Computer Vision-based Integrated Framework for Industrial Worker Safety Compliance and Automated Attendance Monitoring using YOLOv8 and Facial Recognition

Pooja Mishra
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

Pratiksha Shevatekar
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

Hemant Nawghare
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

Ojas Satao
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

Ashitosh Bachute
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

Nakul Kapre
Computer Department
Dr.D.Y.Patil Institute of Engineering Management and Research
Pune, Maharashtra

ABSTRACT
Industrial and construction sites are always prone to safety-related issues because of dangerous equipment, labor movement, and a lack of standardization in the use of personal protective equipment (PPE). In addition, workforce attendance monitoring is still carried out using traditional paper-based attendance systems or biometric systems that are prone to proxy attendance, system delays, and maintenance problems. The current system addresses safety monitoring and attendance monitoring as two separate issues, leading to disjointed workforce management and poor decision-making.

This paper proposes a computer vision-based intelligent monitoring system that integrates real-time PPE use monitoring and automated workforce attendance monitoring using facial recognition. The proposed system uses the YOLOv8 deep learning-based object detection model for helmet and safety vest detection from CCTV cameras and facial embedding recognition models for worker identification

verification. By addressing both issues in a single system, the proposed system allows automated workforce attendance monitoring and real-time safety usage monitoring without any additional hardware infrastructure.

The system is intended to be implemented on top of the existing surveillance systems and has the capability to process live camera feeds, video files, and multi-camera systems. The experimental results carried out on the simulated industrial environment show high accuracy and real-time processing capabilities, making it suitable for Industry 4.0 environments. The proposed system encourages transparency in the workplace, minimizes safety incidents, prevents attendance fraud, and maximizes efficiency through scalability and best data practices.

Keywords
Computer Vision, Industrial Safety Monitoring, PPE Detection, YOLOv8, Facial Recognition, Workforce Attendance Automation, Deep Learning, Smart Construction Sites, Industry 4.0.

1. INTRODUCTION
The industrial and construction sector remains one of the most hazardous work environments in the world, characterized by rapid machine operation, dynamic working processes, and constant human-machine interaction. Even with the implementation of occupational safety standards such as OSHA and ISO 45001, industrial accidents still occur because of the lack of compliance with safety procedures, particularly the misuse of personal protective equipment (PPE). Helmets and safety vests are very significant in injury prevention, but it is hard to comply with them on a large scale in the industrial sector.

The traditional monitoring system is fully dependent on human observation for manual monitoring or analysis of footage from the surveillance camera. Although this is possible on a small scale, it becomes less accurate with a growing number of workers in different operational areas. Human observation is susceptible to human error, fatigue, and decreased monitoring efficiency, thus affecting the proactive part of accident prevention.

Another operational challenge is attendance management. The conventional attendance management system, whether it is in the form of an attendance register, an RFID card, or biometric systems, is susceptible to direct interaction and management. The system is susceptible to proxy attendance, hardware failure, and congestion during peak workforce arrival times. Moreover, the attendance management system is not normally

integrated with safety compliance. This leads to a situation where the workforce is physically present in the facility but not necessarily in compliance with the required safety standards.

Recent advances in artificial intelligence and computer vision have made it possible to develop visual analytics solutions that can analyze complex industrial environments in real-time. Deep learning-based object detection models have demonstrated remarkable performance in detecting multiple objects in a single image or video frame. Among these models, the YOLO (You Only Look Once) models have gained significant attention for their capability to achieve a balance between detection accuracy and processing speed. YOLOv8 introduces architectural advancements that enhance detection accuracy in occluded and varying lighting conditions.

The primary contributions of this work include:

1. Integration of a holistic monitoring system that encompasses safety compliance and attendance monitoring in one system.
2. Real-time PPE detection using YOLOv8 that is optimized for use in a CCTV environment.
3. Use of facial recognition technology that is proxy-proof for attendance marking.
4. Integration of the system with the existing surveillance system without the need for wearable sensors.
5. Addressing ethical issues pertaining to privacy and storage.

