
Academia Open



By Universitas Muhammadiyah Sidoarjo

Table Of Contents

Journal Cover	1
Author[s] Statement	3
Editorial Team	4
Article information	5
Check this article update (crossmark)	5
Check this article impact	5
Cite this article.....	5
Title page	6
Article Title	6
Author information	6
Abstract	6
Article content	7

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Academia Open

Vol. 11 No. 1 (2026): June
DOI: 10.21070/acopen.11.2026.13886

EDITORIAL TEAM

Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

Academia Open

Vol. 11 No. 1 (2026): June
DOI: 10.21070/acopen.11.2026.13886

Article information

Check this article update (crossmark)



Check this article impact (*)



Save this article to Mendeley



(*) Time for indexing process is various, depends on indexing database platform

Automated PPE Compliance Verification Using YOLOv11l Spatial Logic: Verifikasi Kepatuhan APD Otomatis Menggunakan YOLOv11l dan Logika Spasial

Muhammad Minhaj Effendi, effendiminhaj@gmail.com (*)

Department of Electronics and Informatics, State University of Malang, Indonesia

Siti Sendari, siti.sendari.ft@um.ac.id

Department of Electronics and Informatics, State University of Malang, Indonesia

(*) Corresponding author

Abstract

General Background: Monitoring personal protective equipment (PPE) usage is a critical component of occupational health and safety (OHS) in construction, yet manual inspection remains inconsistent and prone to error. **Specific Background:** Recent advances in computer vision, particularly YOLO-based object detection, have improved PPE detection accuracy in complex environments. **Knowledge Gap:** However, existing approaches primarily detect PPE presence without verifying its correct usage or associating it with individual workers, leading to inaccurate compliance interpretation. **Aims:** This study develops an automated PPE compliance verification system using YOLOv11l combined with spatial association logic to assess PPE completeness and anatomical correctness at the individual worker level. **Results:** The system was trained on 2,788 construction images and achieved high performance with mAP@50 of 0.979, precision of 0.976, recall of 0.954, and peak F1-score of 0.97, while demonstrating accurate classification across PPE categories including helmets, vests, and shoes. **Novelty:** The integration of zone-based spatial verification enables validation of PPE placement within anatomically defined regions, addressing the limitation of detection-only systems. **Implications:** This approach supports objective, continuous, and reliable safety auditing in construction environments, offering a scalable alternative to manual OHS monitoring.

Highlights

- Multi-class detection identifies workers and safety equipment with high accuracy
- Region-based validation distinguishes proper gear usage from misplacement
- System classifies compliance status through structured decision logic

Keywords

Automated PPE Verification; Construction Safety; Deep Learning; Spatial Association Logic; YOLOv11l

Published date: 2026-04-19

INTRODUCTION

The construction industry remains one of the most hazardous sectors globally, with many accidents caused by unsafe worker behavior and complex site environments. According to the International Labour Organization, millions of occupational accidents occur annually, with a substantial portion attributed to the lack of proper safety measures. In Indonesia, the number of workplace accidents has shown a worrying upward trend, necessitating stricter enforcement of Occupational Health and Safety (OHS) regulations. Personal Protective Equipment (PPE) specifically safety helmets, high-visibility vests, and safety shoes serves as the last line of defense in minimizing injury severity. While regulations such as the Ministry of Manpower Regulation No. 5 of 2018 mandate PPE usage, non-compliance remains a pervasive issue due to limitations in supervision.

Traditional monitoring methods rely heavily on manual inspections by safety officers. This approach is limited by field of view, officer fatigue, and the inability to continuously monitor large areas. Consequently, violations are often missed due to sporadic inspections [6]. This situation has encouraged the use of computer vision and deep learning to automate safety monitoring [7].

Earlier methods, such as HOG, are less effective in dynamic environments, while CNN based deep learning models have proven more accurate and reliable [4]. Object detectors consist of two stage (Faster R-CNN) and on stage (YOLO, SSD) methods. Faster R-CNN is less suitable for real time because of its high computational cost, while YOLO has become the industry standard because it balances speed and accuracy. Review by [8] shows that the latest version of YOLO is increasing effective in detecting small objects, such as shoes and gloves.

