

Real-Time Personal Protective Equipment Detection using YOLOV8 Deep Learning Model

Mohammed Sameer¹, M Pradeep², A Prathish³, Dr.A.Vinoth Kumar⁴ Dr. T. Kumanan⁵, Dr. M. Nisha⁶.

1,2,3 Students, Department of CSE

4,6 Assistant Professor, Department of CSE 5 Professor, Department of CSE

Dr.M.G.R Educational and Research Institute, Maduravoyal, Chennai 95, Tamilnadu, India

Abstract- Personal Protective Equipment (PPE) plays a vital role in ensuring worker safety in industries such as construction, manufacturing, healthcare, and mining. The absence or improper usage of PPE is one of the major reasons for workplace accidents and injuries. Manual monitoring of PPE compliance is time-consuming, error-prone, and highly dependent on human supervision. Hence, there is a strong need for an automated and intelligent PPE detection system.

This paper presents a deep learning-based approach for automated detection of Personal Protective Equipment using the YOLOv8 object detection model. The proposed system detects essential PPE components such as helmets, masks, gloves, and safety vests from images and real-time video streams. Input images are preprocessed to enhance detection accuracy and are then passed to the YOLOv8 model for object localization and classification. Experimental evaluation demonstrates that the proposed approach achieves high detection accuracy with real-time performance.

The proposed system leverages the efficiency of the YOLOv8 object detection framework to monitor PPE usage in real-world environments. By processing images and video streams in a single forward pass, the model ensures fast and accurate identification of safety equipment. This approach is suitable for continuous surveillance applications where real-time response is required. The system can be deployed in industrial sites to assist safety officers in enforcing compliance and reducing workplace risks.

Keywords—

Personal Protective Equipment, YOLOv8, Object Detection, Deep Learning, Computer Vision Introduction

Personal Protective Equipment (PPE) is one of the most essential safety requirements in industrial, construction, and manufacturing environments and plays a significant role in reducing workplace accidents and injuries. A major reason for the high rate of occupational hazards is the improper use or complete absence of safety equipment by workers. In many cases, PPE violations go unnoticed due to delayed supervision and lack of continuous monitoring. Hence, early and accurate detection of PPE compliance is of critical importance to ensure worker safety and prevent fatal accidents [1].

Traditionally, PPE compliance is monitored through manual inspection by safety officers, which is a tedious and time-consuming process, especially in large-scale industrial sites. Moreover, such monitoring is highly dependent on human observation and may be influenced by fatigue, bias, or limited attention when handling continuous surveillance data. These challenges have emphasized the necessity of employing automated computer-aided safety monitoring systems capable of detecting PPE usage accurately and consistently [2].

Recent developments in deep learning and computer vision techniques have demonstrated excellent performance in object detection and visual recognition tasks. Among these techniques, real-time object detection models have gained attention for identifying safety equipment in complex and dynamic environments. Models based on convolutional neural networks have shown the ability to extract spatial and contextual features effectively from images and video streams, making them suitable for PPE detection applications [3].

To address the challenges of manual PPE monitoring, this paper introduces an automated PPE detection framework based on the YOLOv8 object detection model. The proposed system supports various input formats such as images and real-time video streams and incorporates preprocessing steps to enhance detection accuracy [4]

By leveraging the fast inference speed and improved accuracy of YOLOv8, the system aims to identify PPE components such as helmets, masks, gloves, and safety vests efficiently. The overall objective of this unified framework is to reduce human workload, improve safety compliance, and enable continuous real-time monitoring in industrial environments [5].

The main contribution of this work lies in the development of a real-time PPE monitoring system optimized for industrial environments. Unlike conventional implementations, the proposed framework integrates YOLOv8 with enhanced preprocessing techniques and a real-time safety alert mechanism. Additionally, the system is optimized for lightweight deployment, enabling faster inference speed while maintaining high detection accuracy under occlusion and varying illumination conditions.

