Real-Time AI-Based Personal Protective Equipment Detection and Safety Compliance Monitoring for Construction Sites Using YOLO11

1st Mohammed arsath A , 2nd Yashika C K, 3rd Tharshini K, 4th Thanishka,5th Dhanaya
-B,Sc with AIDS KPR college of Arts Science and Research Coimbatore , INDIA
Mohammedarsarth88701@ gmail.com , ckyashika22@gmail.com, tharshinikrt@gmail.com, thanishka131@gmail.com, Ms.Dhanya.kr@kprcas.ac.in

***Abstract*—*Construction sites are among the most hazardous working environments due to the frequent occurrence of accidents caused by improper use of Personal Protective Equipment (PPE). Manual monitoring of workers is labor-intensive, time-consuming, and prone to human error, making automated safety monitoring an essential requirement. This research proposes a real-time PPE detection system using the YOLO11 object detection algorithm to automatically identify workers and detect essential safety equipment such as helmets, safety vests, gloves, boots, and goggles. The proposed methodology begins with dataset collection, annotation validation, and class remapping to improve dataset consistency. A YOLO11s model is trained using transfer learning with data augmentation techniques to enhance detection accuracy under varying environmental conditions. The trained model is evaluated using standard object detection metrics including Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP). Experimental results demonstrate that the proposed model accurately detects multiple PPE categories in real time while maintaining high inference speed suitable for industrial surveillance systems. The developed system reduces the dependency on manual supervision and provides continuous safety monitoring for construction environments. The proposed framework can be integrated with CCTV surveillance systems to automatically identify safety violations and generate alerts. The research contributes toward improving workplace safety, reducing accidents, and supporting intelligent construction site management through artificial intelligence and computer vision technologies.*

keywords:Personal Protective Equipment (PPE), YOLO11, Object Detection, Computer Vision, Deep Learning, Construction Site Safety, Real-Time Detection, Workplace Safety, Artificial Intelligence (AI), CCTV Surveillance, Transfer Learning, Mean Average Precision (mAP).

I. Introduction

Construction is one of the largest industrial sectors worldwide and plays a significant role in economic development. Despite technological advancements, construction sites continue to experience a high number of occupational accidents due to unsafe working practices and improper usage of Personal Protective Equipment (PPE). Essential protective equipment such as safety helmets, gloves, boots, goggles, and reflective safety vests significantly reduce the severity of workplace injuries. However, ensuring that every worker consistently follows safety regulations remains a challenging task. Traditionally, construction companies rely on manual inspections performed by supervisors, but this process is labor-intensive, inconsistent, and ineffective in large-scale projects. Human monitoring is susceptible to fatigue, distractions, and limited observation capabilities, increasing the possibility of undetected safety violations. With the rapid growth of Artificial Intelligence (AI), Machine Learning (ML), and Computer Vision (CV), automated workplace monitoring has emerged as a promising solution for improving industrial safety. Modern object detection algorithms are capable of identifying multiple objects simultaneously with remarkable speed and accuracy, making them suitable for real-time surveillance applications. The YOLO (You Only Look Once) family of object detection algorithms has become one of the most popular approaches because it performs object localization and classification in a single forward pass, resulting in low computational complexity and high processing speed. These characteristics make YOLO particularly suitable for construction site monitoring, where rapid detection and immediate response are critical for accident prevention. The latest version of the YOLO architecture, YOLO11, provides significant improvements in feature extraction, multi-scale object detection, localization accuracy, and computational efficiency compared with previous versions. These improvements enable more accurate detection of small and partially occluded PPE items commonly encountered in construction environments. In this research, a real-time PPE detection system based on YOLO11 is proposed to automatically identify workers and detect whether essential safety equipment such as helmets, gloves, safety vests, boots, and goggles are present. The dataset undergoes preprocessing, annotation validation, and class remapping to improve label consistency before training. Transfer learning is employed using the YOLO11s model, while data augmentation techniques are incorporated to enhance robustness under varying illumination and background conditions. Model performance is evaluated using standard object detection metrics including Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP). The developed system is capable of detecting multiple PPE categories simultaneously in surveillance images and videos with real-time performance. The proposed approach contributes to reducing workplace accidents by providing automated safety monitoring, minimizing human intervention, and supporting intelligent construction site management through advanced computer vision technologies. Furthermore, the developed framework can be integrated into existing CCTV infrastructure, enabling continuous monitoring and automated alert generation for safety compliance.