Recent studies have adopted newer YOLO iterations to improve performance. Introduced YOLOv7 with trainable bag-of-freebies, setting a new benchmark for real-time detectors. The YOLO model continues to be developed, for example, YOLOv5, with its attention mechanism that improves helmet detection accuracy under occlusion [9], and a modified YOLOv8 that is more effective in dealing with complex backgrounds in industrial environments [10].

However, a gap remains between object detection and its presence verification. Most systems only detect the presence of PPE, such as helmets, without analyzing their relationship to the worker [11],[12],[13]. This results in the "separated object" problem, where PPE is detected but not associated with an individual. Furthermore, studies that examine the overall PPE adequacy for each worker are still limited. The regularity of complete PPE sets (helmets, vests, and boots) is often overlooked for individual workers.

The challenge lies not only in detection but also in spatially associating PPE items with workers to accurately assess compliance. Despite the advancement of deep learning based PPE detection, most studies still view it as general object detection, simply recognizing items like helmets or vests without ensuring correct and complete use by each worker. As a result, detected PPE objects are often not spatially associated with specific individuals, leading to false compliance interpretations.

To clarify this research gap, Table 2 summarizes representative PPE detection studies and highlights their limitations in terms of completeness verification and spatial association. This comparison demonstrates the necessity of a verification-oriented framework rather than detection-only solutions.

Table 1. Comparison of PPE Detection and Compliance Verification Approaches

Study	Method	Classes Detected	Completeness Verification with Spatial Association Logic
Nath et al. (2020)	YOLOv3	Helmet, Vest	No
An et al. (2023)	Improved YOLOv5	Helmet	No
Songn et al. (2024)	Improved YOLOv8	Helmet	No
Samperante et al. (2025)	YOLOv8	Helmet, Vest, Shoes	No
Proposed Method	YOLOv11l + Logic	Helmet, Vest, Shoes	Yes

Note: successfully detected shoes but treated them as independent objects without binding them to the worker ID for verification.

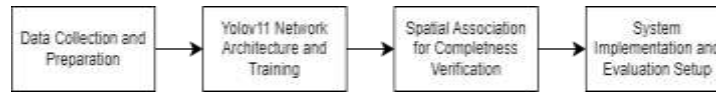
Moreover, existing detection systems fail to account for the anatomical correctness of PPE placement. A helmet detected in the frame does not guarantee it is worn on the worker's head it could be carried in hand or placed on a table. Similarly, a vest might be draped over equipment rather than worn on the torso. Spatial reasoning limitations can lead to verification errors, as PPE cannot simply be detected; it must also meet standards.

This research proposes a YOLOv11l based system with spatial association logic linking PPE to workers to improve the accuracy of individual assessments. Its contributions include: (1) reliable multi class detection for small and obstructed objects; (2) zone based verification that ensures the completeness and accuracy of PPE positioning; and (3) comprehensive performance evaluation to ensure system reliability for automated audits.

METHOD

This study uses a systematic deep learning development flow, including four main stages: data collection, training models, implementing verification logic, and evaluating systems, as shown in Figure 1.

Figure 1 . Research Flow Diagram



Data Collection and Preparation

The dataset was curated to represent realistic construction environments, containing 2,788 images. Data sources included controlled photo acquisition simulating various worker postures (standing, walking, squatting) and diverse lighting conditions. The dataset was annotated using the Roboflow platform, defining four distinct classes:

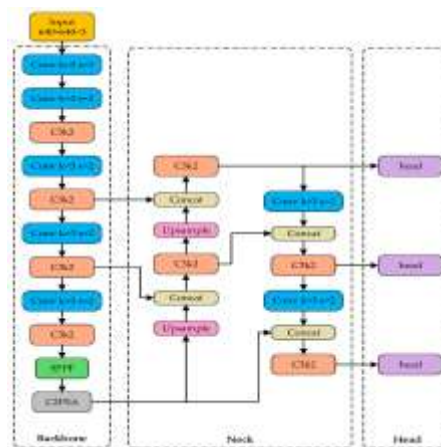
1. **Person:** The entire body of the worker.
2. **Safety Helmet:** Head protection gear.
3. **Safety Vest:** High-visibility bodywear.
4. **Safety Shoes:** Protective footwear.

To enhance model robustness, data augmentation techniques were applied, including random rotation ($\pm 15^\circ$), brightness adjustment ($\pm 25\%$), and horizontal flipping. The dataset was partitioned into Training (70%), Validation (20%), and Testing (10%).

