 가능하다.



그림 7. Case 3 : helmet을 손에 들고 있는 경우 no-helmet으로 정확하게 인식
 Figure 7. Case 3 : Holding a helmet in hand is recognized as no-helmet.

V. 결론

본 연구에서 우리는 이미지 딥러닝 기반 건설현장 근로자 안전모 미착용 인식을 위한 방법론을 제안하였다. 안전모 미착용 탐지를 위해 YOLOv9c 모델을 적용하였고, 안전모 착용 및 미착용 이미지가 있는 CIS 공개 데이터 셋을 적용하여 안전모 미착용 근로자 검출을 위한 최적의 변수와 데이터 비율을 선정하였다. 다양한 학습과 검증을 통해 도출된 사례해석 결과를 통해 얻은 정보는 다음과 같다.

- helmet과 no-helmet의 비율이 편향적이지 않은 경우가 가장 객체를 잘 감지하였다.
- 정석적 이미지를 분석해 본 결과 Case 1과 2인 경우 안전모를 착용하지 않은 경우 helmet이라고 표

시되거나 no-helmet를 인식하지 못하는 오류가 발생했다.

- 혼동 매트릭스와 이미지 비교분석을 통해 Case 3-5인 경우 FP, FN을 최소화하면서 객체를 잘 탐지했다고 볼 수 있다.
- helmet 및 no-helmet 데이터 수가 1000장 이상 적용하여 모델을 구축할 때 헬멧과 미헬멧을 적절하게 감지하는 것을 알 수 있다.

우리는 본 연구 결과를 통해 이미지 또는 동영상 프레임에 대한 실시간 모니터링과 연계한 딥러닝 기반 작업자 안전모 미착용 탐지를 통해 작업자 안전사고를 감소시킬 수 있는 효과적인 방안이 마련될 것이라 기대한다. 또한, 향후 연구 결과를 활용하여 통제관리시스템과 연계하는 방안이 필요할 것이다.

References

- [1] Ministry of Employment and Labor, Industrial Accident Statistics, 2023
- [2] K.W. Lee, S.W. Lee, H.S. Kim, H.K. Jung, "Deep Learning-Based Worker Personal Protective Equipment and Face Identification System," *Journal of Knowledge Information Technology and Systems*, Vol. 17, No. 3, pp. 385 - 394, 2022. DOI : 10.34163/jkits.2022.17.3.001
- [3] T.-W. Kang, B.-K. Kim, and Y.-S. Jung, "Deep learning platform architecture for monitoring image-based real-time construction site equipment and worker," *Journal of KIBIM*, vol. 11, no. 2, pp. 24 - 32, 2021. DOI: 10.13161/kibim.2021.11.2.024
- [4] K.W. Lee, H.S. Kim, H.K. Jung, and S.W. Lee, "Deep Learning-Based Worker Personal Protective Equipment and Face Identification System," *Journal of Knowledge Information Technology and Systems*, vol. 17, no. 3, pp. 385 - 394, 2022 DOI: 10.34163/jkits.2022.17.3.001
- [5] M. H. Kim et al., "Accident Detection System for Construction Sites Using Multiple Cameras and Object Detection," *The journal of the convergence on culture technology*, Vol. 9, No. 5, pp. 605 - 611, 2023. DOI : 10.17703/JCCT.2023.9.5.60
- [6] Y.G. Yoon and T.K. 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 : 10.17703/JCCT.2022.8.5.735
- [7] B. Zhang, C.-F. Sun, S.-Q. Fang, Y.-H. Zhao, and S. Su, "Workshop Safety Helmet Wearing Detection Model Based on SCM-YOLO," *Sensors*, Vol. 22, No. 17. MDPI AG, p. 6702, 2022. DOI : 10.3390/s22176702
- [8] S. B. Jang, "A Deep Learning Performance Comparison of R and Tensorflow," *The journal of the convergence on culture technology*, Vol. 9, No. 4, pp. 487 - 494, 2023. DOI : 10.17703/JCC T.2023.9.4.487
- [9] C. Chien, R. Ju, K. Chou, and J. Chiang, "YOLOv9 for fracture detection in pediatric wrist trauma X ray images," *Electronics Letters*, vol. 60, no. 11. Institution of Engineering and Technology (IET), 2024. DOI : 10.1049/ell2.13248
- [10] R. Ye, G. Shao, Q. Gao, H. Zhang, and T. Li, "CR-YOLOv9: Improved YOLOv9 Multi-Stage Strawberry Fruit Maturity Detection Application Integrated with CRNET," *Foods*, Vol. 13, No. 16. MDPI AG, p. 2571, 2024. DOI : 10.3390/foods13162571

※ 이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임.
(No.2021R111A2050912)