International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Deep Learning Framework for Automated Worker Helmet Detection and Safety Compliance Monitoring Using the Yolo Object Detection Model.

[1]Nayudu Malleswari
[1]M.Tech, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
nmalleswaric@gmail.com

[2]Mr. Ch.Venkata reddy
[2]Assistant Professor, Department of Computer Science and Engineering
DNR College of Engineering & Technology (Autonomous)
Bhimavaram – 534202, Andhra Pradesh, India
venkat95dnr@gmail.com

Abstract*— Construction and industrial environments involve substantial occupational risks, particularly from falling objects, collisions, and unsafe work practices. Safety helmets are therefore a critical component of personal protective equipment, but conventional compliance monitoring relies heavily on manual inspection and periodic supervision. Recent advances in deep learning and computer vision have enabled automated identification of workers, helmets, and non-compliance conditions from images and video streams. Among available object-detection approaches, the You Only Look Once (YOLO) family has attracted considerable attention because of its favorable balance between detection accuracy, computational efficiency, and real-time inference capability. This review examines the development of YOLO-based worker helmet detection from YOLOv5 to more recent architectures, emphasizing small-object detection, complex backgrounds, occlusion, illumination variation, model lightweighting, temporal tracking, and deployment at construction sites. Existing studies demonstrate that YOLOv5 can provide high-speed detection, while improved YOLOv8 models address small and distant helmet targets through attention and feature-fusion mechanisms. More recent YOLOv9-based approaches integrate pose estimation and multi-object tracking to provide continuous helmet-wearing-state monitoring. The review further analyzes datasets, evaluation metrics, compliance decision mechanisms, alert generation, edge deployment, and practical limitations. Finally, research gaps are identified toward robust, privacy-aware, context-aware, and continuously operating safety-compliance systems.

Keywords*— worker safety, safety helmet detection, construction safety, deep learning, YOLO, object detection, computer vision, PPE monitoring, compliance monitoring, real-time detection.

I. INTRODUCTION

Construction and industrial workplaces involve significant safety hazards, making proper use of personal protective equipment (PPE), particularly safety helmets, essential for preventing head injuries. However, conventional helmet-compliance monitoring depends largely on supervisors and manual CCTV inspection, which becomes difficult in large and continuously operating construction environments.

Recent advances in computer vision and deep learning have enabled automated detection of workers and safety helmets from images and video. Among these approaches, the YOLO family has gained considerable attention because of its real-time object-detection capability. Early studies using YOLOv5 demonstrated the feasibility of automated helmet detection, with reported results of up to 95.8% mAP, while lightweight variants achieved high inference speeds. Subsequent YOLOv8-based studies have focused on improving small-object detection, attention mechanisms, and robustness in complex construction environments.

More recent research has extended helmet detection beyond individual-frame recognition by integrating object detection with pose estimation and multi-object tracking. These approaches aim to associate helmets with individual workers, maintain compliance states across video frames, and reduce false alerts caused by occlusion and temporary detection failures.

Therefore, this review examines YOLO-based worker helmet detection from six perspectives: detection accuracy, small-object recognition, environmental robustness, temporal compliance monitoring, computational efficiency, and practical deployment. It further identifies major research challenges and future directions for reliable automated safety-compliance monitoring.

TABLE I
SCOPE OF THE REVIEW AND MAJOR RESEARCH DIRECTIONS

Research dimension	Main problem	Representative approaches	Review observation
Helmet detection	Identify helmet/head classes	YOLOv5, YOLOv8	Real-time detection is feasible
Small-object detection	Helmets occupy few pixels at distance	Attention, feature fusion, SR	Important for wide-area CCTV

