Research Article

Improved Yolov5s-Based Helmet Recognition in Complex Scenes

Siva Rama Krishna Sarma, Veerubhotla[1], Kiran Kumar, Pamidi[2],

[1]Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, Guntur(D) sharmavsrk@kluniversity.in
[2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, Guntur(D) 2401050152cse@gmail.com

Abstract

Enforcing helmet use can be important in improving traffic safety and reducing head injuries, particularly in the difficult context of the real world. The traditional detection systems are often not able to find tiny objects, occlusion, and dynamic lighting, thus, lowering its reliability and efficiency. A helmet detection dataset is provided with annotated pictures of both the helmet and no-helmet categories, and was borrowed off the Helmet Detection YOLOv8 Dataset v3. When maintaining parallel inputs in the classification and detection tasks, preprocessing involves the organization of the dataset, annotations of bounding boxes in the YOLO format, normalization, and the creation of a data configuration file. Some of the deep learning models have been applied, including YOLOv5s6u, YOLOv5x6u, YOLOv8, YOLOv9, YOLOv11, and YOLOv26 to detect effectively, and Faster R-CNN as a reference. The performance is measured with precision, recall, and mAP. By having the largest mAP of 0.930 and equal precision (0.868) and recall (0.874), YOLOv5s6u is the best-performing model, showing great detection ability in difficult conditions. In the case of helmet compliance monitoring systems, the proposed approach significantly enhances detection and real time applicability.

Keywords: Helmet Detection, Object Detection, Real-Time Monitoring, Computer Vision, Deep Learning, Traffic Safety."

1. Introduction

Helmet compliance monitoring will help to enhance traffic safety and minimize the seriousness of head injuries. The use of automated vision-based monitoring systems is being implemented since the existing method of manual inspections is no longer sufficient in light of the increased urbanization and the density of vehicles. Recent developments have shown how well intelligent detection frameworks can recognize helmet use in a variety of scenarios [1], [2]. These systems aim at ensuring constant surveillance and helping the police by providing accurate and prompt information. Nevertheless, in real-world applications, due to the dynamic nature of traffic situations, the attainment of consistent performance remains a challenge [3].

Although tremendous progress has been made, the existing helmet detection technologies have extreme limitations in complex scenarios. The presence of variations in both the light, occlusion, motion blur, and crowded traffic conditions often leads to reduced detection accuracy and wrong results [4], [5]. Moreover, many systems are not able to recognize small or partially visible objects, which are common in the real-life environment. They are also not scalable and robust, which also restricts their applicability in large-scale deployment [6], [7]. These issues point to the fact that the laboratory and real-life performance are widely different and there is a need to work on more efficient and effective solutions [8]. The present research project aims at developing a smart helmet recognition system that can be effective in diverse dynamic settings to address these challenges. The main goals will be to warrant consistent real time tracking, improve generalization under diverse circumstances and improve detection accuracy. The approach focuses on the efficient object localization and robust feature representation, maintaining scalability to the real-world application. The system aims at offering dependable and predictable performance in complex traffic conditions by filling the gap between theory of development and practice [9].

The work is significant as it can enhance automated safety surveillance and provide a possibility to make well-informed decisions on traffic management. Higher levels of detection can enhance effectiveness in enforcement, reduce the chances of accidents, and promote compliance with safety regulations. The technology also helps in the development of smart surveillance systems that are able to handle the intricacies of the real world. Along with the enhancement of the efficiency of operations, these developments are also needed to establish safer transportation networks and reduce the adverse impact of traffic accidents on society [10].

2. Related Work

The goal of recent developments in helmet detection has been to increase precision and resilience in challenging situations. Wang et al. [11] introduced an updated detection system suitable in coal mine scenarios that demonstrated better performance in low-visibility scenarios. Similarly, An et al. [12] proposed a multi-scale fusion approach that employed attention processes in the attempt of enhancing the feature representation and detection accuracy. Mo et al. [13] increased detection capabilities at construction sites, which highlighted the importance of multi-object safety monitoring, by adding helmet and harness recognition. Even though these approaches raise the detection accuracy, they are not very broadly applicable as they often rely on specific areas of safety or special environments.