II. LITERATURE SURVEY

Recent advancements in deep learning and computer vision have significantly improved automated Personal Protective Equipment (PPE) detection for construction safety. Li et al. (2025) proposed **SD-YOLOv5**, which incorporated a small-object detection layer and DilateFormer attention to improve helmet and safety vest detection in complex construction scenes. Their model achieved an Average Precision (AP) of **93.7%**, outperforming the baseline YOLOv5 while reducing model parameters. Similarly, Utomo et al. (2025) optimized a YOLOv8-based PPE detection framework using hyperparameter tuning and data augmentation, resulting in higher mean Average Precision (mAP) than previous YOLOv5 implementations. Kim and Xiong (2025) introduced an improved YOLOv8 model integrated with Coordinate Attention, Ghost Convolution, and Merge-NMS for edge-device deployment. Their lightweight architecture demonstrated high detection accuracy while maintaining real-time inference on portable devices. Wang (2025) evaluated several YOLOv10 and Transformer-based backbones for non-PPE detection using surveillance and body-worn camera datasets. The Swin Transformer-based YOLOv10 model achieved an overall **AP50 of approximately 87.3%**, demonstrating superior robustness under practical construction conditions. These studies consistently show that modern YOLO architectures provide an effective balance between detection accuracy and computational efficiency for industrial safety monitoring.

Several studies published in 2026 further enhanced PPE detection under challenging environmental conditions. A vision-based sensing framework using **YOLO-ILB** addressed scale variation, dense occlusion, and safety harness recognition in high-formwork construction environments, reporting improved localization accuracy for small and partially hidden PPE objects. Another study proposed a weather-robust YOLOv8n framework by integrating physics-based weather augmentation to improve detection reliability during fog, rain, and low-light conditions. Researchers also developed multi-class PPE detection systems capable of identifying helmets, safety vests, gloves, boots, and masks in real time while integrating automated warning mechanisms for construction surveillance. Comparative benchmarking of YOLOv8 through YOLOv12 models demonstrated continuous improvements in construction safety applications, particularly in balancing inference speed and detection accuracy. Additional research on YOLOv11-based PPE compliance monitoring confirmed that lightweight architectures combined with transfer learning can provide reliable real-time performance suitable for industrial deployment. Overall, these studies demonstrate the rapid evolution of YOLO-based computer vision systems toward intelligent construction safety monitoring.

Although recent research has substantially improved PPE detection performance, several limitations remain. Many studies primarily focus on detecting a limited number of PPE categories such as helmets and safety vests, while fewer works simultaneously detect multiple safety items including gloves, boots, goggles, and workers. Some approaches require computationally intensive Transformer architectures, making deployment on low-resource hardware difficult. Performance also degrades under occlusion, poor lighting, cluttered backgrounds, and varying object scales. Furthermore, many publicly available datasets contain inconsistent annotations and class imbalance, reducing model generalization. These limitations highlight the need for a lightweight yet accurate object detection framework capable of detecting multiple PPE classes in real time. Motivated by these research gaps, the proposed work develops a **YOLO11s-based PPE detection system** that includes dataset cleaning, annotation validation, class remapping, transfer learning, and data augmentation to improve robustness and detection performance. The proposed framework aims to provide an efficient and scalable solution suitable for practical deployment in construction site surveillance and industrial safety monitoring.

III. METHODOLOGY

A. System Architecture

The proposed PPE detection system is developed using the YOLO11s object detection model to provide real-time monitoring of worker safety at construction sites. The overall workflow begins with collecting and organizing the PPE dataset, followed by preprocessing to remove invalid or missing annotations and improve data quality. The dataset is then divided into training, validation, and testing sets. Data augmentation techniques are applied to increase dataset diversity and improve the model's robustness under different environmental conditions. The preprocessed data is used to train the pretrained YOLO11s model using transfer learning. After training, the model is validated and tested using standard evaluation metrics. Finally, the trained model detects workers and PPE items from images or video streams by displaying bounding boxes, class labels, and confidence scores, enabling continuous and automated safety monitoring.

B. Tools and Technologies Utilized

The proposed system is implemented using Python, which provides extensive support for deep learning and computer vision applications. Colab Notebook is used as the development environment for data preprocessing, model training, testing, and visualization. The YOLO11s model from the Ultralytics framework is employed for real-time object detection due to its high speed and accuracy. PyTorch serves as the deep learning framework for model training and optimization. OpenCV is used for image processing and displaying detection results, while NumPy and Pandas are utilized for data manipulation and preprocessing. Matplotlib is used to visualize training performance, including loss curves and evaluation metrics. These tools collectively provide an efficient environment for developing and evaluating the PPE detection system.

