Helmet Detection Using YOLOv8 and Detection Transformer

E. Bhanu[1], Kamaluru Mahammad[2], P. Jaagruthi[3], Sugali Harijawahar Naik[4], Yellanur Amruth[5] and Gummadi Gnanesh[6]

{21691A3121@mits.ac.in[1], mahammadk@mits.ac.in[2], 21691A3165@mits.ac.in[3], 21691A3155@mits.ac.in[4], 21691A3111@mits.ac.in[5], 21691A3148@mits.ac.in[6]}

UG Student, Department of Artificial Intelligence, Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India[1, 3, 4, 5, 6]
Assistant Professor, Department of Artificial Intelligence, Madanapalle Institute of Technology & Science, Madanapalle, Andhra Pradesh, India[2]

Abstract. Today's traffic systems including motorcycle safety is taken as the most important issue in reducing the number of deaths and injuries by wearing a helmet while in an accident. In this paper, a deep learning based smart helmet detection system is proposed with two state of the art object detection networks YOLOv8 and the Detection Transformer (DETR). The system is real-time, and it serves for law enforcement and traffic control departments. We train the models based on a large dataset of traffic surveillance photos in diverse environmental and situational conditions, such as different light, rider orientation, upper body occlusions. YOLOv8 is used due to its high quality and fast detection times, making it suitable for real-time operation. On the other hand we choose DETR to use instead for its strong scene context model which delivers superior detection accuracy in hard contexts. The evaluation results illustrate the trade-off between the models: DETR offers more accurate detection in challenging visual conditions, while YOLOv8 provides more efficient processing, making it suitable for dynamic surveillance. The results showed that both models could be embedded in intelligent transportation systems to enhance helmet compliance and safe driving behavior.

Keywords: computer vision, YOLOv8, DETR, deep learning, object detection, helmet detection, road safety, traffic surveillance.

1 Introduction

In motorcycle accidents generally result in catastrophic injuries, simply because there is so little to protect riders in the event of a collision. Therefore, smart helmet detection systems have become a necessity. The distance-based helmet detection algorithms have developed greatly with deep learning techniques such as transformer-based architectures and CNNs to improve scalability and accuracy of detections.

he challenging helmet detection problem involves determining whether riders are wearing helmets under various lighting conditions, occlusions, or different perspectives. Classical methods, such as SVM and Haar cascades, suffer from poor generalization. since they depend on handcrafted features. However, the evolution of deep learning models (e.g., YOLO (You Only Look Once), SSD, and Faster R-CNN) has significantly improved detection, making

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2357978

these models the dominant methods for object detection. One of them is YOLO which is popular because of its accuracy and fast inference. The latest one, YOLOv8, carries on increasing the detection efficiency to combine modifications to the loss function, feature pyramid networks, and backbone. DETR (Detection Transformer) integrates self-attention to capture long-range dependencies and improves object detection in complex scenes. DETR is designed to clean up the detection pipeline by removing the requirement for anchor boxes and is capable of better localization accuracy than traditional object detectors. It results in increased detection accuracy and a notable decrease in false detections, especially in crowded scenes. The aim of the proposed research is to strengthen road safety enforcement in order to reduce deaths related to failure of helmet use, by integrating YOLOv8 for speed and DETR for accuracy.

2 Related works

Han et al. (2024) proposed YOLOv8s-SNC, an architecture-enhanced version of YOLOv8 targeting helmet detection. Their architecture is designed to minimize redundant computation by including structural normalization and convolutional improvements. Experimental results demonstrated that the proposed model enhanced accuracy and efficiency, while keeping lightweight on the model; therefore, being it suitable for real-time construction safety applications [1]. For control of industrial environments with complex dynamics, Wang et al. (2024) proposed a unified detection and attention mechanism model with adaptive spatial and channel attention modules, named YOLOv8-ADSC. The above model was further tested in industrial worksite and it can offer an improved detection accuracy even under cluttered background such as machinery cluttered background or poor light conditions. The adaptability of YOLOv8 was shown in this work along context-aware improvements [2].