YOLOv11 Network Architecture and Training

We selected the YOLOv11 Large (YOLOv11L) model for its superior feature extraction capabilities. YOLOv11 is an Ultralytics based experimental architecture with a CSP and PANet backbone that is capable of preserving feature details for detecting small objects such as safety shoes [16]. The model was configured with an input resolution of 640×640 pixels, a batch size of 16, and trained for 100 epochs using the Stochastic Gradient Descent (SGD) optimizer.

Figure 2 . Backbone YOLOv11 Architecture



The training loss function consists of three components: Box Loss, Class Loss, and Distribution Focal Loss. The total loss is defined as:

$$L_{total} = \lambda_{box}L_{box} + \lambda_{cls}L_{cls} + \lambda_{dfl}L_{dfl}$$

Where lambda represents the hyperparameters weighting each loss component. Box loss utilizes Ciou (Complete Intersection over Union) to ensure accurate localization, which is crucial for our spatial association logic.

Spatial Association for Completeness Verification

This system excels because it uses zone based verification logic, dividing the worker's bounding box into three anatomical regions to ensure standardized PPE placement, going beyond simple global IoU based object detection.

Regional Zone Partitioning

Academia Open

Vol. 11 No. 1 (2026): June
DOI: 10.21070/acopen.11.2026.13886

For each detected Person bounding box $B_P = (x_p, y_p, w_p, h_p)$, where h_p is the height of the person, the system defines three distinct zones:

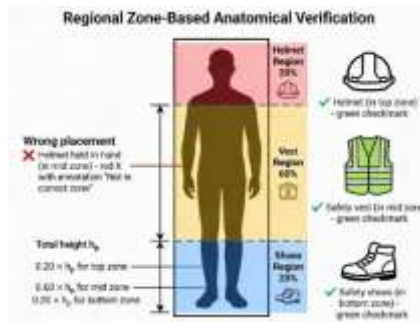
$$Z_{top} = (x_p, y_p, w_p, 0.20 \times h_p)$$

$$Z_{mid} = (x_p, y_p + 0.20 \times h_p, w_p, 0.60 \times h_p)$$

1. Top Zone (Helmet Region): Occupies the upper 20% of the person's bounding box, representing the head area.
2. Mid Zone (Vest Region): Occupies 60% of the bounding box, spanning from shoulder to waist, where high-visibility vests are typically worn.
3. Bottom Zone (Shoes Region): Occupies the lower 20% of the bounding box, corresponding to the foot area.

$$Z_{bottom} = (x_p, y_p + 0.80 \times h_p, w_p, 0.20 \times h_p)$$

Figure 3 . Regional Zone Partitioning for Anatomical PPE Verification



This partitioning is grounded in anthropometric principles and reflects the physical requirements of OHS regulations, ensuring that PPE detection is not only spatially associated but also positionally valid.

Zone-Specific IoU Computation

For each detected PPE item, the system computes the Intersection over Union (IoU) specifically with its corresponding anatomical zone:

$$\text{IoU}_{\text{zone}}(B_{\text{PPE}}, Z_{\text{region}}) = \text{Area}(B_{\text{PPE}} \cap Z_{\text{region}}) / \text{Area}(B_{\text{PPE}} \cup Z_{\text{region}})$$
 The

PPE item is considered correctly associated with the worker only if:

1. Helmet: $\text{IoU}(B_{\text{helmet}}, Z_{\text{top}}) \geq T_{\text{helmet}}$
2. Vest: $\text{IoU}(B_{\text{vest}}, Z_{\text{mid}}) \geq T_{\text{vest}}$
3. Shoes: $\text{IoU}(B_{\text{shoes}}, Z_{\text{bottom}}) \geq T_{\text{shoes}}$

Where T_{helmet} , T_{vest} , T_{shoes} are empirically determined IoU thresholds. Based on validation experiments, we set $T_{\text{helmet}} = 0.5$, $T_{\text{vest}} = 0.4$, and $T_{\text{shoes}} = 0.3$, accounting for the varying sizes and occlusion patterns of each PPE type. Compliance Decision Logic

A worker is classified as "COMPLIANT (Safe)" if and only if all three conditions are satisfied simultaneously:

- * Status = COMPLIANT if $(\text{Helmet} \in Z_{\text{top}})$ AND $(\text{Vest} \in Z_{\text{mid}})$ AND $(\text{Shoes} \in Z_{\text{bottom}})$ *
- Status = NON-COMPLIANT otherwise