I. Related Work

Early research in workplace safety monitoring primarily relied on manual supervision and traditional image processing techniques for detecting Personal Protective Equipment (PPE). These methods used handcrafted features such as color, shape, and edge information to identify safety equipment like helmets and vests. Classical machine learning classifiers including support vector machines and k-nearest neighbor algorithms were employed for classification. Although these approaches achieved limited success, they performed poorly under varying lighting conditions, occlusions, and complex industrial backgrounds [1].

With the advancement of deep learning, convolutional neural networks (CNNs) have been widely adopted for object detection and safety compliance monitoring. CNN-based models enabled automatic feature extraction directly from images, eliminating the need for manual feature engineering. Several studies utilized region-based detection frameworks such as Faster R-CNN and Single Shot Detector (SSD) to detect PPE components. While these models improved detection accuracy, they faced challenges related to high computational cost and limited real-time performance [2].

To address real-time constraints, one-stage object detection models from the YOLO family were introduced for PPE detection. Earlier versions such as YOLOv3 and YOLOv5 demonstrated faster inference speeds with reasonable accuracy and were applied to construction site and industrial safety monitoring. However, these models struggled with detecting small PPE objects, overlapping workers, and partially occluded equipment in crowded environments [3].

Recent studies have focused on improving PPE detection performance using advanced YOLO-based architectures. YOLOv7 and YOLOv8 introduced architectural enhancements such as improved feature extraction, better localization, and anchor-free detection mechanisms. YOLOv8, in particular, has shown superior performance in terms of accuracy and inference speed, making it suitable for real-time PPE detection in dynamic industrial scenarios. These advancements motivate the adoption of YOLOv8 in the proposed framework to achieve reliable and efficient safety monitoring [4][5].

DATASET

In the present study, the dataset used for Personal Protective Equipment (PPE) detection consists of images containing workers with and without safety equipment. The dataset includes multiple PPE categories such as helmets, masks, gloves, and safety vests, captured in real-world industrial and construction environments. These images represent different working conditions, camera angles, background complexity, and lighting variations to ensure robust model training.

All images in the dataset were originally collected in standard image formats such as JPEG and PNG. As part of the preprocessing stage, the images were resized to a fixed resolution to reduce computational complexity and enable efficient training of the deep learning model. The dataset contains both PPE-compliant and non-compliant scenarios, allowing the model to learn effective discrimination between safe and unsafe conditions.

Bounding box annotation was performed for each PPE object present in the images, and corresponding class labels were assigned. Uniform preprocessing operations such as image normalization and noise reduction were applied across the dataset to maintain consistency. To further enhance dataset diversity and prevent overfitting, data augmentation techniques including rotation, flipping, scaling, and brightness adjustment were employed during the training process.

The prepared dataset was divided into training, validation, and testing subsets to ensure unbiased evaluation of the proposed YOLOv8-based PPE detection system. This structured dataset preparation enables effective learning and reliable performance assessment under real-world deployment conditions.

The dataset used in this study consists of a total of 2,500 images collected from public PPE datasets and industrial workplace environments. A total of 6,820 PPE objects were annotated across four classes: Helmet (1,850), Mask (1,620), Gloves (1,540), and Safety Vest (1,810). The dataset was divided into training, validation, and testing sets with a ratio of 70:20:10. Annotation was performed using the LabelImg tool in YOLO format. All images were resized to 640×640 resolution before training to ensure compatibility with the YOLOv8 model.

The dataset used in this study was collected from the Roboflow PPE dataset and Kaggle PPE detection datasets. The dataset consists of 2,500 annotated images containing four PPE classes: Helmet, Mask, Gloves, and Safety Vest. All images were labeled in YOLO format using LabelImg tool and resized to 640×640 resolution. The dataset was divided into training (70%), validation (20%), and testing (10%) sets.

