

An Auditable Real-Time Safety Compliance Detection System using YOLOv8l and Web Integration

Umair Aziz

umairaziz682@gmail.com

Department of Computer Science & Information Technology, Mirpur University of Science & Technology (MUST), Mirpur, Pakistan

Javeria Rehman

javeriarehman100@gmail.com

Department of Computer Science & Information Technology, Mirpur University of Science & Technology (MUST), Mirpur, Pakistan

Maryam Issa

maryamissa124@gmail.com

Department of Computer Science & Information Technology, Mirpur University of Science & Technology (MUST), Mirpur, Pakistan

Muhammad Ahsan Ali

mahsanali189@gmail.com

Department of Computer Science & Information Technology, Mirpur University of Science & Technology (MUST), Mirpur, Pakistan

Asma Jabeen

asma.jabeen@must.edu.pk

Department of Computer Science & Information Technology, Mirpur University of Science & Technology (MUST), Mirpur, Pakistan

Corresponding Author: Umair Aziz umairaziz682@gmail.com

Received: 04-02-2026

Revised: 27-02-2026

Accepted: 05-03-2026

Published: 21-03-2026

ABSTRACT

Conventional approaches to enforcing Personal Protective Equipment (PPE) safety policies in workplaces depend heavily on human observation. This manual routine is inherently variable, demands significant manpower, and is prone to oversight. This paper describes a real time Safety Compliance Detection System that combines a tailored YOLOv8l object detector, a Euclidean-distance tracker, and a Flask web application with a PostgreSQL database. The detection algorithm was trained on a curated collection of 3,235 industrial site images with 14,341 annotations across seven PPE classes: helmet, vest, mask, boots, gloves, goggles, and face shield. A EuclideanDistTracker with a 100-pixel threshold tracked workers across video frames, assigning unique identifiers and enabling continuous compliance monitoring without duplicate alerts. The 7-class network produced an overall mean average precision (*mAP@0.5*) of 0.928 and maximum F1 score of 0.94 at confidence threshold of 0.548 on the test dataset. Moreover, the inference pipeline achieved 21.5 FPS on a Tesla T4 GPU. These outcomes establish the architecture as a highly accurate, continuous, and traceable tool for managing occupational hazards.

Keywords: PPE detection, YOLOv8, deep learning, object tracking, workplace safety, Flask, PostgreSQL

INTRODUCTION

Safety in the workplace is a serious issue. In 2021, there were 1,465 industrial accidents and 200 deaths in Kazakhstan [1]. The Occupational Safety and Health Administration (OSHA) requires the use of PPE to prevent accidents at work [2]. According to the International Labour Organization (ILO), 2.3 million deaths from injuries around the world [3] and the U.S. Bureau of Labor Statistics reported 2.5 million

nonfatal injuries in private industries [4]. In the healthcare sector, frequent use of PPE like masks is also crucial for infection control [5].

While PPE is effective in preventing head injuries (combined protective helmets, safety helmets), respiratory injuries (mask, respirator), and injuries from vision impairment (safety goggles, goggles) its use is difficult to track. Physical safety checks and rounds by the managers are still being used in most facilities. These checks are time consuming, subjective and not very scalable especially in large facilities because they are based on manual observations [6]. So little help is offered by sensor-based solutions like the RFID tags [7] or Wi-Fi portals [8]. Simply they detect the equipment, but not the proper use. Also, workers usually don't feel comfortable carrying additional sensors.

Computer vision, combined with deep neural networks, offers a scalable approach to this problem. Historical approaches had relied largely on color and gradient histograms, but had problems with lighting variability [9]. Other methods, such as region-based models R-CNN [10], its Fast [11] and Faster [12] versions, were highly accurate but not real-time for videos. The advent of single-stage detectors (mainly YOLO [13] and SSD [14]) marked a move towards real-time performance, with further variants of YOLO (Yolo3 [15], to Yolo4 [16]) steadily increasing the balance between accuracy and speed.