Benchmarking and comparison analysis to evaluate the detection performance between models has also been explored. In a comprehensive benchmarking study, Agrawal et al. [14] found that various detection structures had variations in performance. In order to achieve proactive safety-monitoring, Liu et al. [15] designed a real-time warning system that included tracking and detection technologies. Li et al. [16] focused on municipal engineering environments and devoted their attention to protective equipment detection to ensure a higher detection consistency. Nevertheless, scaling, real-time processing efficiency and flexibility to diverse traffic conditions are some of the problems that these systems often face.

Several studies have mentioned automation and lightweight design of models to be used in real-time applications. Hemalatha et al. [17] showed that an automated monitoring system in construction situations is feasible in practice. Ma et al. [18] introduced a low-weight detection framework that can operate in resource-constrained settings and offered a speedup to the inference, sacrificing only a reasonable degree of accuracy. Hao et al. [19] presented a multi-target detection approach that is capable of identifying multiple safety factors simultaneously. Despite recent developments, trade-offs between computing economy and detection accuracy remains a problem.

Other challenges that have been addressed in the recent past include long-range and small-object detection. Fan et al. [20] developed a lightweight model of long-range helmet detection, enhancing the detection performance in large environments. These techniques enhance detection in certain situations, but they still have drawbacks such poor generalization, susceptibility to environmental changes, and uneven performance in actual traffic situations.

All said and done the current practices reveal that there is a great improvement in the accuracy and efficiency of the helmet detection. Nevertheless, there are still issues with real-time scalability, resilience in complicated scenarios, and consistent performance in various environments. These shortcomings highlight the need to have an effective and more practical way of detection that would be in a position to handle the unpredictability on the real world. The present research aims to overcome these shortcomings by emphasizing more accurate detection, scalability, and adaptability that can be employed to create more useful and effective helmet compliance monitoring systems.

3. Materials And Methods

The proposed system aims to develop an accurate and real-time helmet detection system in complex situations using the Helmet Detection YOLOv8 Dataset v3 on Roboflow Universe. To facilitate correct object localization, it begins with the preparation of the dataset, which includes labeling photos in the YOLO format. Preprocessing is necessary to normalize, scale and create a data configuration file in order to assure parallel inputs that will be detected and classified consistently. The basic approach utilizes a variety of deep learning models such as the more sophisticated versions of YOLO such as YOLOv5s6u, YOLOv5x6u, YOLOv8, YOLOv9, YOLOv11, and YOLOv26 and Faster R-CNN as a reference. To effectively distinguish between helmet and no-helmet classes, these models use convolutional feature extraction and anchor-based detection techniques. To ensure the accuracy is balanced, precision, recall, and mean Average Precision are used to evaluate the performance. Effective helmet compliance monitoring in a variety of environmental situations is made possible by the system's robust small-object detection, enhanced localization, and trustworthy real-time inference.

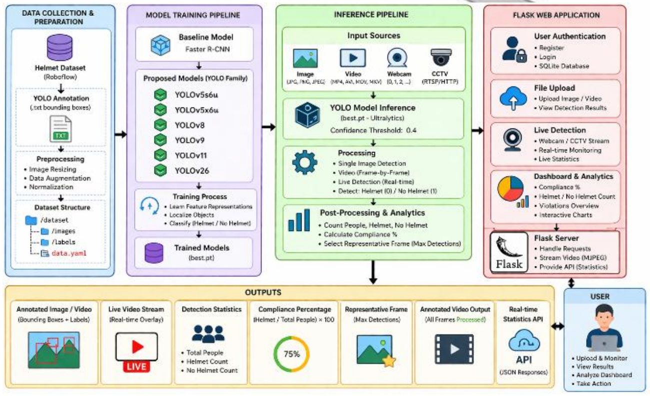

Fig. 1. System Architecture

The overall system design is represented in Fig. 1 that illustrates a structured chain of data input to the formation of the final output. It presents the stages of data gathering, pre-processing, model training and inference and integration with a web based user interface. The architecture focuses on the seamless exchange of data among different components, including input processing, result visualization, and result processing. It also shows how trained models can be used to analyze in real-time to facilitate efficient helmet detection, compliance monitoring and statistical reports using a single scalable system.