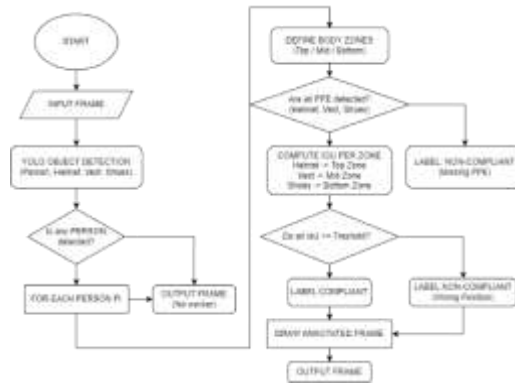
If any PPE item is missing or detected outside its designated zone, the worker is flagged as "NON-COMPLIANT (Unsafe)" with a visual annotation indicating which item is absent or misplaced.

Advantages of Regional Zone Verification

This approach offers several advantages over global IoU:

1. Prevents Misuse: PPE that is simply carried, not worn, cannot fool the system because the zones check for anatomical placement.
2. Reduces False Positives: Objects detected outside the zone, such as a helmet on a desk, are not associated with the worker.
3. OHS Complaint: Verification assesses actual compliance, ensuring PPE is worn correctly, not simply present in the worker's vicinity.
4. Pose Adaptive: Percentage based zones adjust to the pixel size of each worker supporting flow is shown in Figure 3.

Figure 4 . Flowchart



The overall verification process can be summarized as follows. For each input frame, the YOLOv11l detector first localizes all relevant objects (persons, helmets, vests, and shoes). For every detected person, the corresponding bounding box is partitioned into three body regions (top, middle, and bottom), and each PPE item is associated with its designated region using zone-specific IoU thresholds. A worker is labeled Compliant only if a helmet, vest, and shoes are all detected and correctly positioned in their respective zones; otherwise, the worker is labeled NonCompliant due to missing or misplaced PPE. Finally, the system renders an annotated frame with color-coded person bounding boxes (green for compliant, red for noncompliant) and textual feedback indicating the required PPE status.

Evaluation Metrics

System performance was evaluated using standard object detection metrics :

- Precision (P): Accuracy of positive predictions. ▪
- Recall (R): Ability to find all positive instances.
- Mean Average Precision (mAP@50): The average precision across all classes at an IoU threshold of 0.5. ▪
- F1-Score: The harmonic mean of Precision and Recall, providing a balanced view of performance.

Result and Discussion

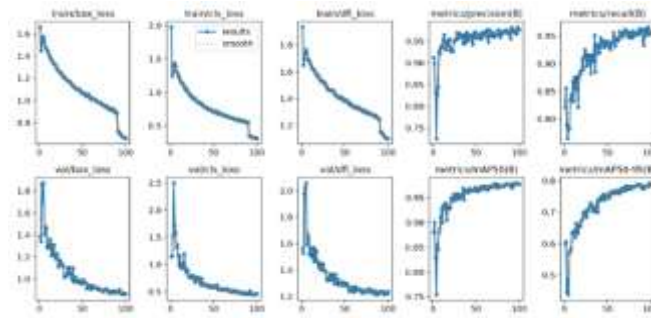
The experimental results validate the efficacy of the proposed system in both quantitative detection metrics and qualitative verification scenarios.

Result

Training Convergence

The training process over 100 epochs showed stable convergence. As illustrated in Figure 4, the box loss (*train/box_loss*) and classification loss (*train/cls_loss*) decreased consistently, while the precision and recall metrics improved, indicating that the model successfully learned the feature representations of the PPE items without overfitting.

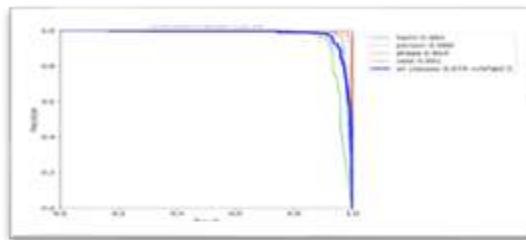
Figure 5 . Training Data Evaluation Results



Detection Performance

The quantitative evaluation reveals exceptional detection accuracy. The Precision-Recall (PR) curve in Figure 5 demonstrates an overall mAP@50 of 0.979.