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Research dimension	Main problem	Representative approaches	Review observation
Complex environments	Occlusion, clutter, fog, low light	CA, improved neck/backbone, augmentation	Remains a major source of missed detections
Compliance recognition	Associate helmet with worker	Spatial relationship, head-helmet matching	More meaningful than isolated helmet detection
Temporal monitoring	Maintain state across frames	StrongSORT, DeepSORT, temporal association	Reduces unstable frame-level decisions
Lightweight deployment	Limited edge resources	YOLOv5n/s, pruning, lightweight modules	Accuracy-speed trade-off remains important
Alert management	Convert detection into safety action	Rule-based alerts, web dashboards, notifications	Requires persistence and confidence thresholds
Generalization	Different sites/cameras/weather	Diverse datasets and augmentation	External validation remains insufficient

II. EVOLUTION OF YOLO-BASED HELMET DETECTION

A. Conventional Computer-Vision-Based Monitoring

Before deep-learning object detection became dominant, safety monitoring commonly relied on handcrafted image-processing techniques such as color segmentation, edge detection, shape analysis, and template matching. Such approaches can work under controlled conditions but generally depend on assumptions concerning helmet color, illumination, viewing angle, and object scale.

Construction sites violate many of these assumptions. Helmets may have different colors, workers can be partially hidden behind equipment, and camera viewpoints can vary substantially. Consequently, handcrafted methods have limited scalability for real-world deployment.

Deep convolutional neural networks improved this situation by learning visual features directly from training data. Instead of manually defining helmet characteristics, the detector learns discriminative representations from annotated images.

B. YOLOv5-Based Detection

YOLOv5 became an important baseline for construction safety applications because it offered multiple model scales, allowing developers to choose between computational efficiency and detection accuracy. Research comparing YOLOv5 variants showed that the larger YOLOv5x model can provide high detection accuracy, while smaller models can support real-time applications.

The primary advantage of this approach is that detection is performed in a single inference pipeline. For safety monitoring, this permits a camera stream to be analyzed continuously rather than requiring separate image-classification and localization stages.

However, helmet detection presents a difficult problem because helmets are often small relative to the complete image. A worker may occupy only a small portion of a wide-angle CCTV frame, and the helmet itself may contain very few pixels. Consequently, YOLOv5-based systems can experience false negatives when workers are distant or partially occluded.

Researchers have therefore introduced modified YOLOv5 architectures. Attention mechanisms can increase the contribution of relevant helmet regions, while feature-pyramid and multiscale approaches improve representation of small targets. Other work has introduced lightweight modules for embedded deployment. A YOLOv5-CS model, for example, was designed for embedded equipment and reported 31 FPS at 640 × 640 resolution on a Jetson Nano while substantially reducing model parameters.

III. SMALL-OBJECT DETECTION AND COMPLEX-SCENE ROBUSTNESS

Helmet detection is fundamentally a small-object detection problem when cameras are installed at a distance. The visual size of a helmet decreases as the worker moves farther from the camera. At the same time, construction environments contain scaffolding, cranes, cables, machinery, shadows, and other objects that may resemble helmet-like features.

Several studies have therefore modified the feature-extraction and feature-fusion stages of YOLO. Attention mechanisms are frequently used because they allow the network to emphasize informative regions while suppressing irrelevant background information.

An improved YOLOv5 approach called Helmet-Yolov5 introduced attention, atrous spatial pyramid pooling, and enhanced upsampling to improve small-object detection and multiscale feature representation. The authors reported improved accuracy compared with the original YOLOv5 while maintaining high inference speed.

Similarly, YOLOv8-based research has increasingly focused on small and distant helmet targets. The YOLOv8n-SLIM-CA model incorporates coordinate attention and lightweight modifications to improve detection under complex backgrounds and for small targets at long distances.

Another YOLOv8 study introduced attentional scale-sequence fusion and Dynamic Head mechanisms, specifically targeting small safety helmets and multilevel feature fusion.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

These studies reveal an important trend: simply replacing an older YOLO version with a newer version does not automatically solve helmet-detection challenges. Performance is strongly influenced by feature resolution, target scale, training diversity, camera placement, augmentation, and environmental conditions.