Wu et al. (2025) developed YOLOv8-CGS for construction safety monitoring. They employed a combination of YOLOv8 and graph theory techniques for better robustness in some realistic actual construction site. The approach showed both a significant false positives reduction at high inference speed by running YOLOv8-CGS, meaning that it should be useful for practical purposes [3]. Fu et al. (2024) addressed the specific difficulties of UGME by proposing YOLOv8n-ADS. The system enhanced the performance of helmet detection for mines in dusty, dark and small spaces. Their modifications include specialized anchor-free detectors and multi-scale feature concatenation to enhance robust detection in challenging situations [4].

Fan et al. (2024) progressive helmet detection proposed LG-YOLOv8 which integrated lightweight modules and feature enhance strategies to YOLO-DNN. They reduced the computation required by the model, and were able to retain high accuracy, which made such real-time helmet recognition on an embedded device (such as surveillance camera) possible. This work focused on the trade-off between energy efficiency and correctness regarding safety critical applications [5]. Liu et al. (2025) built the LSH-YOLO, a lightweight helmet detection model that emphasizes on speed-critical safety supervision. Through simplifying the structure based on structural simplification methods and hardware-friendly layers, their model was within a hair's breadth of state-of-the-art accuracy with significantly less latency. It positions itself as a candidate to address safety monitoring concerns in smart city and industrial IoT applications [6].

Jiao et al. (2025) studied helmet detection using YOLOv8 from UAVs, showing that aerial monitoring can make the safety surveillance enlarged to the large construction sites. They were able to bring YOLOv8 to the context of high-resolution drone imagery, for large-area helmet compliance monitoring with practically zero manual verification. This paper expands the YOLO models for safety inspection with officers/fixed camera from high resolution and perspective operation [7]. Shoman et al. (2024) concentrated on traffic monitoring-based detection of helmet violation with the help of YOLOv and their corresponding deep convolutional GAN generated images. The GAN-based augmentation significantly increased the training data and better generalized across a wide range of motorcycle riding conditions. Their work validated the importance of synthetic data to counter real-world imbalantages [8].

Liu et al. (2024) proposed DST-DETR, which integrates transformer-based detection and YOLO-style strategy. They integrated dehazing with RT-DETR to improve visibility in fog scenes, which effectively improved the detection performance under extreme conditions. This work demonstrated that transformers can assist YOLO in low quality detections by environmental degradation [9]. Chen et al. (2025)) built also a transformer-based for helmet detection with deformable attention mechanism. Through the use of the deformable transformers, our model has stronger ability feature extraction with various scales and occlusions. Their findings confirmed that transformer-based models can remedy YOLO's inability to model features of tracking helmet compliance [10].

Prakash-Borah et al. (2024), was also an extension on the helmet detection study that combined YOLOv8 to detect helmets and automatically extracted number plates. This practical two-fold system created compliance with the helmet law and also provided motor vehicle law enforcement with a link between ticketed violators and specific motorcycle operators. The work emphasizes significance of multimodal system in urban safety and coverage monitoring [11]. Deng et al. (2024) proposed Helmet Net, which is the extension of YOLOv8 to detect helmets. Their architecture stacked residual modules together with newly added up sampling layers, achieving better accuracy and fewer computation costs. This experiment also shows that the YOLOv8 is efficient and robust with a particular task in helmet detection [12].

Af-yolov8 (chen and wang, 2025) is a lighter detection with simple attention fusion. Their study demonstrated that AF-YOLOv8 was suitable for the efficient and accurate detection of helmets and reflective vests. This article shows how the YOLOv8 models can be extended from helmets to general personal protective equipment distributions [13]. Liu et al. (2024) offered an improved YOLOv5 framework in Scientific Reports for helmet detection. Their contributions to feature pyramid networks and the design of an activation function resulted in stable detection in a range working environment conditions. This confirms the implementation of deep learning in the real helmet detection for industrial safety field [14].