Figure 6 . Precision-Recall Curve Achieving mAP@50 of 9,79



To visually demonstrate the model's detection capabilities on unseen data, Figure 6 displays sample predictions from the validation batch. The model accurately localizes multiple PPE items (helmets, vests, shoes) simultaneously across different worker poses and lighting conditions.

Figure 7 . Visual detection results on validation batch showing



precise bounding boxes for all classes

Detailed performance metrics for each class are summarized in Table 1. As shown, the model achieves high precision across all [ISSN 2714-7444 \(online\), https://acopen.umsida.ac.id](https://doi.org/10.21070/acopen.11.2026.13886), published by [Universitas Muhammadiyah Sidoarjo](https://www.muhammadiyah.ac.id)
Copyright © Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

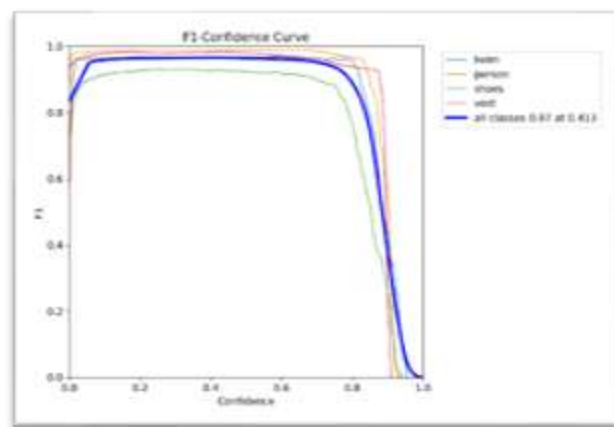
classes. The "Safety Vest" class performs best (mAP 0.992) due to its distinct visual features and high contrast colors. Even for the most challenging class, "Safety Shoes," the model maintains a robust mAP of 0.958, proving the effectiveness of the YOLOv11 architecture for small object detection.

Table 2. Validation Results

Class	Precision (P)	Recall (R)	mAP50	mAP50-95
Safety Helmet	0,981	0,963	0,985	0,812
Safety Vest	0,991	0,987	0,992	0,854
Safety Shoes	0,954	0,921	0,958	0,725
Person	0,988	0,978	0,991	0,865
All Classes	0,976	0,954	0,977	0,787

The F1-Confidence curve (Figure 7) indicates that the model operates optimally at a confidence threshold of approx. 0.413, achieving a peak F1-Score of 0.97. This suggests a very strong balance between precision (low false alarms) and recall (low missed detections).

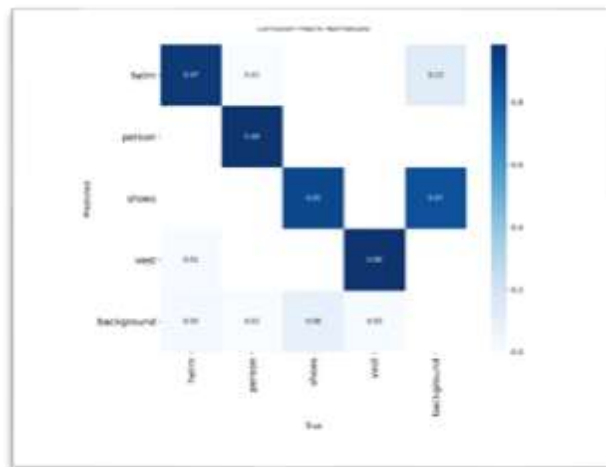
Figure 8 .F1-Confidence Curve Peaking at 0,97



Confusion Matrix Analysis

The normalized confusion matrix (Figure 7) further elucidates the classification accuracy. The model achieves 99% accuracy for identifying persons and 98% for vests. The primary source of error is in the "Shoes" class, where 8% of instances were misclassified as background (false negatives). This is a known challenge in computer vision for small, ground-level objects, yet the 91% true positive rate for shoes remains sufficient for effective safety auditing.

Figure 9. Confusion Matrix Normalize



Qualitative Verification

The spatial association logic was tested on inference data. The system successfully distinguished compliant and noncompliant workers. Figure 8 demonstrates four scenarios: (A) a fully equipped worker detected as compliant, (B) a worker correctly identified as noncompliant due to a missing vest, (C) a worker wearing only safety shoes and a vest, which is classified as

noncompliant because the helmet is absent, and (D) a worker detected with shoes and vest while holding a helmet in the hand, which is also labeled noncompliant since the helmet does not lie in the designated head region.