Yet, most prototypes lack characteristics necessary for industry deployment. The lack of a system for tracking workers in progress, secure logging, and role-based access control means safety auditors are denied continuous records. Thus, we seek to address this gap by developing highly accurate PPE detection system for 7 classes in real-time and integrate it into a secure web platform to deliver automated and audit-ready compliance histories for safety governance of organizations.

Section II reviews the recent methods for PPE monitoring and tracking. The proposed methodology, data pre-processing, and the system design, which is a three-tier architecture with YOLOv8l inference engine and Flask web interface, are discussed in Section III. In Section IV, we report the model's efficacy, and its deployability. Section V concludes study and suggests directions for future research.

RELATED WORK

In recent years, automatic PPE monitoring has seen a shift in technologies. Initially, they mostly relied on sensors. Barro-Torres et al. proposed a real-time monitoring system that uses wireless wearable sensors and radio-frequency (RF) signals to monitor PPE [8]. This method was further extended by Kelm et al., who developed a mobile, passive RFID portal to quickly scan safety gear upon entry to construction sites [7]. These approaches were not susceptible to visual obstructions, but only checked for the presence of tagged objects within their reader's range, and incurred substantial maintenance costs and potential worker discomfort.

Seo et al. offered an overview of vision-based health and monitoring in worker sites. They found that the most prominent weakness of conventional approaches is being susceptible to lighting and background changes [6]. Rubaiyat et al. proposed an automatic helmet detection system using a convolutional neural network (CNN)-pipeline in a controlled environment [9]. Wu and Zhao also used deep convolutional neural networks for helmet detection [17]. Pradana et al. implemented a CNN-based classifier to categorize combinations of safety equipment (such as glasses and helmets); but they only experimented with a controlled indoor setting, with consistent backgrounds [18].

An improvement in detection accuracy was achieved by region-based deep detectors. Girshick et al. proposed R-CNN, a detector that automatically produced proposals for object detection using selective search and classified these proposals using a CNN [10]. Girshick followed this up with Fast R-CNN, which allowed the sharing of convolutional features between proposals [11]. Ren et al. subsequently proposed Faster R-CNN, which introduced a Region Proposal Network (RPN) to generate proposals [12]. Fang et al. developed a Faster R-CNN to identify non-hardhat use in long-range videos [19]. Akbarzadeh et al. presented a network-within-a-network structure to detect PPE on large construction

sites via frame segmentation [20]. However, two-stage detectors could not achieve high enough frame rates for continuous video surveillance from multiple cameras.

To overcome latency issues, researchers came up with single-stage detectors. Liu et al. proposed the Single Shot MultiBox Detector (SSD) that made predictions for object classes and offsets over multi-scale features [14]. Redmon et al. had the YOLO model, which framed the detection problem as regression [13]. The model evolved rapidly: YOLOv3 ran multi-scale predictions through a deeper network [15], while YOLOv4 used mosaic data augmentation and other changes to improve accuracy [16]. Xie et al. adopted CNN models for real-time hard-hat detection [21]. Protik et al. proposed a real-time PPE detection system with TensorFlow and YOLOv4 [22]. Ma et al. used a lightweight version of YOLOv4 for the detection of multi-class PPE [23]. Wang et al. presented fast YOLOv5-based detectors of personal protection equipment in the construction environment and the CHV dataset [24]. Fan et al. proposed an ensemble deep learning model for helmet detection [25]. Loey et al. used YOLOv2 and ResNet-50 for detecting helmets for face masks during the COVID-19 outbreak [26]. Ding et al. used CNNs and long short-term memory (LSTM) networks for unsafe act detection [27]. Ferdous and Ahsan developed a PPE detector for construction using YOLO [28]. Alateeq et al. used YOLOv5 to detect hazards on Saudi construction sites [29].

Recent developments also include anchor-free methods and graph neural networks (GNNs). Wang proposed an anchor-free convolutional neural network which did not use pre-defined anchors to facilitate predictions about box sizes [30]. In particular scenarios, Zhao and Barati implemented GNNs for fast PPE detection and evaluation in electrical substations [31], while Marquez-Sanchez et al. combined smart wearables and edge computing for safety monitoring [32].

