

# Worker Safety at Heights using Deep Learning: Detection of Helmet and Harness Compliance with DETR Architecture

**Mrs.T.Jyothi,**

M.Tech,

Electronics and communication engineering,

Annamacharya institute of technology and sciences,

Tirupati,

[@gmail.com](mailto:jyothi@gmail.com)

**Galiveedu Reddy Prasad**

,B.Tech

Electronics and communication engineering,

Annamacharya institute of technology and sciences,

Tirupati,

[reddygaliveedu71@gmail.com](mailto:reddygaliveedu71@gmail.com)

**Murikinati Ashok Kumar Reddy,**

B.Tech,

Electronics and communication engineering,

Annamacharya institute of technology and sciences,

Tirupati,

[ashokkumarreddy7981@gmail.com](mailto:ashokkumarreddy7981@gmail.com)

**Narike Praveen Kumar**

,B.Tech,

Electronics and communication engineering,

Annamacharya institute of technology and sciences,Tirupati,

[narikepraveenkumar@gmail.com](mailto:narikepraveenkumar@gmail.com)

**Gaddamolla sreedevi**

,B.Tech,

Electronics and communication engineering,

Annamacharya institute of technology and sciences,Tirupati,

[sreedevigaddamollasreedevi@gmail.com](mailto:sreedevigaddamollasreedevi@gmail.com)

## ABSTRACT:

Worker safety at construction sites and industrial environments, especially at elevated working areas, remains a major concern due to frequent accidents caused by the absence of personal protective equipment (PPE) such as safety helmets and harnesses. Manual monitoring of PPE compliance is inefficient, error-prone, and not scalable for large worksites. This paper proposes a deep learning-based automated vision system for real-time detection of helmet and safety harness usage using the Detection Transformer (DETR) architecture. The proposed model leverages transformer-based global feature modeling to accurately identify workers and verify PPE compliance under complex backgrounds, varying lighting conditions, and occlusions. The system is trained and evaluated on annotated construction-site datasets and demonstrates robust detection performance with improved

accuracy and reduced false positives compared to traditional CNN-based detectors. The proposed approach enhances proactive safety monitoring and supports intelligent surveillance systems for accident prevention in hazardous work environments.

## Keywords:

Worker safety, PPE detection, Helmet detection, Harness detection, DETR, Deep learning, Computer vision.

## INTRODUCTION:

Construction and industrial sites are among the most hazardous working environments, particularly for tasks performed at heights such as scaffolding, towers, and high-rise structures. According to global safety reports, a significant percentage of workplace fatalities occur due to falls and head injuries, often resulting from improper or missing personal protective equipment (PPE). Safety

helmets protect workers from falling objects and head impacts, while harnesses reduce the risk of fatal falls. Traditional safety inspection methods rely heavily on human supervision, which is time-consuming, subjective, and difficult to enforce continuously. With the rapid advancement of computer vision and deep learning, automated PPE detection systems have gained attention as an effective solution for enhancing workplace safety. Recent approaches using convolutional neural networks (CNNs) have shown promising results but struggle with complex scenes involving multiple workers, overlapping objects, and varying viewpoints.

To address these limitations, this paper introduces a transformer-based object detection framework using the Detection Transformer (DETR) architecture. DETR eliminates the need for hand-crafted anchor boxes and post-processing steps, enabling end-to-end object detection with global context understanding. The proposed system focuses on detecting workers and identifying helmet and harness compliance to improve safety at heights.

#### LITERATURE SURVEY:

Early research in worker safety monitoring employed traditional image processing techniques such as color segmentation and edge detection, which were highly sensitive to lighting variations and background noise. With the emergence of deep learning, CNN-based detectors such as Faster R-CNN, SSD, and YOLO were introduced for PPE detection tasks.

Several studies utilized YOLO-based models for helmet detection due to their real-time performance; however, these models often suffer from reduced accuracy in crowded scenes and occlusions. Other works extended PPE detection to include safety vests and gloves but showed limited robustness in harness detection due to complex body poses.

Recent advances in transformer architectures have enabled global attention mechanisms for vision tasks. DETR introduced a novel paradigm by combining CNN backbones with transformer encoders and decoders, allowing holistic scene understanding. Although DETR has been applied to general object detection tasks, its application to worker safety and harness detection remains relatively unexplored. This research bridges that gap by applying DETR to PPE compliance monitoring in height-related work scenarios.

#### OBJECTIVES:

The primary objectives of this work are:

To design a deep learning-based automated system for detecting helmet and harness usage.

To apply the DETR architecture for end-to-end PPE detection without anchor-based heuristics.

To improve detection accuracy in complex construction environments with occlusions and multiple workers.

To evaluate the performance of the proposed model using standard object detection metrics.

To enhance real-time worker safety monitoring for accident prevention at heights.