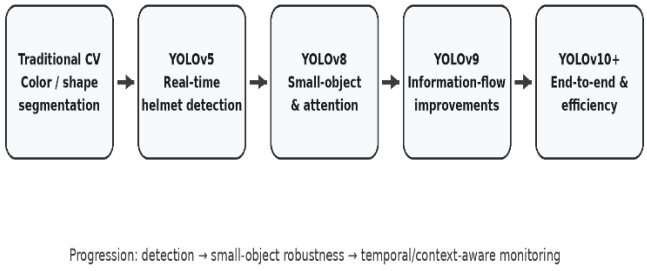


Fig. 1. Evolution of YOLO-based approaches for automated worker helmet detection and safety compliance monitoring.

YOLOv9 introduces Programmable Gradient Information and the GELAN architecture to address information loss during deep feature transformation. YOLOv10 subsequently investigated end-to-end real-time detection using consistent dual assignments and an NMS-free design, targeting improved latency and efficiency.

IV. YOLOv8, YOLOv9 AND RECENT DETECTION DIRECTIONS

YOLOv8 represents a major step in the practical development of real-time object detection. Its different model sizes enable users to select an appropriate computational configuration. In helmet-detection research, YOLOv8 has been investigated because construction helmets are frequently small and difficult to distinguish from background objects.

An IEEE conference study used YOLOv8 for construction-site helmet detection and considered difficult conditions including fog, rain, and low illumination. This direction is important because laboratory performance does not necessarily translate into field performance.

YOLOv9 introduces a different architectural philosophy. Its PGI mechanism is designed to preserve useful information through deep network transformations, while GELAN improves parameter utilization. These properties are potentially relevant to safety-helmet detection because small targets require preservation of fine-grained information.

More importantly, recent helmet-monitoring research has moved from frame-level object detection toward worker-level temporal state estimation. The YOLOv9 + YoloPose + StrongSORT framework first extracts worker head regions using pose keypoints, matches helmet detections to corresponding heads, and then uses tracking information across frames. This enables the system to determine whether a particular worker is continuously wearing a helmet rather than simply detecting a helmet somewhere in each frame.

This represents a significant conceptual transition:

Object detection → person-helmet association → temporal tracking → compliance decision.

TABLE II
COMPARATIVE ANALYSIS OF YOLO-BASED HELMET DETECTION STUDIES

Study / approach	Model / technique	Dataset / evaluation	Reported result	Main contribution	Main limitation
Safety helmet detection system [1]	YOLOv5x	5,000-image Hard Hat dataset	mAP 92.44%	Real-time helmet detection	Limited scene diversity
Real-time helmet detection [2]	YOLOv5n –x	7,063 images	YOLOv5x mAP 95.8%; YOLOv5n 70.4 FPS	Accuracy-speed comparison	Frame-level detection
StrongSORT helmet tracking [3]	YOLOv5s + StrongSORT	Helmet video/detection dataset	mAP@0.5 95.4% after improved loss	Continuous tracking	Computational overhead
Improved Helmet-YOLOv5 [4]	YOLOv5 + attention + ASPP	Construction-site images	Improved over baseline	Small-object and complex-scene detection	More architectural complexity
YOLOv5-CS [5]	Lightweight YOLOv5	Embedded deployment	31 FPS on Jetson Nano	Reduced model size	Accuracy reduction
YOLOv8 construction detection [6]	YOLOv8 m	Construction-site images	High-speed detection reported	Helmet/jacket detection	Low-light and occlusion challenges
YOLOv8n-SLIM-CA [7]	YOLOv8 + CA + lightweight design	Complex scenes	Improved small-target detection	Long-distance helmet detection	Requires specialized tuning
YOLOv8n-ASF-DH [8]	YOLOv8 + ASF + DyHead	Helmet dataset	Improved feature representation	Multiscale small-target detection	Increased architectural complexity
YOLOv9 + YoloPose + StrongSORT [9]	YOLOv9 + pose + tracking	Self-built dataset	98.89% accuracy	Continuous wearing-state monitoring	Dataset generalization
YOLOv10 [10]	NMS-free YOLO	COCO benchmark	Improved efficiency	End-to-end real-time detection	Helmet-specific evidence remains limited

Note: The performance values reported in Table II are taken from the respective studies and should not be interpreted as directly comparable. Differences in datasets, image resolutions, class definitions, hardware configurations, training procedures,

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

and evaluation protocols may significantly affect the reported performance.