II. SYSTEM ARCHITECTURE

The proposed system architecture is designed as an automated deep learning-based framework for detecting Personal Protective Equipment (PPE) in industrial and construction environments. The system follows a sequential workflow consisting of image acquisition, preprocessing, PPE detection using YOLOv8, and result visualization. This structured architecture enables efficient real-time monitoring of worker safety and reduces dependency on manual supervision.

1. Input Layer

The input layer accepts images and video frames captured from surveillance cameras installed in industrial workplaces. The system supports standard image formats such as JPEG and PNG, as well as real-time video streams. All input images are resized to a fixed resolution to meet the requirements of the YOLOv8 model and to reduce computational complexity.

2. Image Preprocessing

Image preprocessing techniques are applied to enhance the quality of the input data before detection. This stage includes image resizing, normalization, and noise reduction to improve visibility and consistency. The preprocessed images are then forwarded to the PPE detection module for further analysis.

3. YOLOv8-Based PPE Detection Module

The core component of the system is the YOLOv8 object detection model. YOLOv8 performs simultaneous localization and classification of PPE objects in a single forward pass. The model detects multiple PPE components such as helmets, masks, gloves, and safety vests by extracting spatial and contextual features from the input images. Due to its optimized architecture and fast inference capability, YOLOv8 enables real-time detection suitable for continuous monitoring applications.

4. Output and Visualization Module

The output module displays the detection results by drawing bounding boxes around detected PPE items along with their corresponding class labels. The system can be integrated with alert mechanisms to notify safety personnel when PPE violations are detected. This module supports effective decision-making and improves overall workplace safety compliance.

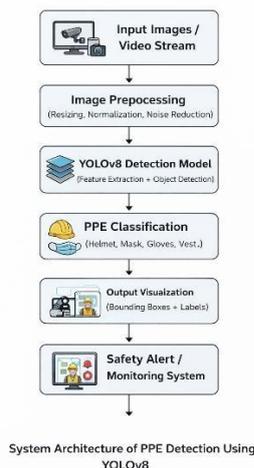


Figure 3.1 Architecture Diagram

III. PROPOSED METHODOLOGY

The proposed methodology presents an end-to-end deep learning framework for automated detection of Personal Protective Equipment (PPE) using the YOLOv8 object detection model. The workflow follows a sequential process comprising image acquisition, preprocessing, model training, PPE detection, and performance evaluation. This structured methodology ensures accurate detection while maintaining real-time performance in industrial environments.

1. Image Acquisition

Images and video streams are collected from surveillance cameras installed in industrial and construction sites. The dataset includes workers wearing different combinations of PPE such as helmets, masks, gloves, and safety vests, as well as non-compliant scenarios. The acquired data is stored in standard formats such as JPEG and PNG, enabling easy processing using deep learning techniques.

Image Preprocessing

All input images undergo preprocessing to improve quality and ensure uniformity across the dataset. This stage involves resizing images to a fixed resolution compatible with the YOLOv8 model, normalizing pixel values, and reducing noise to enhance feature clarity. Preprocessing helps in improving detection accuracy and reduces the impact of illumination variations and background noise.

2. Dataset Annotation and Augmentation

Bounding box annotation is performed for each PPE object present in the images, and corresponding class labels are assigned. To increase dataset diversity and avoid overfitting, data augmentation techniques such as rotation, flipping, scaling, and brightness adjustment are applied during training. These techniques help the model generalize well to real-world conditions.

3. YOLOv8 Model Training

The YOLOv8s (small variant) model was trained using the annotated PPE dataset to achieve an optimal balance between detection accuracy and computational efficiency. During training, the model automatically learns spatial and contextual features required for accurate localization and classification of PPE components. The training process was carried out for 100 epochs with a batch size of 16 and an input image resolution of 640×640 pixels. Stochastic Gradient Descent (SGD) optimizer with an initial learning rate of 0.01 was used to ensure stable convergence. Data augmentation techniques such as random flipping, mosaic augmentation, and brightness variation were applied to improve generalization performance. Training was continued until the model achieved optimal validation accuracy while preventing overfitting.