Although these studies achieve significant performance gains in speed and accuracy, they mostly overlook practical considerations. Few solutions track workers across video frames, therefore leading to duplicate recording of the same hazardous activities. These practical challenges present the main motivation for the current research.

METHODOLOGY

Study Design

We took a deep learning model and created a detection model to detect seven different types of PPE from a selected dataset. We tested the model's effectiveness in practical applications by embedding it in a complex enterprise system that involves computer vision and databases. This allows us to assess real-time compliance using the model.

Dataset Preparation

We used Roboflow platform to prepare the dataset which contains 3,235 images in YOLO format resized to 640x640 pixels. This has 14,341 labels of 7 classes of bounding boxes. We then divided the images into 2,264 for training, 647 for validation and 324 for testing. We unified the labelling and resolution in pre-processing. We employed run-time augmentation (horizontal flips, brightness variations) during training to improve generalization. The class imbalance for each split is shown in Table I.

System Design and Workflow

The system has a three-tier architecture that integrates computer vision and web enterprise management. Our system has an architecture that works from live video streams to the user interface, as illustrated in Fig. 1.

Table 1. Dataset Class Distribution across Splits

Class	Train (2,264)	Val (647)	Test (324)
helmet	2,231	765	300
gloves	2,144	645	309
boots	1,987	596	269
vest	1,797	580	258
mask	1,050	314	162
goggles	547	149	88
face_shield	111	25	14
Total	9,967	3,074	1,300

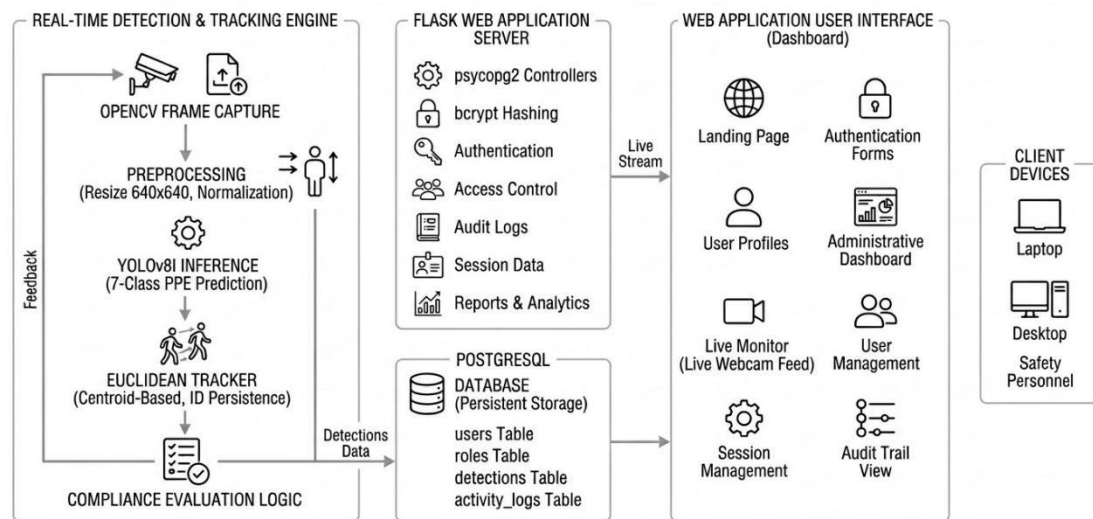


Fig. 1. System Design and Workflow

We use OpenCV to capture video or webcam feed. We also standardize and resize the image (640 x 640 pixes) for the inference engine (YOLOv8l) to detect the seven pieces of PPE. A EuclideanDistTracker for object tracking, using centroid-based tracking and ID maintenance was applied. Next, the compliance checker verifies the safety items' compliance.

The live detection results are reported to the Flask-based web application server which serves as the system's core. Here, psycpg2 controller facilitates interaction with the database, bcrypt password hashing is used for security and role-based credentials are implemented. PostgreSQL database is used to store the roles of users, detection logs and audit logs. Finally, the Flask server streams the video and data to the web-based dashboard from which safety managers can monitor activities and produce reports, on any type of device.