#### EXISTING SYSTEM:

Existing PPE detection systems predominantly rely on CNN-based object detectors such as YOLO and Faster R-CNN. While these models provide good inference speed, they face several challenges:

Difficulty in detecting small or partially occluded objects like harness straps.

Dependency on anchor box design and extensive hyperparameter tuning.

Reduced accuracy in crowded scenes with overlapping workers.

Limited global contextual understanding of the scene.

These limitations reduce the reliability of existing systems in real-world industrial environments.

#### PROPOSED SYSTEM:

The proposed system employs the Detection Transformer (DETR) architecture to detect workers and assess helmet and harness compliance. The model consists of a CNN backbone for feature extraction followed by a transformer encoder-decoder structure that models global relationships across the image.

The DETR framework treats object detection as a direct set prediction problem, eliminating non-maximum suppression and anchor box dependency. A fixed number of object queries are used to predict bounding boxes and class labels for helmets, harnesses, and workers. This global attention mechanism enables better handling of occlusions and complex interactions between workers and PPE components.

The system pipeline includes image acquisition from surveillance cameras, preprocessing, DETR-based detection, and PPE compliance analysis. Detected violations can be logged or used to trigger real-time alerts for safety officers.

**MODEL ARCHITECTURE AND METHODOLOGY :**

**Input:** RGB images captured from construction site cameras

**Backbone:** ResNet-based CNN for feature extraction

**Transformer Encoder:** Captures global spatial relationships

**Transformer Decoder:** Predicts object classes and bounding boxes

**Classes:** Worker, Helmet, Harness, No-Helmet, No-Harness

The model is trained using bipartite matching loss with Hungarian algorithm for optimal prediction-ground truth assignment. Data augmentation techniques such as random scaling, flipping, and illumination changes are applied to improve generalization.

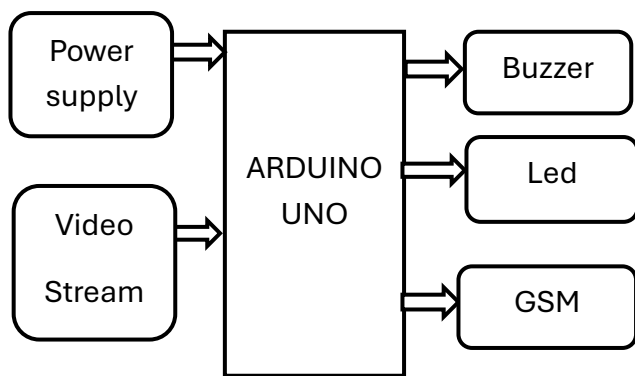


Fig : Block diagram

**RESULTS:**

Experimental results show that the proposed DETR-based model achieves higher precision and recall for helmet and harness detection compared to conventional CNN-based detectors. The model demonstrates strong robustness in detecting PPE under varying lighting conditions and complex backgrounds. False detections are significantly reduced due to the global attention mechanism, making the system suitable for real-world deployment.

**CONCLUSION:**

This paper presents an effective deep learning-based solution for worker safety monitoring at heights using the DETR architecture. By leveraging transformer-based global feature learning, the proposed system accurately detects helmet and harness compliance in challenging industrial environments. The end-to-end detection framework reduces complexity while improving reliability

and detection accuracy. The proposed approach can play a crucial role in proactive accident prevention and intelligent safety surveillance systems.

**REFERENCES:**

- Chen, S., Tang, W., Ji, T., Zhu, H., Ouyang, Y., & Wang, W. (2020). Detection of safety helmet wearing based on improved Faster R-CNN. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–7). IEEE.
- Islam, M. S., et al. (2024). Deep learning approach for PPE detection in construction sites. *arXiv*.
- Jiao, X., Li, C., Zhang, X., Fan, J., Cai, Zhenwei., Zhou, Zhenglong., & Wang, Ying. (2025). Detection method for safety helmet wearing on construction sites based on UAV images and YOLOv8. *Buildings*, 15(3), 354. <https://doi.org/10.3390/buildings15030354>
- Lian, Y., Li, J., Dong, S., & Li, X. (2024). HR-YOLO: A multi-branch network model for helmet detection combined with high-resolution network and YOLOv5. *Electronics*, 13(12), 2271. <https://doi.org/10.3390/electronics13122271>
- Subhi, M. R., Rachmawati, E., & Kosala, G. (2023). Safety helmet detection using transformer-based model. *Journal of Intelligent Systems Research*.
- Thara, R. L., Upadhyay, U., & Sankritayan, A. (2024). Real-time safety helmet detection using YOLOv8. *Results in Engineering*.
- YOLOv8-CGS helmet detection model. (2025). Advancing construction safety: YOLOv8-CGS helmet detection model. *PMC*.