A) Dataset Collection

The dataset utilized in the Helmet Detection YOLOv8 Dataset v3 is the annotated photos about people wearing and not wearing helmets. It has a range of samples under various lighting conditions, backgrounds, and angles, as it has two classes: helmet and no helmet is the source of the dataset used for helmet detection. It has a range of samples under various lighting conditions, backgrounds, and angles, as it has two classes: helmet and no helmet. The YOLO (.txt) bounding box annotations enable accurate identification of objects. The dataset is optimal in real-time detection work in complex scenarios due to its structured annotation format and variability, which enhance the model generalization.

B) Data Preprocessing

Preprocessing guarantees consistency, improves the learning performance, and provides reliable performance to both classification as well as detection tasks in complex environments by configuring and standardizing the data by establishing configurations and systematically separating the data sets.

Data Configuration Preparation: An ordered configuration file was created to define the datasets and class location where the detection framework is to use. This step besides ensuring that training, validation and testing samples are easily obtained will ensure that data is well organized. The more the categories of the classes are well defined, the more accurate the interpretation of the labels and there is less ambiguity when learning the model. Good configuration also leads to reproducibility, and scalability to bigger data sets or new classes without impacting the workflow consistency on data handling and processing.

Parallel Dataset Structuring: Two overlapping sets of data were kept to facilitate the classification and detection. Category-level recognition is made possible by the classification dataset, which comprises of photos coupled with matching labels. Conversely, the detection dataset has photos that have corresponding bounding box annotations that accurately localize items. This systematic separation allows the system to use both global and localized information and enhance the flexibility of experimentation and evaluation. Such an approach improves the overall performance by helping to support multiple learning goals and critically analyze the results of detection.

C) Algorithms

Faster R-CNN: As a baseline detection model, Faster R-CNN generates region proposals, which are further classified and localized. In spite of the inefficiency of higher computational complexity in real-time applications, its two-step method improves the detection accuracy and feature representation.

YOLOv5s6u: The YOLOv5s6u simultaneously predicts objects and their classes to perform single-stage detection. Its improved architecture boosts small object detection and assists to attain uniform performance in challenging and dynamic environments by providing a trade-off between speed and accuracy.

YOLOv5x6u: YOLOv5x6u improves the accuracy and consistency of detection by reaching the complex spatial properties due to a deeper architecture. Although the increased complexity of the models can affect the computation speed, they can be localized and categorized in an accurate and reliable way in a large variety of conditions.

YOLOv8: YOLOv8 enables the successful localization of objects without any defined boundaries, suggesting better aggregation of the features and the best detection approach. Its simplified structure allows faster inference, higher detection in numerous challenging scenarios, and enhances generalization.

$$\mathcal{L} = \lambda_{box}\mathcal{L}_{box} + \lambda_{obj}\mathcal{L}_{obj} + \lambda_{cls}\mathcal{L}_{cls} \qquad (1)$$

YOLOv9: YOLOv9 enhances the learning of features and information, which encourages the consistency and strength of detection. It contributes to the consistent performance in dynamic contexts and also sustaining the economy of computation in managing complex patterns of visuals, and variable scale objects.

$$\mathcal{L}_{YOLOv9} = \lambda_{box}\mathcal{L}_{box} + \lambda_{obj}\mathcal{L}_{obj} + \lambda_{cls}\mathcal{L}_{cls} + \lambda_{dfl} + \mathcal{L}_{dfl} \qquad (2)$$

YOLOv11: YOLOv11 will also enhance detection efficiency and accuracy with better feature extraction and prediction techniques. It guarantees consistent performance in cases where there is occlusion, variation in scale and complex environments by aiding proper localization and categorization.