2. SYSTEM OVERVIEW AND OPERATIONAL FRAMEWORK

In the prevailing industrial setting, there is a requirement for monitoring systems that will be able to monitor the safety of the workers on a continuous basis and the records of the workers. The prevailing monitoring systems consider the monitoring of the safety of the workers and the monitoring of the workers' attendance as two separate tasks. The proposed system will provide a new computer vision operational approach that will seek to fill the prevailing gap by incorporating the monitoring of the usage of personal protective equipment and the monitoring of the workers' attendance through facial recognition.

The system operates by utilizing the existing surveillance systems such as the CCTV cameras or webcams, thus not requiring the installation of biometric hardware. The video streams are harvested from the industrial setting and are processed by utilizing a multi-component pipeline that entails frame acquisition, preprocessing, detection, recognition, and decision analytics. Each component has separate tasks that can interact with each other.

At the perception level, the video frames are processed using preprocessing methods such as resizing, illumination normalization, and noise reduction. In the industrial environment, the video frames are vulnerable to dust, motion blur, illumination changes, and perspective projections by the cameras. As a result, adaptive preprocessing methods improve the accuracy of detecting the video frames.

The safety monitoring system applies the YOLOv8 object detection model to locate workers and classify PPE types such as helmets and reflective safety vests. The bounding box localization method allows the system to detect multiple workers in a crowded environment. Once the lack of PPE is detected, the system records incidents of violation based on temporal indicators and camera locations.

In parallel with the safety detection process, the attendance automation system applies facial recognition for identity verification by embedding. The faces detected from the worker area are represented by feature encoding networks that encode numerical embeddings to denote the facial features of people. The embeddings are then matched with the databases of authorized workers according to similarity scores for identity verification.

With the integration of both pipelines, it is possible to authenticate the attendance events as well as the status of safety compliance. Thus, an employee will be rewarded not only for attendance but also for safety compliance. The violation notices, attendance details, and statistical details are all stored in a centralized database and can be retrieved through monitoring dashboards by authorized supervisors.

From the usability point of view, the system provides a number of operational modes. These include live camera surveillance, video analysis, and image testing. The supervisors can obtain violation images, attendance information, and safety compliance data without video analysis.

3. LITERATURE REVIEW

The integration of computer vision technology in the industrial safety monitoring system has seen a tremendous increase with the development of deep learning-based object detection and recognition systems. During the early stages of research, the primary objective of the safety monitoring system was the manual observation of the site aided by video recording of the site, where the safety personnel analyzed the video recording to detect any violation. Although useful for documentation purposes, the system was not capable of taking proactive measures and had a high labor cost.

Recent research work has shown the effectiveness of YOLO-based models for helmet, safety vest, and worker detection in a construction site environment. The recent developments in YOLOv7 and YOLOv8 models have led to the adoption of anchor-free detection methods and enhanced feature aggregation techniques, making the models more robust to occlusions and scales. These methods are suitable for high frame rate inference and can be adopted continuously in surveillance systems.

Simultaneous research work has also been focused on facial recognition for worker identification and access control. Deep learning models such as FaceNet and DeepFace rely on embedding similarity matching, making it possible to have efficient identification even with moderate pose variations. Contactless attendance systems developed using these models have shown improved efficiency compared to biometric scanners and RFID systems.

However, most implementations of research work remain disconnected from the specific domain tasks. PPE detection systems usually provide safety notifications without relating them to specific employee identities. Conversely, attendance automation systems are merely concerned with the verification of presence without considering the safety policy compliance in the workplace.

The latest trends in Industry 4.0 emphasize the need for completely integrated intelligent monitoring systems with the ability to process workforce behavior, productivity, and safety simultaneously. Research surveys have identified the need for multimodal monitoring systems with the ability to address

transparency and accident prevention in smart construction settings.