V. FROM HELMET DETECTION TO SAFETY COMPLIANCE MONITORING

A major limitation of simple object detection is that identifying a helmet does not necessarily establish that a worker is wearing it.

For example, consider a construction image containing:

- three workers,
- two helmets,
- one helmet lying on equipment.

A detector may correctly identify all three helmet objects but cannot automatically conclude which worker is compliant. Therefore, safety compliance requires an additional association mechanism.

A general compliance pipeline can be expressed as

$$I_t \rightarrow D_t \rightarrow A_t \rightarrow T_t \rightarrow C_t \quad (1)$$

where:

- I_t represents the input video frame,
- D_t represents object detection,
- A_t represents person-helmet association,
- T_t represents temporal tracking,
- C_t represents the final compliance state.

For each worker w_i, the system can estimate:

$$C_i = \begin{cases} 1, helmet\ associated\ with\ worker \\ 0, helmet\ absent\ or\ non-compliant \end{cases} \quad (2)$$

However, making the decision from a single frame can generate false alarms. A helmet may temporarily disappear because of occlusion, motion blur, or an imperfect detection. Therefore, temporal confirmation is preferable.

A practical compliance rule can be defined as:

$$C_i^* = 1\ if\ \frac{1}{N}\sum_{t=1}^{N} C_{i,t} \geq \tau \quad (3)$$

where N is the number of observed frames and τ is the minimum compliance ratio.

This temporal perspective is consistent with recent YOLOv9 + YoloPose + StrongSORT research, where previous-frame information is used to improve helmet-wearing-state consistency and reduce missed detections caused by occlusion.

VI. TEMPORAL TRACKING AND VIDEO-BASED MONITORING

Construction sites are dynamic environments. Workers walk behind equipment, change direction, bend down, and interact with other workers. A detector operating independently on every frame may therefore generate inconsistent predictions.

Tracking algorithms address this problem by assigning persistent identities to detected workers. StrongSORT and related approaches can associate detections across consecutive frames. One helmet-monitoring study combined YOLOv5 with StrongSORT and reported that the tracking system could update object trajectories at approximately 0.05 s per frame.

The major advantage of tracking is that compliance becomes a **state-estimation problem** rather than a collection of independent classifications.

A worker can therefore be represented as:

$$W_i = \{I\ D_i, B_i, H_i, S_i\} \quad (4)$$

where I D_i represents the worker identity, B_i represents the bounding box, H_i represents the helmet association, and S_i represents the compliance state.

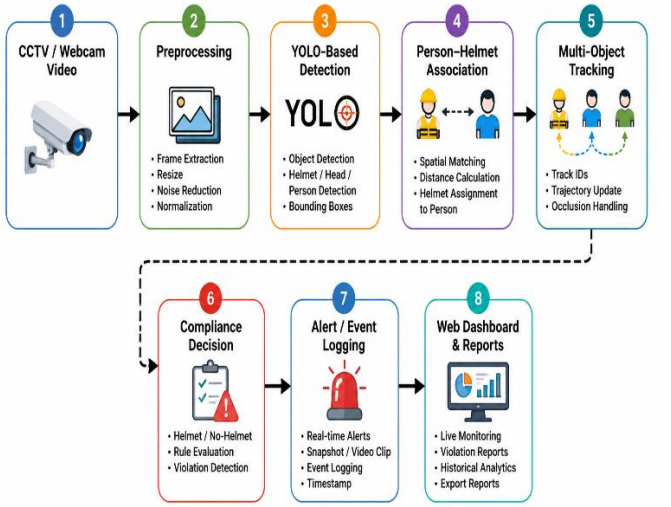


Fig. 2. General architecture of automated worker helmet detection and safety-compliance monitoring.