Wei et al. (2024) continued that the stability of diversified environment in such deep learning-based helmet detection research had been succeeded. They made use of large dataset, implemented optimization-based detection using YOLO to obtain higher accuracy and robustness. Indeed, their work reveals that the performance improvement is quite dependent on dataset cleanliness and model fine-tuning [15]. Lin (2024) fine-tuned YOLOv8 for helmet detection in IEEE Access by improving its backbone and feature extraction layers. They reported significant improvement in construction safety applications and demonstrated the

YOLOv8 as a powerful backbone for high-precision safety applications [16].

Cheng (2024) proposed a powerful helmet detection network based on YOLOv8 combined with the transformer modules. Their incorporation enhanced the generalization capability of the model on addressing occlusions, head poses and more realistic scenarios. This mixture design further attributes YOLOv8- Transformer to be applicable in an object detection task for safety-critical applications [17]. Zhou et al. (2024) introduced the Helmet-YOLO model, a high-precision real-time detection model. Their results showed that the verification speed and detection accuracy had been significantly improved, Helmet-YOLO become a competitor of traditional YOLOv8 models for real-time CSMS application [18].

Baoju et al. (2025) work on helmet identification at heavy machinery plants, where environment complexity and occlusion are the main challenges. They demonstrated that YOLO family detectors were suitable for industrial-scale applications including enforcement of safety protocol and regulation at on-site environment [19]. Finally, Sudharsanan et al. (2025) compares YOLOv8 with older detectors such as Faster R-CNN, and YOLOv3. Comparison results showed the YOLOv8 was faster and more accurate, which also confirmed it state of the art in real-time helmet detection for ITS and workplace safety. [20]).

3 Methodology

3.1 Dataset and Preprocessing

The dataset consists of images depicting helmeted and non-helmeted riders, collected from public repositories and real-world traffic footage. Preprocessing steps include:

- **Image Resizing**: Standardizing dimensions to 640×640 pixels.
- **Data Augmentation**: Applying transformations like flipping, rotation, brightness variation, and contrast normalization.
- **Noise Reduction**: Utilizing Gaussian filtering to improve image quality.
- **Bounding Box Refinement:** Enhancing annotations for better localization accuracy.

3.2 Model Architectures

3.2.1 YOLOV8 Architecture

The YOLOv8 is a state-of-the-art real-time object detection model which is a part of the YOLO (You Only Look Once) family. It comes with many new features. The architecture leverage a decoupled head layout that decouples the classification and regression tasks, leading to increase in precision as well as speed. The backbone of KCNet is a lighter and more efficient one which is employed to enrich feature extraction, and the neck is a Path Aggregation Network (PAN) that integrate features at different scales. YOLOv8 introduces anchor-free detection, leading to a simpler architecture that achieves better performance across objects of various sizes. Owing to the end-to-end structure, high accurate with less computational operation is obtained, which is suitable for real-time application.

3.2.2 DETR Architecture

The Detection Transformer (DETR) is a new method of object detection that incorporates Transformer architecture directly in the detection pipeline. Unlike traditional object detection model architectures or anchor boxes, DETR describes object detection as a direct set prediction problem. It includes a convolution network (usually ResNet) to extract features, and relies on a Transformer encoder-decoder to process the features globally. The decoder utilizes a fixed number of learnable object queries to attend over image regions and to perform multi-task prediction of bounding boxes and class labels. Based on this, the attention-based mechanism of DETR can bring long-range dependencies and contextual information, so as to achieve the better object locations and categories, especially in the complex scenes. Fig 1 shows the architecture diagram of combination of YOLO and DETR.