Figure 10 . Verification result: (a) Compliant worker detected with full PPE; (b) Non compliant worker detected with missing vest alert.



(a)

(b)

Figure 1. Verification result: (c) Non compliant worker detected without helmet; (d) Non compliant worker detected



(a)

(b)

RESULTS AND DISCUSSION

The experimental results indicate that the proposed system functions effectively not only as a PPE detection model but as a comprehensive compliance verification framework. While the high detection accuracy achieved by the YOLOv11 backbone confirms the robustness of the model, the primary contribution of this study lies in the integration of spatial association logic that enables per-worker PPE completeness assessment. This shows that the system is not only oriented towards technical performance, but also towards practical relevance in the context of implementing occupational safety in the field. Deep learning techniques have increasingly been adopted for construction safety monitoring due to their ability to automatically detect unsafe conditions and worker behaviors .

Unlike detection-oriented systems that infer compliance based solely on the presence of PPE objects in an image, the proposed approach explicitly enforces anatomical correctness through region-based verification. This system aligns detected PPE with defined body zones, preventing false compliance, such as PPE simply being carried or placed near the worker. This approach increase the validity of the results by ensuring verification under real world conditions, aligning with OHS standards.

Zone based partitioning provides a computational alternative to efficient pose estimation or keypoint methods. By dividing the PPE region proportionally within the worker's bounding box, the system adjust for height, posture, and camera perspective

while still enabling realtime processing, critical for resource constrained construction operations.

Unlike conventional PPE systems that simply detect items individually, this framework combines multi class detection with spatial verifications logic, resulting in more accurate and structured pass/fail decision, making it an automated safety audit tool rather than a passive detection tool.

Despite its high overall performance, safety footwear remains challenging due to occlusion and poor visual conditions, the system prioritizes a conservative principle, where false negative errors are safer than permitting risky behavior.

Limitations also arise in condition of heavy occlusion or low lighting. Integrating temporal tracking or additional context could improve detection consistency without changing the core design. Although tested in construction, this framework can be applied to other sectors such as manufacturing, mining, and logistics that have similar safety standards. When directly compared to previous research, this approach demonstrates more comprehensive conceptual advantages. Most previous studies based on YOLO or object detection only detect the presence of PPE without verifying its use thus risking presence bias. This study addresses this issue with body zone based spatial verifications, minimizing misinterpretation of improperly applied PPE.

Compared to pose estimation methods, this approach is simpler and computationally efficient, suitable for real time implementation on construction sites. In addition to detection, this framework generates explicit decisions (pass/fail), making it more contextual for safety evaluation. However, performance can degrade when helmet occlusion occurs, vests blend into the background. Pr when worker density is high. Computationally, while zoning is less demanding than pose estimation, evaluations such as FPS and memory usage are still necessary to ensure performance.

In terms of implementation, this system has the potential to be integrated with CCTV cameras and IoT-based safety management systems or web dashboards. However, implementation faces challenges such as lighting variations, hardware limitations, camera calibration requirements, and user acceptance. Therefore, its implementation requires not only technical readiness but also adequate operational support and safety policies.

CONCLUSION

This study presents an automated PPE compliance verification system that extends conventional object detection into a structured, per-worker safety assessment framework. By integrating the YOLOv11 deep learning architecture with anatomically grounded spatial association logic, the proposed system verifies not only the presence of PPE items but also their correct placement and completeness for individual workers.

The system demonstrates strong detection performance, achieving an overall mAP@50 of 0.979 and a peak F1-score of 0.97, while effectively reducing false compliance cases through region-based verification. Unlike conventional detection, this framework supports consistent compliance with OHS standards and is suitable for continuous, low bias safety audits, with a safety first approach. Although shoe detection is still hampered by occlusion and low lighting the system's performance remains adequate for field monitoring.

Further development will focus on integrating temporal tracking, expanding the dataset to nighttime and extreme environments, and adding a real time alert system to enhance safety management. These enhancements are expected to further strengthen the applicability of the proposed system for automated construction safety supervision.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to all parties who supported this research. Including providing data, academic input, and moral support. We hope the results of this research will be beneficial for the development of occupational safety technology in the construction sector.