VII. DATASETS, TRAINING AND EVALUATION

Dataset quality is one of the most important factors influencing helmet-detection performance. Existing studies use different datasets, class definitions, image resolutions, and annotation strategies. For example, one YOLOv5 study used a 5,000-image Hard Hat dataset containing Helmet, Person, and Head classes but focused its experiments on Helmet and Head classes. Another YOLOv5 study used 7,063 images and divided them into training, validation, and testing subsets.

A robust review of helmet detection should therefore consider more than accuracy. Important metrics include:

Precision

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Precision indicates how many predicted helmet detections are correct.

Recall

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

Recall is particularly important in safety applications because missed non-compliance events can have serious consequences.

F1-score

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision+Recall} \quad (7)$$

Intersection over Union

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

$$\text{IoU} = \frac{\text{Area}(\text{Bp} \cap \text{Bg})}{\text{Area}(\text{Bp} \cup \text{Bg})} \quad (8)$$

where Bp and Bg represent predicted and ground-truth bounding boxes.

Mean Average Precision

mAP aggregates average precision across classes and IoU thresholds. It is widely used in object-detection research.

However, comparing mAP values across papers should be done carefully. Different studies use different datasets, image sizes, class definitions, hardware, and IoU thresholds. Therefore, a reported 95% mAP in one study does not automatically demonstrate superiority over a 92% result from another study.

VIII. ENVIRONMENTAL ROBUSTNESS AND REAL-WORLD CHALLENGES

Despite strong reported results, several challenges remain before YOLO-based helmet detection can become a universally reliable safety system.

A. Occlusion

Workers may be partially hidden by machinery, scaffolding, vehicles, or other workers. The helmet may disappear temporarily from the camera view.

B. Small and Distant Helmets

A distant worker may occupy only a small number of pixels. Fine helmet features can be lost during image resizing and deep feature extraction.

C. Illumination

Construction sites may operate during early morning, night, or under strong sunlight. Shadows and reflections can change the appearance of helmets.

D. Weather

Rain, fog, dust, and smoke can reduce image quality. YOLOv8 construction-site research specifically considered challenging conditions such as fog, rain, and low light.

E. Helmet Color and Design

Helmet color is not necessarily standardized across organizations. A model trained primarily on one color distribution may experience domain-shift problems at another site.

F. Worker Crowding

Multiple workers standing close together create overlapping bounding boxes and make person-helmet association more difficult.

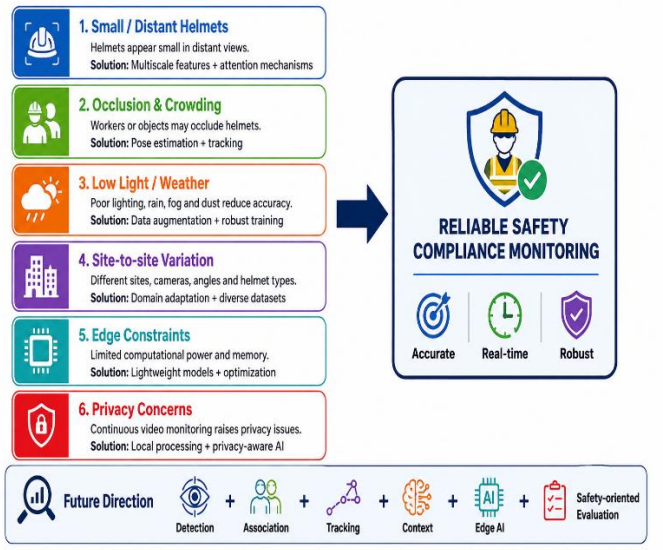


Fig. 3. Major challenges and research directions for robust YOLO-based safety-compliance monitoring.