YOLOv26: YOLOv26 is an advanced detection algorithm that aims to maximize performance and scalability. It integrates well behavioral feature learning with simplified prediction, and thus, it is possible to accurately detect small and distant objects and yet flexible in various conditions.

4. Experimental Results

Precision: Precision is the proportion of samples or incidents that are correctly determined as positives. Thus, the precision can be determined by the following formula:

$$Precision = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FP}} \qquad (3)$$

Recall: Recall in machine learning is a measure used to evaluate a model that determines its ability to locate all the relevant instances of a specific class. It provides information of how much instances of a certain class a model can cover and is calculated as the ratio of those cases actually predicted to be positive to the number of actual positives.

$$Recall = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \qquad (4)$$

mAP: Mean Average Precision (MAP) is one of the ranking quality statistics. It takes into account the quantity and ranking of pertinent recommendations. The arithmetic mean of the Average Precision (AP) at K for every user or query is used to compute MAP at K.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \qquad (5)$$

Table.1 Performance Evaluation

ML Model	Precision	Recall	mAP
Faster-RCNN	0.862	0.843	0.876
YOLOv8	0.848	0.895	0.924
YOLOv5s6	0.868	0.874	0.930
YOLOv5x6	0.894	0.856	0.924
YOLOv11	0.822	0.868	0.903

Table 1 shows that YOLOv5s6 is the most reliable model to use in identifying helmets in a complex environment due to its balanced precision, recall and mAP that provides good detection accuracy and consistency.

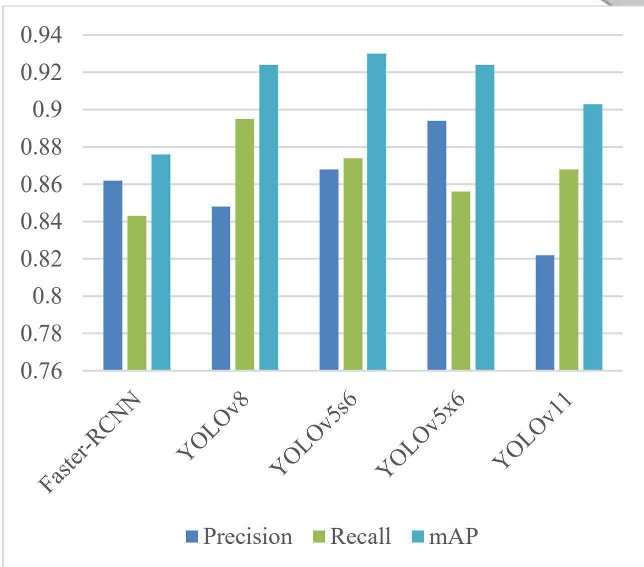

Fig.2 Comparison Graph

YOLOv5s6 is increasingly more effective than other models since it has a trend that is well-balanced in all assessment measures, as depicted in Fig. 2. This comparison graph shows that it has more stability in detection and overall performance in managing complex and diverse helmet detection environments.

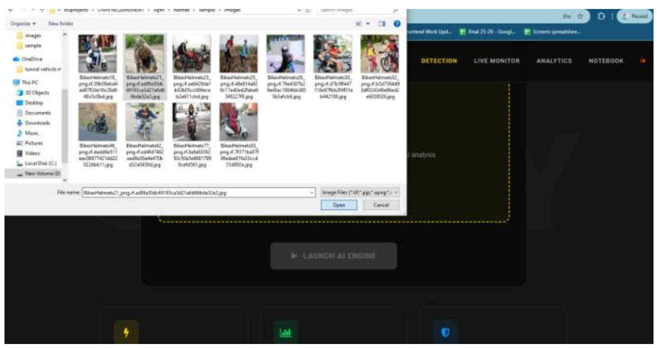

Fig.3 Upload Input Image

The interface of the upload of input photos or videos is presented in Fig. 3. This will enable users to feed data on the area of helmet detection, to ensure easy interaction and efficient data submission to further processing and analysis.