References

- [1] L.; Liu et al., "Multi-Task Intelligent Monitoring of Construction Safety Based on Computer Vision," Buildings 2024, Vol. 14, Page 2429, vol. 14, no. 8, p. 2429, Aug. 2024, doi: 10.3390/buildings14082429.
- [2] ILO, "Occupational Safety and Health Statistics (OSH database)." Accessed: Nov. 07, 2025. [Online]. Available: <https://ilostat.ilo.org/methods/concepts-and-definitions/description-occupational-safety-and-health-statistics/>
- [3] BPJS Ketenagakerjaan, "Kecelakaan Kerja Makin Marak dalam Lima Tahun Terakhir." Accessed: Nov. 07, 2025. [Online]. Available: <https://www.bpjsketenagakerjaan.go.id/berita/28681/Kecelakaan-Kerja-makin-Marak-dalam-Lima-Tahun-Terakhir>
- [4] J. Wu, N. Cai, W. Chen, H. Wang, and G. Wang, "Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset," Autom. Constr., vol. 106, Oct. 2019, doi: 10.1016/j.AUTCON.2019.102894.
- [5] Kementerian Ketenagakerjaan RI, "Permenaker No. 5 Tahun 2018 Tentang Keselamatan Dan Kesehatan Kerja Lingkungan Kerja." Accessed: Nov. 07, 2025. [Online]. Available: <https://peraturan.go.id/id/permenaker-no-5-tahun-2018>
- [6] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," Autom. Constr., vol. 112, Apr. 2020, doi: 10.1016/j.AUTCON.2020.103085.

Academia Open

Vol. 11 No. 1 (2026): June

DOI: 10.21070/acopen.11.2026.13886

- [7] J. Lee and S. Lee, "Construction Site Safety Management: A Computer Vision and Deep Learning Approach," *Sensors (Basel)*, vol. 23, no. 2, p. 944, Jan. 2023, doi: 10.3390/s23020944.
- [8] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimedia Tools and Applications* 2022 82:6, vol. 82, no. 6, pp. 9243–9275, Aug. 2022, doi: 10.1007/S11042-022-13644-Y.
- [9] Q. An, Y. Xu, J. Yu, M. Tang, T. Liu, and F. Xu, "Research on Safety Helmet Detection Algorithm Based on Improved YOLOv5s," *Sensors* 2023, Vol. 23, Page 5824, vol. 23, no. 13, p. 5824, Jun. 2023, doi: 10.3390/S23135824.
- [10] X. Song, T. Zhang, and W. Yi, "An improved YOLOv8 safety helmet wearing detection network," *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/S41598-024-68446-Z.
- [11] V. S. K. Delhi, R. Sankarlal, and A. Thomas, "Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques," *Front. Built Environ.*, vol. 6, Sep. 2020, doi: 10.3389/FBUIL.2020.00136.
- [12] B. Mahaur and K. K. Mishra, "Small-object detection based on YOLOv5 in autonomous driving systems," *Pattern Recognit. Lett.*, vol. 168, pp. 115–122, Apr. 2023, doi: 10.1016/J.PATREC.2023.03.009.
- [13] N. A. N. M. Nazli, N. Sabri, R. Aminuddin, S. Ibrahim, S. Yusof, and S. D. N. M. Nasir, "A real-time system for detecting personal protective equipment compliance using deep learning model YOLOv5," *Procedia Comput. Sci.*, vol. 245, no. 1, pp. 647–656, Jan. 2024, doi: 10.1016/j.procs.2024.10.291.
- [14] J. Samperante, M. Agus, W. Putra, A. G. Permana, S. Informasi, and M. Informatika, "Implementasi Arsitektur Yolo V8 Dalam Mendeteksi Alat Pelindung Diri (APD) Di Sektor Konstruksi Dan Industri," vol. 1, no. 1, 2025.
- [15] B. Dwyer, J. Nelson, and J. Solawetz, "Roboflow: Computer vision developer framework." Accessed: Dec. 13, 2025. [Online]. Available: <https://roboflow.com/>
- [16] G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics (Version 8.0.0)." Accessed: Dec. 13, 2025. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [17] R. Padilla, S. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms." Accessed: Dec. 13, 2025. [Online]. Available: https://www.researchgate.net/publication/343194514_A_Survey_on_Performance_Metrics_for_Object-Detection_Algorithms
- [18] J. Liu, H. Luo, and H. Liu, "Deep learning-based data analytics for safety in construction," *Autom. Constr.*,