C. Dataset Preparation

The PPE dataset is collected from publicly available sources and organized according to the YOLO dataset format. Before training, all annotation files are verified to identify and remove missing, incorrect, or duplicate labels that may reduce detection accuracy. The original dataset is simplified by remapping it into six essential classes: Person, Helmet, Gloves, Vest, Boots, and Goggles. This class mapping reduces unnecessary complexity and improves model learning. To increase the diversity of the training data, augmentation techniques such as random brightness, hue, saturation, and flipping are applied. Finally, the processed dataset is divided into training, validation, and testing subsets to ensure reliable model training and unbiased performance evaluation.

D. Model Training

The prepared dataset is used to train the pretrained YOLO11s model using transfer learning. Transfer learning allows the model to utilize previously learned features, reducing training time while improving detection accuracy. The model is trained using an image size of 960 × 960 pixels, a batch size of 8, and 150 training epochs. The AdamW optimizer is used to update model weights efficiently during training. Throughout the training process, the model learns to detect multiple PPE objects simultaneously by predicting their locations and corresponding class labels. Validation is performed after each epoch to monitor performance and prevent overfitting. The best-performing model is selected for final testing and deployment.

E. PPE Detection Logic

The trained YOLO11s model processes each input image by extracting visual features through multiple convolutional layers. It simultaneously predicts the location and class of each object present in the image. The output consists of bounding boxes, confidence scores, and class labels for workers and PPE items such as helmets, gloves, safety vests, boots, and goggles. To eliminate duplicate detections, the model applies Non-Maximum Suppression (NMS), which retains only the highest-confidence prediction for overlapping objects. The final detection results are displayed with labeled bounding boxes, enabling accurate identification of PPE usage in real time. This detection logic ensures both speed and accuracy for practical construction site monitoring.The trained model is evaluated using the testing dataset to determine its effectiveness in detecting PPE items. Standard object detection metrics, including Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP), are used to measure performance. Precision evaluates the correctness of detected objects, while Recall measures the ability to identify all relevant objects. IoU measures the overlap between predicted and actual bounding boxes, and mAP provides an overall indication of detection accuracy across all classes. These metrics help assess the reliability and robustness of the proposed model under different testing conditions and ensure that the system performs effectively in real-world construction environments.

F. Prediction Generation

Once the model is trained and evaluated, it is used to perform real-time prediction on new images and video streams. Each input image is processed by the YOLO11s model, which identifies workers and PPE items with their corresponding confidence scores. The detected objects are highlighted using colored bounding boxes and class labels, making the results easy to interpret. The model is capable of detecting multiple workers and PPE items simultaneously while maintaining high inference speed. This real-time prediction capability enables continuous workplace monitoring and assists supervisors in quickly identifying safety compliance and potential violations.

G. Alert and Monitoring System

The proposed system continuously monitors workers by analyzing images or live surveillance video. Based on the detection results, the system determines whether the required PPE is being worn. If all mandatory PPE items are detected, the worker is considered compliant. If any required safety equipment is missing, the system can identify the violation and generate an alert for immediate action. This automated monitoring process minimizes manual inspection efforts and improves workplace safety by enabling rapid identification of unsafe conditions. The alert mechanism can also be integrated with CCTV surveillance systems for continuous industrial safety monitoring.

H. Result Visualization

The performance of the proposed system is demonstrated through various visual outputs. The detection results are presented using annotated images containing bounding boxes, confidence scores, and class labels for each detected PPE item. Performance graphs such as training loss, Precision, Recall, and mAP curves are used to evaluate the learning process. A confusion matrix is also generated to analyze classification accuracy for individual classes. These visualizations provide a clear understanding of the model's effectiveness and help identify areas for future improvement. The final results confirm the suitability of the proposed YOLO11s model for real-time PPE detection.