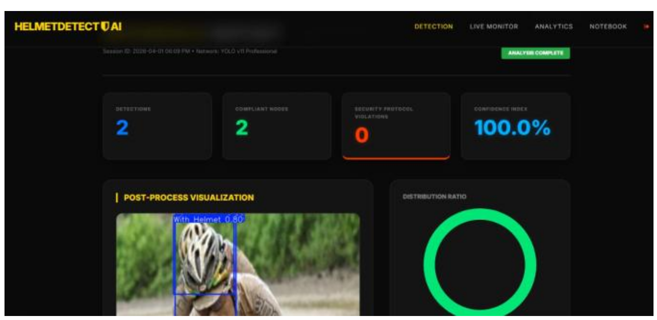

Fig.4 Detection Results – with Helmet

Fig. 4 shows the output image of two detections with full compliance and bounding boxes around helmeted individuals, confidence rating, distribution ratios, and comparable analytics to accurately monitor and evaluate.

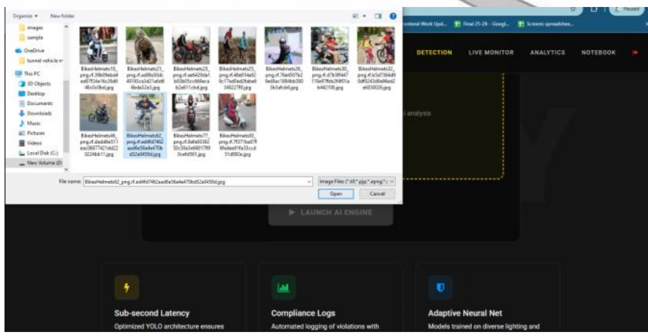

Fig.5 Upload Input Image

The input interface in Fig. 5 allows users to input photos or videos to be detected, which provides an easy-to-use interface to initiate the analysis and ensures smooth interaction with the system.

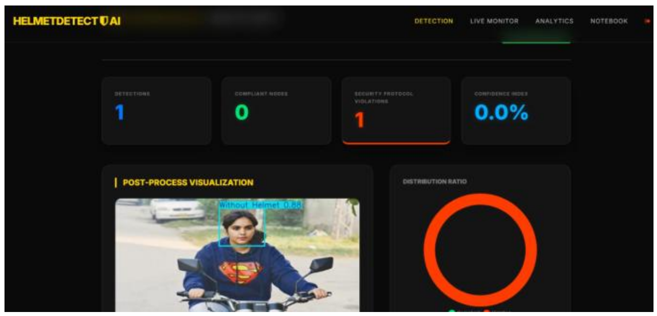

Fig.6 Detection Results – without Helmet

To achieve an efficient compliance evaluation, Fig. 6 shows the output image with a solitary finding of a violation, indicating a non-helmet case with the help of bounding boxes, as well as the confidence score, distribution ratio, and comparative analytics.

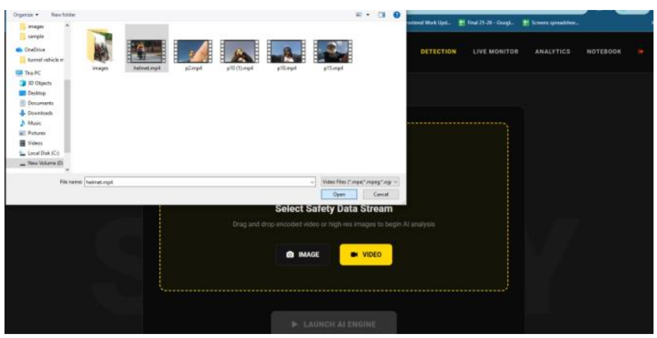

Fig.7 Upload Input Image

The uploading interface of the images and videos is presented in Fig. 7, which enables the user to upload multiple formats to be detected, and ensure that various real-world monitoring situations are handled.