Table 1. Literature Review

Paper Title	Journal Name	Authors & Publication Date	Methodology	Limitation
Computer Vision for Safety Management: A Case Study in the Construction Industry	Proceedings of the International Conference on Industrial Engineering and Operations Management (IEOM)	Nguyen Quoc Khanh, 2020	Uses computer vision to monitor PPE and worker behavior on construction sites; employs real-world case study to validate automated safety compliance.	Limited to a single case study, may not generalize to diverse construction environments or large-scale deployment.
A Review of Computer Vision-Based Monitoring Approaches for Construction Workers' Work-Related Behaviors	IEEE Access (Review Paper, 2024)	Multiple Authors, Aug 2024	Conducts an extensive literature review of computer vision technologies for behavioral monitoring, including algorithm performance and safety impact.	Lacks original experimental validation or real-world implementation of reviewed methods.
Work Safety Assessment through Contextual Analysis with Computer Vision	IEEE Transactions on Industrial Informatics	A. Javed, T. Mahmood, M. Jeon, 2020	Develops and tests contextual vision models to assess worker safety via automated scene and activity recognition in hazardous environments.	Performance may degrade under occlusion, poor lighting, and dynamic industrial conditions not fully covered in testing.
A Computer Vision and Point Cloud-Based Monitoring Approach for Automated Construction Tasks Using Full-Scale Robotized Mobile Cranes	Journal of Intelligent Construction	Xiao Pan, Tony T. Y. Yang, Ruiwu Liu, Yifei Xiao, Fan Xie, Sept 2024	Processes synchronized image and point cloud data to track crane pose, operation, and progress; algorithms detect and estimate automated activities in real-time using multi-sensor data.	Tested only on full-scale robotized cranes; limited applicability to smaller machinery or varied construction scenarios.

Despite the above advancements, some issues still remain unaddressed. Industrial settings are characterized by worker occlusions, facial expressions obscured by protective gear, and lighting conditions that affect recognition accuracy. Dataset constraints also affect the generalization performance of models on different industrial regions. Ethical issues associated with facial recognition also demand privacy-preserving deployment strategies.

4. METHODOLOGY

4.1 System Architecture Overview

The monitoring framework is based on a modular design that is scalable for use in an industrial environment. The framework is made up of interconnected modules that deal with video acquisition, preprocessing, PPE detection, facial recognition, decision analytics, and reporting visualization.

The video feed is achieved through the installation of strategically positioned surveillance cameras. The video feed is processed for preprocessing, which entails the standardization of the resolution and the improvement of the image quality. Adaptive histogram equalization techniques are used to mitigate the impact of illumination changes, which are common in the industrial environment.

After the video has been preprocessed, it is then processed for PPE detection. The PPE detection module is developed using the YOLOv8 algorithm. The algorithm is trained on the labeled dataset with workers wearing helmets and safety vests from different angles. This is done to ensure faster convergence using transfer learning without influencing the accuracy of the detection.

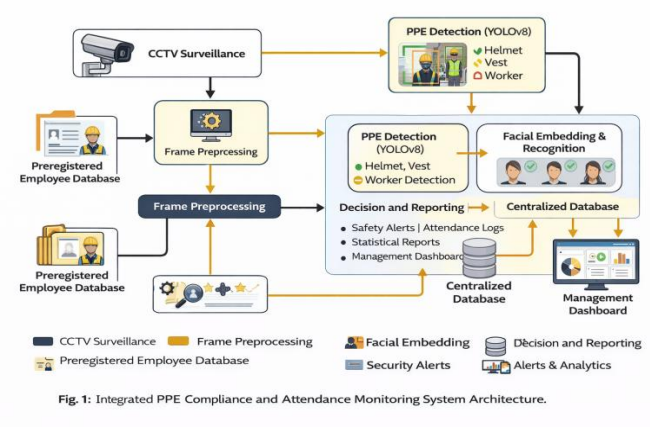

Fig 1: Integrated PPE Compliance and Attendance Monitor System

The regions of interest for people are then processed for facial recognition. A face detection algorithm is applied to extract the facial regions from the background noise. The feature embeddings are calculated using a face recognition network that is pre-trained to minimize intra-class variance and maximize inter-class variance.