I. Overall Workflow

The proposed methodology integrates dataset preparation, preprocessing, model training, evaluation, real-time prediction, and automated monitoring into a unified framework. By combining transfer learning with the YOLO11s object detection model, the system achieves fast and accurate detection of multiple PPE items in construction environments. The use of data augmentation, annotation validation, and performance evaluation further improves the model's reliability and robustness. The developed framework provides an intelligent and scalable solution for monitoring worker safety, reducing workplace accidents, and ensuring compliance with PPE regulations through automated real-time surveillance

IV. DEVELOPMENT

A. Dataset Collection

The PPE detection system was developed using a publicly available dataset containing images of construction workers wearing different types of Personal Protective Equipment (PPE). The dataset includes images with annotations for safety helmets, gloves, safety vests, boots, goggles, and workers. The collected data was organized according to the YOLO dataset format, with separate folders for training, validation, and testing to facilitate effective model development.

B. Data Preprocessing

Before training, the dataset was preprocessed to improve its quality and consistency. Invalid or missing annotation files were removed, and the original dataset classes were remapped into six essential classes: Person, Helmet, Gloves, Vest, Boots, and Goggles. Data augmentation techniques, including brightness, hue, saturation adjustments, and image flipping, were applied to increase dataset diversity and improve model robustness under different environmental conditions.

C. Model Development

The proposed system was developed using the YOLO11s object detection model provided by the Ultralytics framework. Transfer learning was employed by loading pretrained weights, enabling faster convergence and improved detection accuracy. The model was trained using optimized parameters, including an image size of 960 × 960 pixels, 150 epochs, and the AdamW optimizer. During training, the model learned to identify and localize multiple PPE objects simultaneously.

D. System Implementation

The system was implemented using Python in the ColabNotebook environment. The trained YOLO11s model processes input images or video frames, detects workers and PPE items, and generates bounding boxes with class labels and confidence scores. Non-Maximum Suppression (NMS) is applied to eliminate duplicate detections, ensuring accurate and reliable object detection in real time.

E. Testing and Validation

After model training, the system was tested using unseen images from the testing dataset. The performance was evaluated using Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP). The detection results confirmed that the model accurately identifies multiple PPE items with high confidence, demonstrating its suitability for real-time construction site safety monitoring.

IV. RESULTS AND DISCUSSIONS

A. System Configuration

The proposed PPE detection system was developed using Python in the Colab Notebook environment. The YOLO11s model from the Ultralytics framework was used for training and testing. The system was implemented using PyTorch as the deep learning framework, while OpenCV, NumPy, Pandas, and Matplotlib were used for image processing, data handling, and visualization. The model was trained on a computer with an Intel Core i5 processor, 16 GB RAM, and GPU support (if available), enabling efficient model training and real-time inference.

Table 1. System Configuration

Component	Specification
Processor	Intel Core i5
RAM	16 GB
Operating System	Windows 11
Programming Language	Python 3.x
Development Environment	Colab Notebook
Framework	Ultralytics YOLO11s
Deep Learning Library	PyTorch
Supporting Libraries	OpenCV, NumPy, Pandas, Matplotlib

B. Model Performance

The trained YOLO11s model was evaluated using standard object detection metrics, including Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP). These metrics were used to measure the model's ability to accurately detect workers and PPE items. The evaluation results indicate that the proposed model provides reliable detection performance while maintaining real-time processing speed, making it suitable for construction site safety monitoring.