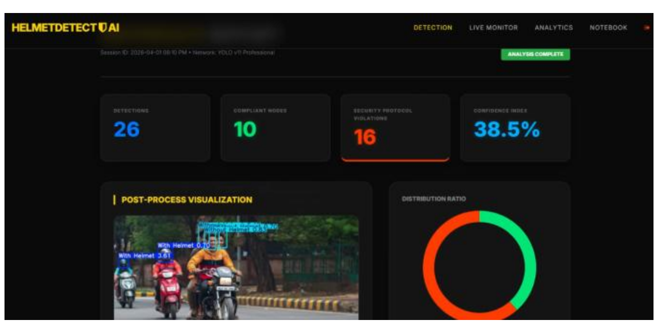

Fig.8 Detection Results

The output of the video analysis (Fig. 8) contains bounding boxes, confidence scores, distribution ratios, and comparison analytics to monitor and evaluate the results comprehensively, as well as various detections with and without compliance situations and violation situations.

5. Conclusion

The system is expected to overcome the inefficiencies in the traditional methods of vision-based approach to enhance the compliance of the helmets through offering accuracy and real-time detection in challenging environments. It utilizes the Helmet Detection YOLOv8 Dataset v3 that contains photos with annotations of the helmet and no-helmet classes. Using convolutional feature extraction and anchor-based detection for accurate localization and classification, the methodology combines several advanced YOLO versions, such as YOLOv5s6u, YOLOv5x6u, YOLOv8, YOLOv9, YOLOv11, and YOLOv26, with Faster R-CNN as a baseline. Performance evaluation involves the use of the following measures to make a fair evaluation of detecting capability; precision, recall and the mean Average Precision. With a mAP of 0.930, YOLOv5s6u outperforms the other models that have been put into practice, showing excellent overall detection performance in difficult situations. To enhance usability, real-time inference, in terms of image, video, and live-stream processing, is also combined with such statistical outputs as compliance % and detection counts. In practical traffic monitoring systems, the developed system assists in automated surveillance, improves the safety enforcement, and promotes informed decision-making by delivering precise and effective detection.

With the addition of transformer-based designs and attention mechanisms, the future works can focus on the increase in detection robustness in adverse conditions such as low light conditions, strong occlusion, and high-traffic conditions. The detection reliability may be enhanced with the addition of multimodal data, e.g., heat or depth sensors. Quantization and model trimming are some of the optimization strategies that can be employed to be deployed on edge devices. A further addition of Generalization is to introduce variety of real-world variations to the dataset, which includes the Helmet Detection YOLOv8 Dataset v3. Moreover, with the assistance of advanced tracking algorithms and behavior analysis, it will be possible to provide continuous monitoring, detect anomalies, and more intelligent traffic safety enforcement systems.