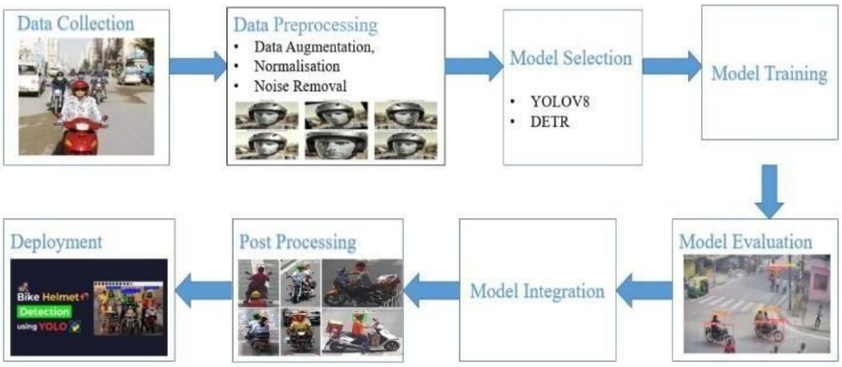

Fig. 1. Architecture Diagram of Combination of YOLO and DETR.

3.2.3 Algorithm

1. Data Gathering: Collect data from surveillance cameras and online resources.
2. Pre-processing: Resizing, augmentation, and noise reduction are utilized.
3. Feature Extraction: adopt the CNN backbone from YOLOv8, and the transformer-based encoder from DETR for feature extraction.
4. Object Detection: YOLOv8 fast detects helmets, and DETR fuses localization.
5. Post-Processing: Use NMS to remove the redundant detections.
6. Going back up: Classification: Decide if helmet is present using confidence scores.
7. Performance Evaluation: CALCULATE precision, recall, and map.
8. Deployment: Deploy the learned model in a real aerial traffic surveillance system.

3.2.4 Training and Evaluation

The training stage has a process pipeline as:

- Augmentation techniques used to ensure dataset diversity.

- Feature extraction: YOLOv8 proposes helmets, DETR then refines the classes.
- Object Detection & Filtering: Non-Maximum Suppression (NMS) is used to discard the redundant bounding boxes.
- Hyperparameter Optimization: Adam optimizer finetune Learning Rate and weight decay.
- Characteristics of performance: Precision, recall, F1-score, map, and inference speed are investigated.

4 Results and Evaluation

4.1 Figures and Tables

A comparative analysis of YOLOv8 and DETR can be summarized in terms of their strengths. The results are presented below:

Lay outing Figures: Fig 2, 3, 4 and table 1 are fundamental to help to understand and summarize the results in helmet detection with YOLOv8 and DETR.

Fig. 2. Sample results of helmet detection using YOLOv8 and DETR at different angles.

Fig. 3. Sample results of helmet detection using YOLOv8 and DETR with low light.

Fig. 4. Sample results of helmet detection using YOLOv8 and DETR.

4.1.1 Performance Comparisons of YOLOv8 and DETR

Extensive experimental results were conducted to compare the YOLOv8 with Detection Transformer (DETR) for the helmet detection problem. Models were evaluated in terms of precision, recall, mAP@50 and the inference time. These measures represent a trade-off between the detection and real-time ability of the models. These evaluation results summarized in the following table 1.

Table 1. Performance Comparison of YOLO and DETR Object Detection Models.

Model	Precision (%)	Recall (%)	mAP@50 (%)	Processing Time (ms)
YOLO	85	82	83	25
DETR	88	83	85	50

4.2 Model Performance Optimization

This section discusses methods to improve the accuracy and efficiency of the proposed helmet recognition system.

- Hyperparameter Tuning: Tune learning rate, batch size, and weight decay for model convergence.
- Feature Fusion: Leveraging of LLFMs and HLFMs for better robustness in detection.

- Attention Mechanisms: Investigating self-attention layers in order to refine both object localization and classification. Post-Processing Improvements: Applying state-of-the-art Non-Maximum Suppression (NMS) techniques to reduce false positives.
- Quantization & Pruning: Compact models, low computational cost and a high level of accuracy.

The primary goal of this optimization is to improve the efficiency of real-time helmet detection across different conditions using YOLOv8 and DETR.