IX. LIGHTWEIGHT MODELS AND EDGE DEPLOYMENT

Real-world construction monitoring requires continuous processing of video streams. Consequently, computational efficiency is nearly as important as detection accuracy.

Large YOLO models may provide improved accuracy but can require substantial computational resources. Lightweight models such as YOLOv5n and YOLOv5s are more appropriate for edge devices. Experimental comparisons have demonstrated that YOLOv5n can achieve high frame rates, whereas larger models generally provide better detection accuracy.

Lightweight architectural modifications are another important research direction. YOLOv5-CS, for example, reduced the parameter count and was evaluated on a Jetson Nano, demonstrating the possibility of deploying helmet detection directly on embedded hardware.

Recent YOLO research has increasingly focused on efficiency. YOLOv10 proposes NMS-free end-to-end detection and reports improved latency-efficiency trade-offs relative to previous YOLO variants.

For construction safety, however, the most accurate model is not necessarily the best model. A practical deployment should optimize:

Accuracy + FPS + latency + memory + energy + reliability.

A detector that achieves very high mAP but cannot process video with acceptable latency may be unsuitable for real-time safety monitoring.

X. ALERT GENERATION AND SAFETY-COMPLIANCE MANAGEMENT

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

The final objective of helmet detection is not simply to draw bounding boxes. The system should transform visual information into actionable safety information.
A practical monitoring platform can include:

- live camera monitoring;
- worker detection;
- helmet detection;
- worker-helmet association;
- violation confirmation;
- audible/visual alerts;
- notification to safety officers;
- timestamped violation records;
- image/video evidence;
- compliance statistics;
- historical reporting.

The recently published system corresponding closely to the title of this review combines YOLO detection with a Flask web application, video upload, webcam monitoring, automatic notifications, and violation recording.
Such integration demonstrates an important distinction between **object-detection research** and **safety-management systems**. A detector provides the visual prediction, whereas a complete compliance platform requires application-level logic, event management, reporting, and human intervention.

XI. CRITICAL ANALYSIS OF EXISTING RESEARCH

The reviewed literature demonstrates substantial progress, but several methodological limitations remain.

1. Dataset dependency

Many studies use relatively specialized datasets. A model can achieve high performance on a dataset while performing poorly on another construction site.

2. Lack of external validation

A major weakness is the limited use of completely unseen construction sites for evaluation. Random image splits may place visually similar frames from the same environment in both training and testing sets.

3. Frame-level evaluation

Traditional mAP and precision/recall measure object detection but do not necessarily measure whether the system correctly identifies **persistent worker-level compliance**.

4. Insufficient temporal evaluation

Video-based safety monitoring should report tracking accuracy, identity consistency, false-alert frequency, and violation-detection latency.

5. Limited environmental diversity

Performance under rain, dust, fog, low light, shadows, and severe occlusion remains a major concern.

6. Accuracy-efficiency trade-off

Large models may provide better detection accuracy but may not be suitable for embedded or edge deployment.

7. Human-in-the-loop requirements

Safety systems should not automatically interpret every low-confidence prediction as a violation. Confidence thresholds, temporal persistence, and human verification should be considered.

XII. RESEARCH GAPS AND FUTURE DIRECTIONS

The literature suggests that future helmet-monitoring systems should move beyond isolated frame-level detection toward context-aware safety intelligence.
First, multimodal safety monitoring could combine video with worker identity, location, restricted-zone information, and site activity. Helmet status could then be interpreted within the broader context of worker behavior.
Second, temporal learning can improve robustness. Instead of treating each frame independently, temporal transformers, recurrent networks, or tracking-based approaches can exploit information from previous observations.
Third, domain adaptation is required for deployment across different sites. A model trained at one construction project may encounter different camera angles, helmet colors, backgrounds, lighting, and worker populations at another project.
Fourth, edge AI should be explored to minimize latency and reduce dependence on continuous cloud connectivity.
Fifth, privacy-preserving techniques are important because continuous worker video monitoring involves sensitive visual information. Processing video locally or transmitting only detection metadata can reduce unnecessary exposure of raw imagery.
Sixth, evaluation should include safety-oriented measures rather than only object-detection metrics. Future studies should report:

- false alarms per worker-hour;
- missed violations;
- violation-detection latency;
- tracking consistency;
- site-to-site generalization;
- performance under environmental degradation;
- computational latency;
- energy consumption;
- privacy/security characteristics.