RESULTS AND DISCUSSION

Quantitative Results

The model was evaluated across all seven PPE classes on the held-out test set. It achieved a global network mAP@0.5 of 0.928, an average precision of 0.899, and an overall recall of 0.904. End-to-end processing operated at 47.4 ms per frame, equating to 21.5 FPS on a Tesla T4 GPU. The model reached an optimal operational balance with a peak F1-score of 0.94 at a 0.548 confidence threshold. Detailed class metrics are presented in Table II.

In industrial safety, it is important to achieve a high precision (no false alarms) and recall (no missed violations). The F1-score serves as the harmonic mean of these two metrics. As illustrated in Fig. 2, the model reached an optimal operational balance with a peak global F1-score of 0.94 at a 0.548 confidence threshold.

The F1-confidence curve shows a steep rise and a flat plateau ($F1 > 0.90$) over a large threshold range (0.40-0.65). This suggests that our model is stable and not extremely sensitive to small variations in confidence scores.

Table 2: Per-Class Detection Metrics on Test Set

Class	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Helmet	0.961	0.980	0.987	0.816
Boots	0.981	0.996	0.989	0.746
Mask	0.961	0.969	0.983	0.716
Gloves	0.947	0.929	0.956	0.738
Vest	0.900	0.903	0.927	0.738
Face Shield	0.716	0.722	0.836	0.485
Goggles	0.825	0.830	0.818	0.475
Average	0.899	0.904	0.928	0.673

A threshold of 0.55 ensures that the Flask web application will be able to successfully eliminate low-scoring background information, while accurately notifying administrators about PPE usage non-compliance.

Precision-Recall Characteristics

The Area Under the Curve (AUC) in the Precision-Recall (PR) graph also offers some insight into class-specific performance. As indicated in Fig. 3, high-frequency, structurally rigid classes like boots, helmets (0.987 mAP@0.5) and masks (0.983mAP@0.5) have curves that closely approach the top-right corner. This suggests that they are almost always detected, regardless of environmental factors.

In contrast, curves for face shields and goggles are less steep. These differences are due to class imbalance in the dataset and difficulty in computer vision of distinguishing transparent objects against varied backgrounds.

Confusion Matrix

A confusion matrix was created to assess the model's classification performance and its robustness to background clutter (Fig. 4).

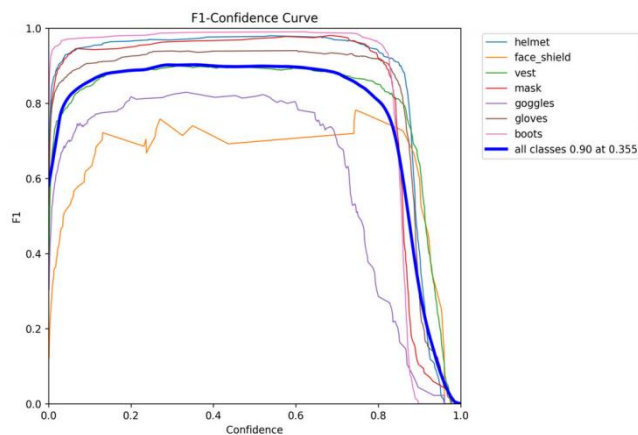


Fig. 2. F1-Confidence curve indicating a peak score of 0.94 at 0.548 threshold.

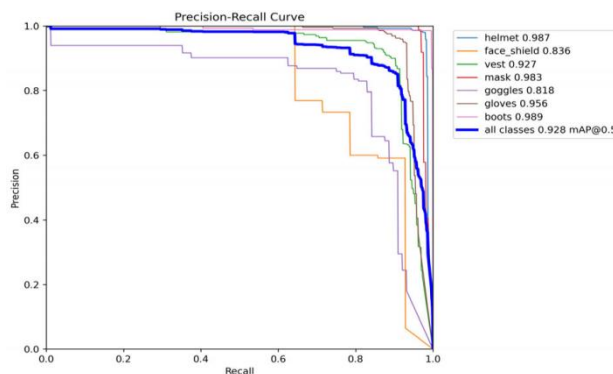


Fig. 3. Precision-Recall curve of 7-classes.