5 Conclusion and Future Scope

The focus of this work is automatic helmet detection based on YOLOv8-DETR network which can take the advantage of the superiority of YOLO and the strengths of DETR, and finally realize the trade-off between real-time operation and accuracy. YOLOv8 is both light weight and optimized so the inference is fast, making it ideal for real time helmet monitoring. Meanwhile, DETR adopts a transformer-based algorithm to achieve higher object detection accuracy, especially in challenging cases including occlusion and overlapped multiple objects.

It is validated from the comparative analysis of helmet detection datasets that YOLOv8 achieves optimal speed performance which are affordable for quick helmet enforcement application. In contrast, DETR has higher precision performance, which can greatly improve its ability to detect helmet wearing in difficult situations e.g. lighting differences and background clutter. By integration of these models, we proposed to improve the safety of the road by means of automated helmet date citing systems.

5.1 Future Scope

The performance, scalability and adaptability of the new system can be further enhanced by some extensions that can be exploited as future work:

- Enhancement of the Training Data – Increasing the dataset size with more image variety of helmet designs (different graphic designs, style, and color), rider demographics, and environmental details leads to the improvement of model generalization.
- Advanced Data Augmentation – Incorporating methods like GAN generated synthetic data, mix-up augmentation and domain adaptation, can breed even more robustness in real-world scenarios.
- Multiple Camera Integration – Using several cameras in sync allowing to a wider area of monitoring of watchlist motorcycle individuals from different perspectives resulting in more consistent detection.
- Interfacing with Traffic Surveillance Systems – The helmet detection model can be interfaced with the smart traffic control systems, it automatically identifies violators and assists the authorities in enforcing the road safety rules.

- OCR-Powered License Plate Recognition – By adding Optical Character Recognition (OCR) to retrieve license plate information, we can automatically log violations and enforce the law.
- Hybrid AI Model – Integrating the CNN-based and the transformer-based structures in a single framework may be able to maximize the computational efficiency without loss of detection performance.
- Policy and Law Enforcement Implications – Partnerships with government can aide in implementation of this technology in monitoring large scale and assist in minimizing road traffic accidents and helmet compliance.
- By tackling these issues, the system can eventually mature into an autonomous, on-line helmet detection system assisting the law enforcement as well as inculcating safer ride behaviour among two-wheeler riders.

References

[1] Han, D., Ying, C., Tian, Z., Dong, Y., Chen, L., Wu, X., & Jiang, Z. (2024). YOLOv8s-SNC: An Improved Safety-Helmet-Wearing Detection Algorithm Based on YOLOv8. *Buildings*, *14*(12), 3883. https://doi.org/10.3390/buildings14123883

[2] Wang, J., Sang, B., Zhang, B., & Liu, W. (2024). A Safety Helmet Detection Model Based on YOLOv8-ADSC in Complex Working Environments. Electronics, 13(23), 4589. https://doi.org/10.3390/electronics13234589

[3] Wu Z, Lei X, Kumar M (2025) Advancing construction safety: YOLOv8-CGS helmet detection model. PLoS One 20(5): e0321713. https://doi.org/10.1371/journal.pone.0321713

[4] Fu, Z., Ling, J., Yuan, X., Li, H., Li, H., & Li, Y. (2024). Yolov8n-FADS: A Study for Enhancing Miners' Helmet Detection Accuracy in Complex Underground Environments. Sensors, 24(12), 3767. https://doi.org/10.3390/s24123767

[5] Fan, Z., Wu, Y., Liu, W., Chen, M., & Qiu, Z. (2024). LG-YOLOv8: A Lightweight Safety Helmet Detection Algorithm Combined with Feature Enhancement. Applied Sciences, 14(22), 10141. https://doi.org/10.3390/app142210141

[6] Liu, Z., Wang, F., Wang, W., Cao, S., Gao, X., & Chen, M. (2025). LSH-YOLO: A Lightweight Algorithm for Helmet-Wear Detection. Buildings, 15(16), 2918. https://doi.org/10.3390/buildings15162918