References

1. Li, Y., Shi, X., Xu, X., Zhang, H., & Yang, F. (2025). Yolov5s-PSG: Improved Yolov5s-based helmet recognition in complex scenes. IEEE Access.
2. Huang, B., Wu, S., Xiang, X., Fei, Z., Tian, S., Hu, H., & Weng, Y. (2023). An improved YOLOv5s-based helmet recognition method for electric bikes. Applied Sciences, 13(15), 8759.
3. Wang, S., Zhang, L., Ma, H., & Zhao, J. (2025). YOLOv8n-ASA: An Asymmetry-Guided Framework for Helmet-Wearing Detection in Complex Scenarios. Symmetry, 17(12), 2124.
4. Huang, B., Wu, S., Xiang, X., Fei, Z., Tian, S., Hu, H., & Weng, Y. (2023). An improved YOLOv5s-based helmet recognition method for electric bikes. Applied Sciences, 13(15), 8759.
5. Liu, J., Zhang, H., Lv, G., Liu, P., Hu, S., & Xiao, D. (2024). Helmet wearing detection algorithm based on YOLOv5s-FCW. Applied Sciences, 14(21), 9741.
6. 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.
7. Zhang, B., & Yuan, Y. (2024, August). An Improved Merge-YOLOv5 Algorithm Applied to Helmet Wear Detection. In 2024 7th International Conference on Pattern Recognition and Artificial Intelligence (PRAI) (pp. 430-436). IEEE.
8. Zhang, Y., Zhang, Y., & Huang, Z. (2026, January). Design and implementation of non-motor vehicle helmet recognition system based on YOLOv5s. In International Conference on Advanced Electronics, Intelligent Technology, and Computing (AEITC 2025) (Vol. 14010, pp. 69-74). SPIE.
9. Hu, T., Huang, W., Wei, J., Liu, Q., Huang, J., & Yu, J. (2024, September). STOE-YOLOv5s: A Hard Hat Wearing Detection Algorithm Based on Improved YOLOv5s. In 2024 6th International Conference on Robotics and Computer Vision (ICRCV) (pp. 6-11). IEEE.
10. Yonghui, L. I., Liang, Y. U. A. N., Li, H. E., Teng, R. A. N., & Kai, L. Y. U. (2024). Research on Safety Helmet Wearing Detection Algorithm in Chemical Industry Park Scenarios. Journal of Computer Engineering & Applications, 60(12).
11. Wang, Z., Mu, D., Li, Z., & Dong, F. (2026). A Safety Helmet Detection Algorithm Based on Improved YOLOv11n for Complex Coal Mine Scenarios. Digital Signal Processing, 106082.
12. An, J., Chen, W., Ni, Y., & Chang, F. (2026). RSPV-YOLO: two-wheeler helmet detection method based on residual channel attention and pixel-guided multi-scale fusion. Journal of Real-Time Image Processing, 23(1), 1-20.

13. Mo, Y., Xu, S., Ao, S., Cai, N., Zhou, S., & Zhou, Y. (2026). Safety helmet and harness detection on construction sites based on deep learning: Y. Mo et al. Signal, Image and Video Processing, 20(1), 41.
14. Agrawal, K., Patel, V. S., Cannon, I., Tran, M. T., & Nguyen, T. V. (2024, December). Motorcycle Helmet Detection Benchmarking. In International Symposium on Information and Communication Technology (pp. 203-215). Singapore: Springer Nature Singapore.
15. Liu, Y., Wu, H., & Li, H. (2026). Research on a Real-Time Warning System for Unsafe Behaviors in Hydraulic Construction Based on DeepSORT and Improved YOLOv5s. Applied Sciences, 16(6), 2960.
16. Li, S., Lv, Y., Li, M., & Wang, Z. (2023). Detection of Protective Apparatus for Municipal Engineering Construction Personnel Based on Improved YOLOv5s. International Journal of Advanced Computer Science and Applications, 14(7).
17. Hemalatha, S., Nithya, P., Pradhipa, S., Gokulraj, S. R., & Akash, A. J. (2026). Automated Construction Helmet Monitoring via YOLO Algorithm. In Applied Mathematics, Automation and Computing (pp. 109-115). CRC Press.
18. Ma, D., Yang, C., Ao, L., Shi, B., & Ma, S. (2024, May). Lightweight Motorcycle Helmet Detection Based on Improved YOLOv7-Tiny. In 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE) (pp. 1126-1132). IEEE.
19. Hao, F., Zhang, T., He, G., Dou, R., & Meng, C. (2024). CaSnLi-YOLO: construction site multi-target detection method based on improved YOLOv5s. Measurement Science and Technology, 35(8), 085202.
20. Fan, J., Guang, J., Yu, N., Chen, J., & Wang, W. (2025, August). Tiny-YOLOv8 long-distance electric vehicle helmet detection model. In 2025 5th International Conference on Computer Vision, Application and Algorithm (CVAA) (pp. 70-73). IEEE.