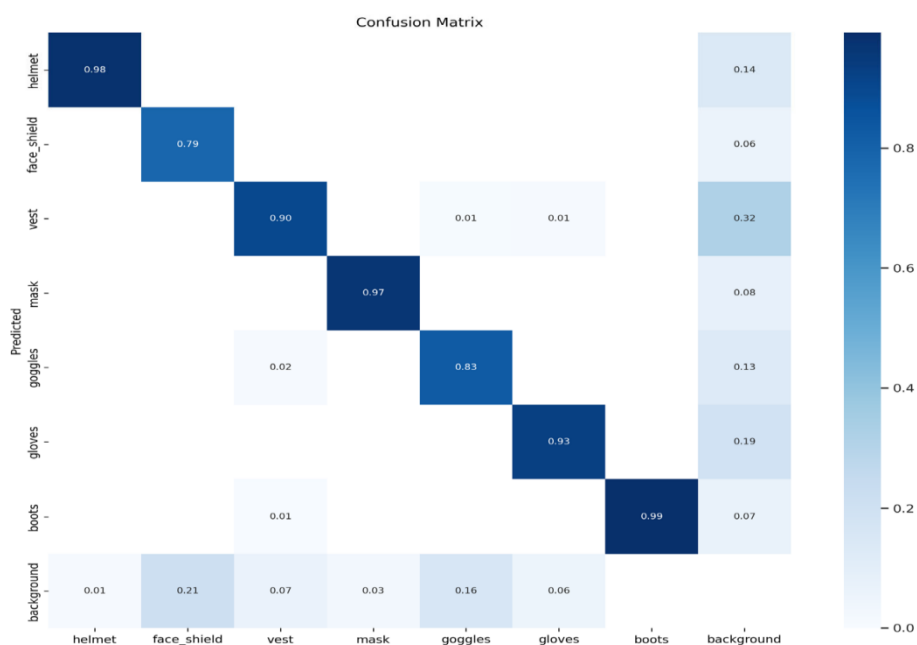


Fig. 4. Confusion matrix for detection model

The diagonal of the matrix is heavily weighted, proving that the model correctly identifies true positives with little confusion among the classes. For example, the model should not confuse a 'helmet' and a 'mask', which confirms that the feature extraction layers of the YOLOv8l backbone are highly discriminative.

The bulk of the errors in the matrix are background related. Occasionally, PPE items were overlooked, due to strong occlusion or bad lighting in the images used. The EuclideanDistTracker can help resolve these issues. For instance, if the YOLO predicts a missed helmet, the ID system still tracks the worker's compliance. This prevents ambiguous warnings to appear on the compliance dashboard.

DISCUSSION

We found high generalization for worldwide protective equipment. We achieved close to 100% accuracy for boots, helmets and masks. This is probably due to their rigidity, shape, and texture, and their common presence in the data. Precision and recall for face shields and goggles were lower, consistent with the class imbalance of the annotations, and difficulty in protecting transparent surfaces against a dynamic background.

Beyond this, our system addressed significant problems encountered by previous work. A Flask server and PostgreSQL database leveraging the trained model transforms video into an enterprise-wide web application. Inference results are stored rapidly in the database and can be logged, tracked, and monitored using roles in enterprise safety programs.

CONCLUSION AND FUTURE PROSPECTS

This research has successfully developed a 7-class Safety Compliance Detection System with an overall test mAP@0.5 of 0.928. Through the combination of a fine-tuned YOLOv8l model with object tracking and a secure web interface, the system translates research into industry-relevant safety auditing. The authors plan to continue this work by balancing the dataset classes for under-represented items and porting the model to edge devices with state-of-the-art appearance-based trackers such as DeepSORT to improve tracking in crowded scenes.

ACKNOWLEDGMENT

We thank the Roboflow community for providing the datasets. We are also grateful to the developers of YOLOv8, Flask, PyTorch, and PostgreSQL for creating the tools that made this work possible.