[7] Jiao, X., Li, C., Zhang, X., Fan, J., Cai, Z., Zhou, Z., & Wang, Y. (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

[8] Shoman, M., Ghoul, T., Lanzaro, G., Alsharif, T., Gargoum, S., & Sayed, T. (2024). Enforcing Traffic Safety: A Deep Learning Approach for Detecting Motorcyclists' Helmet Violations Using YOLOv8 and Deep Convolutional Generative Adversarial Network-Generated Images. Algorithms, 17(5), 202. https://doi.org/10.3390/a17050202

[9] Liu Z, Sun C, Wang X. DST-DETR: Image Dehazing RT-DETR for Safety Helmet Detection in Foggy Weather. Sensors (Basel). 2024 Jul 17;24(14):4628. doi: 10.3390/s24144628.

[10] Chen, S., Sun, H., Wu, Y., Shang, L., & Ruan, X. (2025). A helmet detection algorithm based on transformers with deformable attention module. *Chinese Journal of Electronics, 34*(1), 229–241. https://doi.org/10.23919/cje.2023.00.346

[11] Prakash-Borah, J., Devnani, P., Kumar-Das, S., Vetagiri, A., & Pakray, P. (2024). Real-time helmet detection and number plate extraction using computer vision. *Computación y Sistemas, 28*(1). https://doi.org/10.13053/cys-28-1-4906

[12] Deng, L., Zhou, J. & Liu, Q. Helmet Net: An Improved YOLOv8 Algorithm for Helmet Wearing Detection. Int J Netw Distrib Comput 12, 329–343 (2024). https://doi.org/10.1007/s44227-024-00040-1

[13] Chen, X., Wang, G. AF-YOLOv8: a lightweight algorithm suitable for helmet and vest

detection. J Supercomput 81, 996 (2025). https://doi.org/10.1007/s11227-025-07496-0
[14] Liu, Y., Jiang, B., He, H. et al. Helmet wearing detection algorithm based on improved YOLOv5. Sci Rep 14, 8768 (2024). https://doi.org/10.1038/s41598-024-58800-6
[15] Wei, L., Liu, P., Ren, H. et al. Research on helmet wearing detection method based on deep learning. Sci Rep 14, 7010 (2024). https://doi.org/10.1038/s41598-024-57433-z
[16] Lin, B. (2024). Safety helmet detection based on improved YOLOv8. *IEEE Access, 12,* 28260–28272. https://doi.org/10.1109/ACCESS.2024.3368161
[17] L. Cheng, "A Highly Robust Helmet Detection Algorithm Based on YOLO V8 and Transformer," in IEEE Access, vol. 12, pp. 130693-130705, 2024, doi: 10.1109/ACCESS.2024.3459591.
[18] Zhou, S., Peng, Z., Zhang, H., Hu, Q., Lu, H., & Zhang, Z. (2024). Helmet-YOLO: A new method for real-time, high-precision helmet wearing detection. *IEEE Access.* Advance online publication. https://doi.org/10.1109/ACCESS.2024.3443146
[19] Baoju, L., Xiangqian, W., Qingshan, C. et al. Safety helmet detection methods in heavy machinery factory. Sci Rep 15, 18565 (2025). https://doi.org/10.1038/s41598-025-02824-z
[20] Sudharsanan, R., Sanjay, A., Raghul, P., Hayden Joseph, D., Gopirajan, P.V., Suresh Kumar, K. (2025). Comparative Analysis of YOLOv8 and Legacy Object Detection Algorithms for Real-Time Helmet Detection. In: Rathore, V.S., Murala, S., Tuba, E., Singh, M.M. (eds) Modern Practices and Trends in Expert Applications and Security. MP-TEAS 2024. Lecture Notes in Networks and Systems, vol 1379. Springer, Singapore. https://doi.org/10.1007/978-981-96-5784-1_16