The verification of attendance is carried out by comparing the cosine similarity between the generated embeddings and the registered embeddings of employees in the database. The decision logic is based on the output of both detection pipelines. Attendance is marked only if identity verification is successful, and the status of safety compliance indicates whether alerts are sent. All system outputs are logged along with timestamps and camera IDs.

4.2 Data Preparation and Model Training

The training datasets are made up of PPE image samples annotated for helmets, safety vests, and worker types based on varying illumination and camera angles. Data augmentation methods such as rotation, scaling, brightness adjustment, and addition of Gaussian noise are employed to enhance the generalization ability of the model.

The YOLOv8 model is initialized with weights from large-scale object detection datasets. Fine-tuning is done through transfer learning to fine-tune the layers that handle feature extraction to adapt to the industrial environment. Optimization is done through stochastic gradient descent with adaptive learning rates.

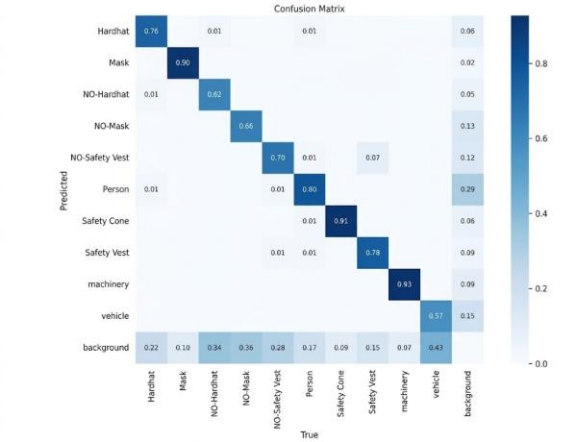

Fig 2: Confusion matrix of PPE kit detection model

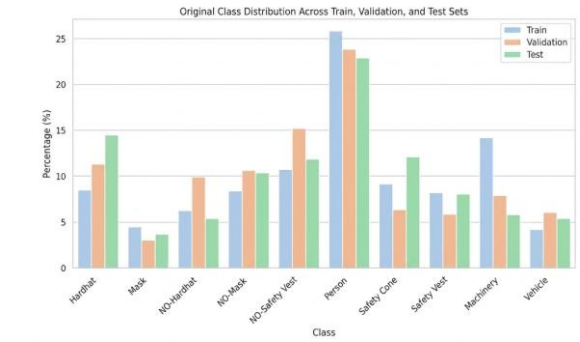

Fig 3: Original class distribution across Train, Validation, and Test Sets

4.3. Decision and Reporting Framework

The occurrence of events is transmitted to a central monitoring console. The attendance database holds data on the employee, the time of entry, and the camera zone. The database for safety violations holds evidence in the form of a snapshot that can be used to validate the audit trail.

The system employs statistical analysis for violation trends, attendance, and percentages of compliance. The system provides automatic notifications to supervisors for real-time PPE violation notifications. Privacy is ensured by encrypting the database and restricting access to the dashboard by authorized personnel.

4.4 Result Analysis

Performance assessment of the proposed integrated workforce monitoring system was carried out through experimental deployment of the system using simulated industrial video streams mimicking real-world working conditions. The purpose of performance assessment was to evaluate accuracy of detection, reliability of recognition, system response, and feasibility of deployment in industrial settings.

The PPE detection module performed well under different lighting conditions and worker density settings. The anchor-free detection framework of YOLOv8 facilitated precise localization of helmets and vests even when multiple workers were present in a single frame. Experimental results revealed that the system was able to perform with consistent detection confidence even when workers were moderately occluded, either partially overlapping each other or interacting with industrial equipment. Accuracy of detection was greatly enhanced using domain-specific fine-tuning and techniques such as brightness adjustment and rotation.

Fig 4: Before Detection

Fig 5: After Detection

The system performed well in real-time inference. On hardware equipped with GPU support, the system was able to process data at speeds that could easily support live CCTV streams without any lag.Buffering and preprocessing of frames were implemented to achieve the optimal system throughput, enabling supervisors to receive notifications almost instantly after PPE offenses. This is a highly significant factor for proactive accident prevention and not only for analysis after accidents.