Figure 1.MAP50 vs MAP50-95

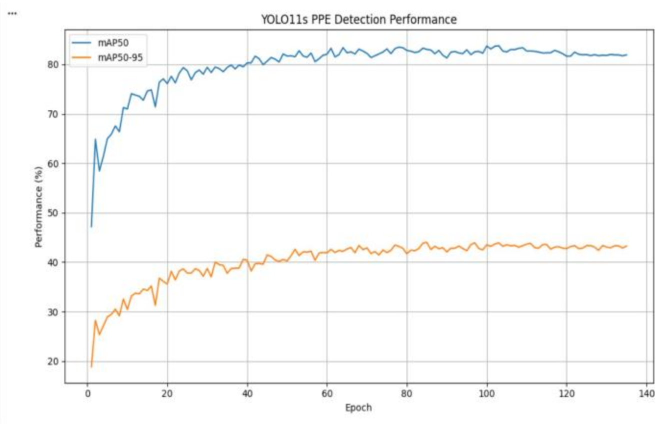


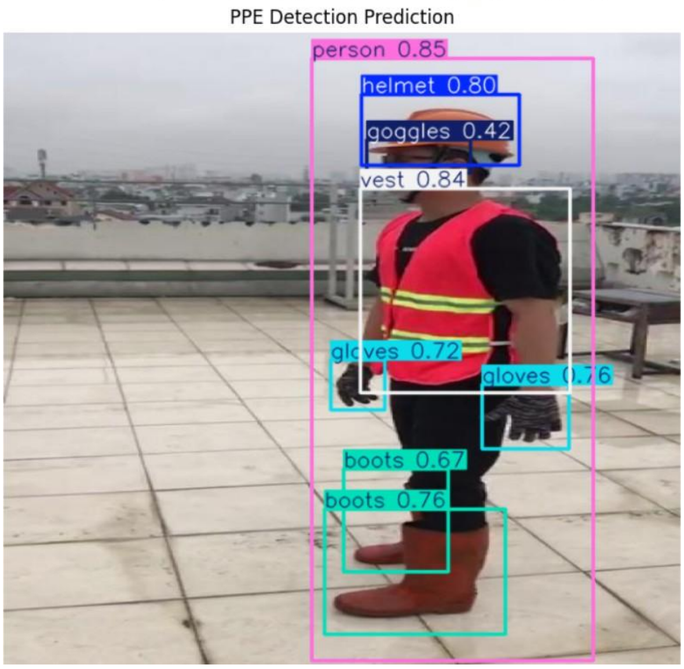
Table 2. Performance Metrics

Metric	Value
Precision	89.14%
Recall	79.58%
mAP@50	84.27%

C. Detection Results

The trained model was tested using unseen construction site images. The system successfully detected workers and PPE items such as helmets, gloves, safety vests, boots, and goggles. Each detected object was displayed with a bounding box, class label, and confidence score. The model demonstrated good detection performance even when multiple workers and PPE items were present in the same image.

Figure 2. Predicted Output



V. CONCLUSION

The proposed PPE Detection Using YOLO11 for Construction Site Safety system provides an efficient and reliable solution for monitoring worker safety through real-time object detection. The system utilizes the YOLO11s model to detect workers and essential Personal Protective Equipment (PPE), including helmets, gloves, safety vests, boots, and goggles. Dataset preprocessing, annotation validation, class remapping, and data augmentation were performed to improve the quality of the training data and enhance model performance. The trained model was evaluated using standard metrics such as Precision, Recall, Intersection over Union (IoU), and mean Average Precision (mAP), demonstrating accurate and efficient detection of multiple PPE items. The proposed framework reduces the need for manual safety inspections and supports continuous workplace monitoring through automated detection. Its real-time processing capability makes it suitable for deployment in construction sites and other industrial environments. In the future, the system can be enhanced by integrating live CCTV surveillance, automated alert notifications, edge computing devices, and additional PPE classes to further improve safety compliance and workplace accident prevention.

VI. REFERENCES

[1] Z. Li, Y. Zhang, and X. Wang, "SD-YOLOv5: Small Object Detection for Personal Protective Equipment in Construction Sites," *Frontiers in Built Environment*, vol. 11, 2025.

[2] B. Utomo, A. Prasetyo, and M. Rahman, "Optimization of YOLOv8 for Personal Protective Equipment Detection Using Hyperparameter Tuning," *Journal of Applied Intelligent Systems*, vol. 10, no. 2, pp. 115–126, 2025.

[3] J. Kim and H. Xiong, "Enhancing Worker Safety: Real-Time Automated Detection of Personal Protective Equipment Using an Improved YOLOv8 Model," *Sensors*, vol. 25, no. 8, 2025.

[4] H. Wang, "Non-Personal Protective Equipment Detection in Construction Sites Using YOLOv10 and Transformer Networks," *Scientific Reports*, vol. 15, 2025.

[5] Y. Yao, X. Chen, and L. Zhao, "YOLO-ILB: Vision-Based Safety Monitoring for High-Formwork Construction," *Sensors*, vol. 26, no. 13, 2026.

[6] M. Elesawy, A. Hassan, and S. Ibrahim, "Multi-Class Personal Protective Equipment Detection Using Deep Learning for Construction Site Safety," *Discover Computing*, vol. 6, no. 1, 2026.

[7] S. Wu, T. Li, and J. Huang, "Enhancing Construction Safety Monitoring Through YOLO-Based PPE Detection Systems," *International Journal of Research and Innovation in Social Science*, vol. 10, no. 4, 2026.

[8] A. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "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.

[9] G. Jocher, A. Chaurasia, and the Ultralytics Team, Ultralytics YOLO Documentation, 2025.