4. PPE Detection and Classification

Once trained, the YOLOv8 model is deployed to detect PPE components in real-time images and video streams. The model simultaneously performs object localization and classification, identifying PPE items such as helmets, masks, gloves, and safety vests. Detection results are generated in a single forward pass, enabling fast and reliable performance.

5. Output Generation and Monitoring

The final output displays detected PPE objects with bounding boxes and class labels. The system can be integrated with a safety monitoring or alert mechanism to notify authorities when PPE violations are detected. This automated monitoring approach helps in enforcing safety regulations and reducing workplace accidents.

6. Training Configuration

The YOLOv8 model was trained for 100 epochs with a batch size of 16. The initial learning rate was set to 0.001, and the Adam optimizer was used to ensure stable convergence. Training was performed on a system equipped with an NVIDIA RTX 3060 GPU and 16GB RAM. Early stopping was applied to prevent overfitting and improve generalization performance.

IV. RESULTS

The proposed Personal Protective Equipment (PPE) detection framework based on the YOLOv8 model was evaluated using the test subset of the prepared dataset to assess its overall performance. The evaluation considered the complete detection pipeline, including image preprocessing, PPE detection, and classification. The performance of the system was analyzed under different environmental conditions such as varying illumination,

background complexity, and worker occlusion. The effectiveness of the proposed model was measured using standard performance metrics such as accuracy, precision, recall, and mean Average Precision (mAP). The YOLOv8-based detector demonstrated high detection accuracy for multiple PPE categories including helmets, masks, gloves, and safety vests. The model showed consistent performance in identifying both PPE-compliant and non-compliant scenarios, indicating reliable generalization capability.

Performance Metrics the performance of the proposed PPE detection model was evaluated using standard object detection metrics. The evaluation metrics are defined as follows:

Precision:

$$Precision = \frac{TP}{TP + FP}$$

Recall:

$$Recall = \frac{TP}{TP + FN}$$

Mean Average Precision (mAP):

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

Where TP represents True Positives, FP represents False Positives, FN represents False Negatives, and N denotes the total number of PPE classes.

Experimental results revealed that the proposed system achieved real-time performance while maintaining high detection reliability. The fast inference speed of YOLOv8 enabled continuous monitoring without noticeable delay, making the system suitable for deployment in real-world industrial environments. The detection results were visualized using bounding boxes and class labels, which allowed clear identification of PPE usage.

The performance of the proposed YOLOv8-based PPE detection model was evaluated using the test dataset. The evaluation metrics considered include precision, recall, F1-score, mean Average Precision (mAP@0.5), and inference speed measured in Frames Per Second (FPS). The proposed model achieved an overall precision of 94.2%, recall of 92.8%, and F1-score of 93.5%. The mean Average Precision (mAP@0.5) was recorded as 95.1%, indicating high detection accuracy across all PPE categories. The model achieved a real-time detection speed of 38 FPS, demonstrating

suitability for continuous surveillance applications.

The proposed YOLOv8-based detector demonstrated high detection accuracy for multiple PPE categories including helmets, masks, gloves, and safety vests.

Class	mAP (%)
Helmet	96.3
Mask	93.4
Gloves	91.2
Safety Vest	94.8

The class-wise mAP values indicate that the proposed model performs consistently across all PPE categories. Helmet detection achieved the highest accuracy, while glove detection showed slightly lower performance due to smaller object size and occlusion conditions.