The effectiveness of attendance verification was determined by the accuracy of facial recognition depending on different facial visibility conditions. Multiple facial captures from different angles during facial registration enhanced the accuracy of facial recognition in the real-world environment. The similarity comparison mechanism based on embedding successfully addressed the proxy attendance issue by confirming the identity of the person before attendance registration.

The accuracy of facial recognition was impacted by safety equipment such as helmets or goggles that partially covered facial features. In a densely populated area or when workers were turned away from the cameras, the confidence level of face detection was sometimes reduced. These factors emphasize the importance of optimal camera placement and diversity in the real-world environment of the system.

The integrated system provided a great advantage in terms of efficiency in monitoring compared to other monitoring systems. The direct link between PPE compliance and the confirmed identity of the workers allowed supervisors to monitor the compliance of individual workers as opposed to anonymous workers. This improves accountability and makes it simpler to offer specific safety training.

The system provided relevant operational data using automated reporting systems. Supervisors can monitor attendance trends, identify hotspots where repeated offenses occur, and assess compliance percentages over time.

The feedback received from the simulated deployment activity showed an improvement in efficiency in monitoring as opposed to the manual monitoring processes. Automated alerts eliminated the need for constant human monitoring, allowing safety officers to focus on high-priority events. Managers also reported a decrease in administrative tasks related to keeping attendance records.

However, some operational considerations were also highlighted. These include dust accumulation, lighting variations, and motion blur, which impacted the effectiveness of the detection system. Bandwidth limitations also impacted the video streaming quality, thus highlighting the importance of edge computing strategies in such large-scale implementations.

Ethical considerations were also highlighted during the testing phase. Facial recognition systems require the secure storage of biometric information to promote worker cooperation and trust. Encryption storage strategies, data retention policies, and role-based access to the dashboard were thus incorporated into the system.

Fig 6: Video Result

Conclusion from the above results, it is evident that the proposed integrated monitoring framework is an efficient and effective solution that can improve industrial safety compliance and worker transparency. The proposed system combines automation, accountability, and operational intelligence into a single intelligent platform.

5. KEY DIFFERENCES

Table 2: Comparison

Comparison Parameter	**Traditional / Earlier Models**	**Proposed Integrated Framework**
System Architecture	Independent systems for PPE detection and attendance	Unified dual-pipeline architecture integrating PPE detection and facial recognition
Monitoring Approach	Safety and attendance handled separately	Simultaneous safety compliance and identity verification
Violation Alerts	Generic alerts without identity mapping	Identity-associated real-time alerts with timestamp and camera zone
Attendance Method	Manual registers, RFID, or biometric scanners	Contactless facial recognition using embedding similarity matching
Proxy Attendance Prevention	Limited; biometric spoofing possible	Strong prevention through face embedding verification

Comparison Parameter	Traditional / Earlier Models	Proposed Integrated Framework
Infrastructure Requirement	Requires additional biometric hardware	Utilizes existing CCTV infrastructure
Real-Time Processing	Often semi-automated or reviewed manually	Fully automated real-time inference using YOLOv8
Detection Model Type	Earlier YOLO versions or region-based CNNs	YOLOv8 anchor-free architecture optimized for speed and robustness
Scalability	Limited to single-zone or isolated deployment	Multi-camera scalable framework for large industrial sites
Data Integration	Fragmented logs for safety and attendance	Centralized database integrating compliance, identity, and analytics

6. CONCLUSION AND FUTURE SCOPE

This research work has provided a complete computer vision solution for surveillance that aims to address two significant issues in the industrial sector: ensuring safety compliance among employees and ensuring correct attendance records. With the combination of PPE recognition through YOLOv8 and identity recognition through facial recognition, this system enables the simultaneous tracking of safety compliance and attendance.