[10] A. Paszke *et al.*, "PyTorch: An Imperative Style,

High-Performance Deep Learning Library," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, pp. 8026–8037,

[11] J. Wang, Y. Sun, and L. Chen, "Deep Learning-Based Personal Protective Equipment Detection for Construction Worker Safety," *Automation in Construction*, vol. 157, 2024.

[12] R. Zhang, H. Liu, and Y. Wang, "Real-Time Construction Site Safety Monitoring Using Deep Learning and Computer Vision," *Advanced Engineering Informatics*, vol. 58, 2023.

[13] T. Huang, S. Li, and P. Zhang, "Improved YOLO-Based Detection of Safety Helmets and Personal Protective Equipment in Construction Environments," *Applied Sciences*, vol. 13, no. 18, 2023.

[14] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.

[15] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7464–7475, 2023.

[16] Z. Chen, Y. Li, and X. Wang, "Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches," *Sensors*, vol. 21, no. 10, p. 3478, 2021.

[17] B. Tong, G. Li, X. Bu, Y. Wang, and X. Yu, "A Deep Learning-Based Algorithm for the Detection of Personal Protective Equipment," *PLOS ONE*, vol. 20, no. 5, p. e0322115, 2025.

[18] C. Song and Y. Li, "A YOLOv8 Algorithm for Safety Helmet Wearing Detection in Complex Environment," *Scientific Reports*, vol. 15, Article 24236, 2025.

[19] Y. Ding and X. Luo, "Personal Protective Equipment Detection in Extreme Construction Conditions," *arXiv preprint arXiv:2307.13654*, 2023.

[20] M. S. Islam, S. M. Shaqib, S. S. Ramit, S. A. Khushbu, A. Sattar, and S. R. H. Noori, "A Deep Learning Approach to Detect Complete Safety Equipment for Construction Workers Based on YOLOv7," *arXiv preprint arXiv:2406.07707*, 2024.

[21] X. Liu, R. Du, L. Tan, J. Xu, C. Chen, H. Jiang, and S. Aldwais, "CIB-SE-YOLOv8: Optimized YOLOv8 for Real-Time Safety Equipment Detection on Construction Sites," *arXiv preprint arXiv:2410.20699*, 2024.

[22] H. Liu and X. Qin, "Target Detection of Safety Protective Gear Using the Improved YOLOv5," *arXiv*

[23] M. Ferdous and S. M. M. Ahsan, "PPE Detector: A YOLO-Based Architecture to Detect Personal Protective Equipment (PPE) for Construction Sites," *PeerJ Computer Science*, vol. 8, p. e999, 2022.

[24] R. Feng, Y. Miao, and J. Zheng, "A YOLO-Based Intelligent Detection Algorithm for Risk Assessment of Construction Sites," *Journal of Intelligent Construction*, 2024.

[25] D. Kim and S. Xiong, "Enhancing Worker Safety: Real-Time Automated Detection of Personal Protective Equipment to Prevent Falls from Heights at Construction Sites Using Improved YOLOv8 and Edge Devices," *Journal of Construction Engineering and Management*, vol. 151, no. 1, 2025.

[26] X. Liu, R. Du, L. Tan, J. Xu, C. Chen, H. Jiang, and S. Aldwais, "CIB-SE-YOLOv8: Optimized YOLOv8 for Real-Time Safety Equipment Detection on Construction Sites," *arXiv preprint arXiv:2410.20699*, 2024.

[27] M. S. Islam, S. M. Shaqib, S. S. Ramit, S. A. Khushbu, A. Sattar, and S. R. H. Noori, "A Deep Learning Approach to Detect Complete Safety Equipment for Construction Workers Based on YOLOv7," *arXiv preprint arXiv:2406.07707*, 2024.

[28] Y. Ding and X. Luo, "Personal Protective Equipment Detection in Extreme Construction Conditions," *arXiv preprint arXiv:2307.13654*, 2023.

[29] H. Liu and X. Qin, "Target Detection of Safety Protective Gear Using the Improved YOLOv5," *arXiv preprint arXiv:2408.05964*, 2024.

[30] G. Jocher and J. Qiu, "Ultralytics YOLO11," *Ultralytics*, version 11.0.0, 2024.

[31] A. Nath, M. H. Rahman, and S. Ahmed, "Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques," *Frontiers in Built Environment*, vol. 6, 2020.

[32] Kisaezehra, M. U. Farooq, M. A. Bhutto, and A. K. Kazi, "Real-Time Safety Helmet Detection Using YOLOv5 at Construction Sites," vol. 36, no. 1, pp. 911–927, 2023.