REFERENCES

- Committee on Labour, Social Protection and Migration of Kazakhstan, "Workplace injuries in Kazakhstan 2021," Kursiv Media, Jan. 2022. [Online]. Available: <https://kz.kursiv.media/2022-01-29/v-kazahstane-rabochie-kalechatsya-na-proizvodstve-v-tri-raza-chasche-chem/>
- U.S. Occupational Safety and Health Administration, "Personal Protective Equipment," OSHA, 2024. [Online]. Available: <https://www.osha.gov/personal-protective-equipment>
- International Labour Organization, "Safety and Health at Work," ILO, 2023. [Online]. Available: <https://www.ilo.org/topics-and-sectors/safety-and-health-work>

- U.S. Bureau of Labor Statistics, "Employer-Reported Workplace Injuries and Illnesses – 2023-2024," *BLS News Release USDL-26-0101*, Jan. 2026. [Online]. Available: <https://www.bls.gov/news.release/pdf/osh.pdf>
- World Health Organization, "Rational use of personal protective equipment for COVID-19 and considerations during severe shortages," WHO Interim Guidance, 2020. [Online]. Available: <https://iris.who.int/handle/10665/338033>
- J.Seo, S. Han, S. Lee, and H. Kim, "Computer vision techniques for construction safety and health monitoring," *Adv. Eng. Informatics*, vol. 29, no. 2, pp. 239-251, 2015. [Online]. Available: <https://doi.org/10.1016/j.aei.2015.02.001>
- A. Kelm, L. Lauaat, A. Meins-Becker, D. Platz, M. J. Khazaee, A. M. Costin, M. Helmus, and J. Teizer, "Mobile passive Radio Frequency Identification (RFID) portal for automated and rapid control of Personal Protective Equipment (PPE) on construction sites," *Automat. Constr.*, vol. 36, pp. 38-52, 2013. [Online]. Available: <https://doi.org/10.1016/j.autcon.2013.08.009>
- S. Barro-Torres, T. M. Fernandez-Carames, H. J. Perez-Iglesias, and C. J. Escudero, "Real-time personal protective equipment monitoring system," *Comput. Commun.*, vol. 36, no. 1, pp. 42-50, 2012. [Online]. Available: <https://doi.org/10.1016/j.comcom.2012.01.005>
- A. H. Rubaiyat, T. T. Toma, M. Kalantari-Khandani, S. A. Rahman, L. Chen, Y. Ye, and C. S. Pan, "Automatic detection of helmet uses for construction safety," in *Proc. 2016 IEEE/WIC/ACM Int. Conf. Web Intelligence Workshops (WIW)*, pp. 135-142, Oct. 2016. [Online]. Available: <https://doi.org/10.1109/WIW.2016.045>
- R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 580-587, 2014. [Online]. Available: <https://doi.org/10.1109/CVPR.2014.81>
- R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 1440-1448, 2015. [Online]. Available: <https://doi.org/10.1109/ICCV.2015.169>
- S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137-1149, 2017. [Online]. Available: <https://doi.org/10.1109/TPAMI.2016.2577031>
- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 779-788, 2016. [Online]. Available: <https://doi.org/10.1109/CVPR.2016.91>
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Comput. Vis. ECCV 2016*, LNCS, vol. 9905, pp. 21-37, Springer, 2016. [Online]. Available: https://doi.org/10.1007/978-3-319-46448-0_2
- J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint, arXiv:1804.02767, 2018. [Online]. Available: <https://doi.org/10.48550/arXiv.1804.02767>