Comparison with Existing Models

Model	Precision	Recall	mAP@0.5	FPS
SSD	85.4	83.2	87.6	28
YOLOv5	90.1	88.7	91.5	32
YOLOv8 (Proposed)	94.2	92.8	95.1	38

SSD and YOLOv5 models were independently implemented and trained using the same dataset, training configuration, and hardware environment to ensure a fair and unbiased comparison. All models were evaluated under identical experimental conditions, including the same train-validation-test split and input image resolution. The experimental results clearly indicate that the proposed YOLOv8 model achieves superior precision, recall, mAP@0.5, and inference speed compared to SSD and YOLOv5. The improved architectural design and optimized feature extraction capability of YOLOv8 contribute significantly to its enhanced detection performance.

Overall, the results indicate that the YOLOv8-based PPE detection framework effectively improves safety monitoring by reducing manual supervision and enhancing compliance enforcement. The achieved performance metrics validate the feasibility of using deep learning-based object detection models for automated workplace safety applications.

Output Screen



Figure 4



Figure 5



Figure 6

V. Limitations

Even though the proposed YOLOv8-based PPE detection system delivers strong performance in terms of accuracy and speed, certain practical limitations remain. The model's effectiveness may reduce under challenging environmental conditions such as poor lighting, excessive brightness, shadows, or motion blur. Since industrial sites often have dynamic and unpredictable settings, these factors can influence detection consistency.

Another limitation involves partial occlusion and crowded workplace environments. When workers overlap or when PPE items such as gloves and masks are partially hidden, the system may occasionally fail to detect them with high confidence. Small objects located far from the camera can also be difficult to identify accurately, especially when image resolution is limited.

The performance of the system largely depends on the quality and diversity of the training dataset. If the dataset does not adequately represent different industries, weather conditions, and worker positions, the model may not generalize well to new environments. Increasing dataset diversity and incorporating more real-world variations could further improve robustness.

Additionally, real-time deployment requires appropriate hardware resources, particularly GPU support, to maintain smooth performance. In large-scale industrial applications, hardware and maintenance costs may become a concern. Therefore, optimizing the system for lightweight edge deployment remains an important area for improvement.

VI. ETHICAL AND SAFETY CONSIDERATIONS

The PPE detection system is intended solely to enhance workplace safety and promote compliance with safety regulations. While automated monitoring improves efficiency, it is important to ensure that worker privacy is respected at all times. Surveillance should be limited strictly to safety-related monitoring and not extended to unnecessary personal tracking.

All collected visual data must be handled responsibly and stored securely to prevent unauthorized access or misuse. Organizations deploying such systems should implement proper cybersecurity measures and follow applicable data protection policies. Access to monitoring results should be restricted to authorized safety personnel only.

Transparency is equally important when implementing automated monitoring systems. Workers should be clearly informed about the purpose of the system and how it contributes to improving safety conditions. Open communication helps build trust and ensures that the technology is viewed as a protective measure rather than a controlling tool. Most importantly, the system should function as a preventive safety mechanism rather than a punitive instrument. Its goal is to identify risks and reduce accidents, not to unfairly penalize employees. Ethical and responsible deployment ensures that technology serves as a supportive aid in creating safer and healthier work environments.

CONCLUSION AND FUTURE WORK

This paper presented an automated deep learning-based framework for detecting Personal Protective Equipment (PPE) using the YOLOv8 object detection model. The proposed system integrates image preprocessing, real-time object detection, and PPE classification into a unified pipeline. By leveraging the fast inference speed and high accuracy of YOLOv8, the system effectively identifies essential safety equipment such as helmets, masks, gloves, and safety vests in industrial environments. The experimental results demonstrate that the proposed approach significantly reduces manual monitoring efforts while improving workplace safety compliance.

The proposed PPE detection system shows strong potential for real-world deployment in construction sites, manufacturing units, and other safety-critical environments. The model performs reliably under varying lighting conditions and complex backgrounds, making it suitable for continuous surveillance applications. The achieved results validate the effectiveness of deep learning-based object detection techniques in enhancing occupational safety and reducing accident risks.