Finally, the development of standardized large-scale helmet datasets containing multiple construction sites, camera viewpoints, weather conditions, worker densities, helmet types, and temporal video sequences would substantially improve reproducibility.

CONCLUSION

YOLO-based deep learning has become an important approach for automated worker helmet detection and construction safety monitoring. The literature demonstrates a clear progression from conventional computer vision toward YOLOv5-based

real-time detection, improved YOLOv8 architectures for small and difficult targets, YOLOv9-based information-preserving models, and integrated detection-tracking frameworks for continuous helmet-wearing-state estimation. Existing studies report strong detection performance, but direct comparison remains difficult because datasets and evaluation protocols differ. The principal unresolved challenges are small-object detection, occlusion, environmental variation, temporal consistency, cross-site generalization, computational efficiency, and privacy. Future safety systems should therefore integrate detection, person-helmet association, tracking, contextual reasoning, edge deployment, and safety-oriented evaluation. Such a transition can transform helmet detection from a simple object-recognition task into a reliable automated safety-compliance monitoring framework.

REFERENCES

[1] M. M. A. A. H. et al., "Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety," *Applied Sciences*, vol. 12, no. 16, Art. no. 8268, 2022.

[2] Kisaezehra, M. U. Farooq, M. A. Bhutto, and A. K. Kazi, "Real-Time Safety Helmet Detection Using Yolov5 at Construction Sites," *Intelligent Automation & Soft Computing*, vol. 36, no. 1, pp. 911–927, 2023, doi: 10.32604/iasc.2023.031359.

[3] Y. Liu *et al.*, "Helmet-Wearing Tracking Detection Based on StrongSORT," *Sensors*, 2023.

[4] C. Shan, H. Liu, and Y. Yu, "Research on improved algorithm for helmet detection based on YOLOv5," *Scientific Reports*, vol. 13, Art. no. 18056, 2023.

[5] J. Li, J. Zhang, X. Zhang, and S. Wang, "Lightweight helmet detection algorithm based on improved YOLOv5," in *Proc. 2nd Int. Conf. Advanced Algorithms and Signal Image Processing (AASIP)*, 2022, Art. no. 124751P, doi: 10.1117/12.2659641.

[6] S. S. Mahmud *et al.*, "Safety Helmet Detection of Workers in Construction Site using YOLOv8," in *Proc. 26th Int. Conf. Computer and Information Technology (ICCIT)*, 2023, doi: 10.1109/ICCIT60459.2023.10441212.

[7] B. Lin, "Safety Helmet Detection Based on Improved YOLOv8," *IEEE Access*, vol. 12, pp. 28260–28272, 2024, doi: 10.1109/ACCESS.2024.3368161.

[8] B. Lin, "YOLOv8n-ASF-DH: An Enhanced Safety Helmet Detection Method," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3435453.

[9] "Detection of helmet use among construction workers via helmet-head region matching and state tracking," *Automation in Construction*, 2025.

[10] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," in *Proc. European Conf. Computer Vision (ECCV)*, 2024, doi: 10.1007/978-3-031-72751-1_1.

[11] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding, "YOLOv10: Real-Time End-to-End Object Detection," in *Advances in Neural Information Processing Systems*, vol. 37, 2024.

[12] M. A. Paz Bodero and P. E. Suarez Chavez, "A Computer Vision-Based System for Detecting Safety Helmet Compliance on Construction Sites Using YOLOv5s," *Congreso Internacional de Ingeniería Industrial de la Universidad de Lima*, 2026, doi: 10.26439/ciii2025.8657.