- A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint, arXiv:2004.10934, 2020. [Online]. Available: <https://doi.org/10.48550/arXiv.2004.10934>
- H. Wu and J. Zhao, "Automated visual helmet identification based on deep convolutional neural networks," in *Comput. Aided Chem. Eng.*, vol. 44, pp. 2299-2304, 2018. [Online]. Available: <https://doi.org/10.1016/B978-0-444-64241-7.50378-5>
- R. D. W. Pradana, R. Y. Adhitya, M. Syai'in, R. M. Sudibyo, D. R. A. Abiyoga, M. A. Jami'in, L. Subiyanto, B. Herijono, A. Wahidin, A. Budianto, and N. H. Rochiem, "MIdentification system of personal protective equipment using Convolutional Neural Network (CNN) method," in *Proc. 2019 Int. Symp. Electronics Smart Devices (ISESD)*, pp. 1-6, Oct. 2019. [Online]. Available: <https://doi.org/10.1109/ISESD.2019.8909629>
- Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T. M. Rose, and W. An, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Automat. Constr.*, vol. 85, pp. 1-9, 2018. [Online]. Available: <https://doi.org/10.1016/j.autcon.2017.09.018>
- M. Akbarzadeh, Z. Zhu, and A. Hammad, "Nested network for detecting PPE on large construction sites based on frame segmentation," in *Creative Construction e-Conf. 2020*, pp. 33-38, 2020. [Online]. Available: <https://doi.org/10.3311/CCC2020-006>
- Z. Xie, H. Liu, Z. Li, and Y. He, "A convolutional neural network based approach towards real-time hard hat detection," in *Proc. 2018 IEEE Int. Conf. Progress Informatics Comput. (PIC)*, pp. 430-434, Dec. 2018. [Online]. Available: <https://doi.org/10.1109/PIC.2018.8706269>
- A. A. Protik, A. H. Rafi, and S. Siddique, "Real-time personal protective equipment (PPE) detection using YOLOv4 and TensorFlow," in *Proc. 2021 IEEE Region 10 Symp. (TENSYP)*, pp. 1-6, Aug. 2021. [Online]. Available: <https://doi.org/10.1109/TENSYP52854.2021.9550808>
- L. Ma, X. Li, X. Dai, Z. Guan, and Y. Lu, "A combined detection algorithm for personal protective equipment based on lightweight YOLOv4 model," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1-11, 2022. [Online]. Available: <https://doi.org/10.1155/2022/3574588>
- Z. Wang, Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao, "Fast personal protective equipment detection for real construction sites using deep learning approaches," *Sensors*, vol. 21, no. 10, p. 3478, 2021. [Online]. Available: <https://doi.org/10.3390/s21103478>
- Z. Fan, C. Peng, L. Dai, F. Cao, J. Qi, and W. Hua, "A deep learning-based ensemble method for helmet-wearing detection," *PeerJ Comput. Sci.*, vol. 6, p. e311, 2020. [Online]. Available: <https://doi.org/10.7717/peerj-cs.311>
- M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustainable Cities Soc.*, vol. 65, p. 102600, 2021. [Online]. Available: <https://doi.org/10.1016/j.scs.2020.102600>
- L. Ding, W. Fang, H. Luo, P. E. Love, B. Zhong, and X. Ouyang, "A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory,"

Automat. Constr., vol. 86, pp. 118-124, 2018. [Online]. Available: <https://doi.org/10.1016/j.autcon.2017.11.002>

- M. Ferdous and S. M. M. Ahsan, "PPE detector: A YOLO-based architecture to detect personal protective equipment (PPE) for construction sites," *PeerJ Comput. Sci.*, vol. 8, p. e999, 2022. [Online]. Available: <https://doi.org/10.7717/peerj-cs.999>
- M. M. Alateeq, F. R. Pp, and M. A. Ali, "Construction site hazards identification using deep learning and computer vision," *Sustainability*, vol. 15, no. 3, p. 2358, 2023. [Online]. Available: <https://doi.org/10.3390/su15032358>
- H. Wang, "Detection of Personal Protective Equipment (PPE) using an Anchor Free-Convolutional Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 2, 2024. [Online]. Available: <https://doi.org/10.14569/IJACSA.2024.0150239>
- M. Zhao and M. Barati, "Substation safety awareness intelligent model: Fast personal protective equipment detection using GNN approach," *IEEE Trans. Ind. Appl.*, 2023. [Online]. Available: <https://doi.org/10.1109/TIA.2023.3234515>
- S. Marquez-Sanchez, I. Campero-Jurado, J. Herrera-Santos, S. Rodriguez, and J. M. Corchado, "Intelligent platform based on smart PPE for safety in workplaces," *Sensors*, vol. 21, no. 14, p. 4652, 2021. [Online]. Available: <https://doi.org/10.3390/s21144652>