This complete system has reduced the need for human surveillance and eliminated the need for biometric attendance systems that involve physical contact. The simultaneous tracking of the surveillance video enables supervisors to track safety compliance in real-time and ensure correct attendance records without any administrative issues. The combination of safety compliance and employee identities enhances employee safety.

The results of the experiment in a real-world setting have proved the accuracy of detection and the capability of real-time processing. The automated notification and analysis of the dashboard have enhanced the decision-making capability of the system, which provides useful information about worker dynamics and safety trends. The modularity of the system also enables scalability in an industrial and construction environment with minimal modifications to the existing infrastructure.

Although the results are encouraging, there are some limitations to this system. The facial recognition system may be affected by the use of protective clothing covering the face or an improperly placed camera. Environmental conditions such as lighting or high dust density may also affect the accuracy of the system. To address these problems, it is required to enhance the dataset and make the necessary modifications to the models.

Some of the aspects that should be investigated in the future include the development of modules for analyzing worker behavior that can detect potentially risky worker behavior such as entering restricted areas or operating machinery incorrectly. Multimodal surveillance systems using Internet of Things (IoT) sensors embedded in worker helmets or industrial machinery could also enhance situational awareness by combining environmental and visual data.

Another promising area for the deployment of edge computing is the reduction of latency and improvement of scalability in remote industrial sites. Methods of privacy-preserving machine learning, like federated learning, could also facilitate collaborative learning of models without the need for centralized storage of biometric data.

Improvements in three-dimensional vision and depth sensing technology could also improve the accuracy of detection in cluttered settings. In the end, intelligent monitoring systems that integrate artificial intelligence, ethical governance, and human oversight will play a pivotal role in the development of Industry 5.0 settings centered on safety, efficiency, and human-machine collaboration.

7. ACKNOWLEDGMENT

The authors would also like to extend their sincere thanks to the Department of Computer Engineering, Dr. D. Y. Patil Institute of Engineering Management and Research, Akurdi, Pune, for providing academic guidance and infrastructure support during the development of this research work. Special thanks are extended to the project mentors and faculty members whose continuous encouragement and technical inputs have helped in the successful completion of this research work.

8. REFERENCES

[1] Nguyen Quoc Khanh, “Computer Vision for Safety Management: A Case Study in the Construction Industry,” Proceedings of the International Conference on Industrial Engineering and Operations Management (IEOM), 2020, pp. 1254–1261.

[2] “A Review of Computer Vision-Based Monitoring Approaches for Construction Workers’ Work-Related Behaviors,” Aug. 2024.

[3] GOV.UK, Construction 2025: Strategy. Accessed Apr. 20, 2023. [Online]. Available: https://www.gov.uk/government/publications/construction-2025-strategy

[4] A. Javed, T. Mahmood, and M. Jeon, "Work Safety Assessment through Contextual Analysis with Computer Vision," IEEE Transactions on Industrial Informatics, 2020.

[5] X. Pan, T. T. Y. Yang, R. Liu, Y. Xiao, and F. Xie, "A computer vision and point cloud based monitoring approach for automated construction tasks using full-scale robotized mobile cranes," Journal of Intelligent Construction, Sept. 2024.

[6] J. Redmon et al., “You Only Look Once: Unified Real-Time Object Detection,” Proc. IEEE CVPR, 2016.

[7] A. Bochkovskiy, C. Wang, and H. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv preprint arXiv:2004.10934, 2020.

[8] G. Jocher et al., “Ultralytics YOLOv8 Documentation,” Ultralytics, 2023.

[9] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition," Proc. IEEE CVPR, 2015.

[10] J. Li et al., "Computer Vision-Based Monitoring Approaches for Construction Workers' Work Behaviors," IEEE Access, 2024.

[11] International Labour Organization (ILO), Occupational Safety Statistics Report, 2023.

[12] GOV.UK, "Construction 2025 Industrial Strategy," Government Publication.

[13] OpenCV Documentation, Open Source Computer Vision Library.

[14] Attendance Recording System Face Detection and Recognition Models Enais Adnan Moses Deli;Hasan Thabit Rashid Kurmasha 2025