As part of future work, the system can be extended by incorporating additional PPE categories and hazardous activity detection. Integration with IoT-based alert systems and access control mechanisms can further improve safety enforcement. Deployment on edge devices such as embedded systems and CCTV cameras is also planned to enable scalable and low-latency real-time monitoring. Additionally, expanding the dataset with more diverse samples can further enhance model robustness and detection accuracy.

VII. REFERENECES

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.
- [2] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *Proc. IEEE CVPR*, pp. 7263–7271, 2017.
- [3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.
- [4] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint arXiv:2004.10934, 2020.
- [5] G. Jocher et al., "YOLOv5," GitHub Repository, 2020. [Online]. Available: <https://github.com/ultralytics/yolov5>
- [6] C. Wang, A. Bochkovskiy, and H. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," arXiv:2207.02696, 2022.
- [7] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8: Ultralytics Next-Generation Object Detection," Ultralytics Documentation, 2023.
- [8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE TPAMI*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [9] W. Liu et al., "SSD: Single Shot MultiBox Detector," *Proc. ECCV*, pp. 21–37, 2016.
- [10] L. Wang, X. Zhang, and H. Chen, "Deep Learning-Based PPE Detection for Construction Safety," *Automation in Construction*, vol. 112, 2020.
- [11] Y. Fang, H. Ding, and P. Luo, "Vision-Based PPE Detection Using Deep Neural Networks," *Safety Science*, vol. 120, pp. 183–191, 2019.
- [12] S. Li, Z. Zhu, and X. Li, "Helmet Detection in Construction Sites Using Deep Learning," *IEEE Access*, vol. 8, pp. 144629–144639, 2020.
- [13] K. Kim and J. Cho, "Real-Time Safety Helmet Detection Using YOLO-Based Architecture," *Sensors*, vol. 21, no.

3, 2021.

- [14] M. Nath, A. Hazarika, and K. Das, "Personal Protective Equipment Detection Using CNN and YOLO," *Proc. Int. Conf. Smart Systems and Inventive Technology*, pp. 856–861, 2021.
- [15] A. J. Pérez, J. C. Pérez, and R. Molina, "Deep Learning-Based PPE Compliance Monitoring System," *Journal of Industrial Information Integration*, vol. 24, 2021.
- [16] R. Girshick, "Fast R-CNN," *Proc. IEEE ICCV*, pp. 1440–1448, 2015.
- [17] T. Lin et al., "Microsoft COCO: Common Objects in Context," *Proc. ECCV*, pp. 740–755, 2014.
- [18] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," *Proc. IEEE*, vol. 107, no. 8, pp. 1548–1566, 2019.
- [19] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [21] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *Proc. IEEE CVPR*, pp. 3431–3440, 2015.
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proc. IEEE CVPR*, pp. 770–778, 2016.
- [23] A. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv:1704.04861, 2017.
- [24] S. Huang, Z. Li, and Y. Wang, "Vision-Based Safety Monitoring Using Deep Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, 2021.
- [25] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *Proc. IEEE CVPR*, pp. 886–893, 2005.
- [26] M. Everingham et al., "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.
- [27] A. Gupta and R. Jain, "IoT-Enabled Smart Safety Monitoring System Using Computer Vision," *Future Generation Computer Systems*, vol. 107, pp. 402–412, 2020.
- [28] S. Sreeraj and P. Kumar, "Automated PPE Detection in Industrial Environments Using Deep Learning," *Procedia Computer Science*, vol. 171, pp. 1050–1057, 2020.
- [29] M. Zhao, L. Li, and Y. Zhang, "Real-Time Worker Safety Monitoring Using Vision-Based Methods," *IEEE Access*, vol. 9, pp. 126345–126356, 2021.
- [30] R. S. Kumar and V. Rajesh, "Deep Learning- Based Intelligent Safety Monitoring System for Industrial Applications," *Journal of Safety Science and Engineering*, vol. 4, no. 2, pp. 98–108, 2022.