Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 12 NO: 3 PAGE: 773-782 YEAR: 2023
DOI:10.17798/bitlisfen.1297952

Deep Learning-Based Automatic Helmet Detection System in Construction Site Cameras

Adem KORKMAZ[1*], Mehmet Tevfik AĞDAŞ[2]

[1] Bandırma Onyedi Eylül University, Gonen Vocational School, Department of Computer Technology
[2] Munzur University, Cemisgezek Vocational School, Department of Computer Technology
(ORCID: 0000-0002-7530-7715) (ORCID: 0000-0002-5608-6240)

Keywords: Helmet detection, Object detection, YOLOv8, Personal protective equipment.

Abstract
Ensuring worker safety in high-risk environments such as construction sites is paramount. Personal protective equipment, particularly helmets, is critical in preventing severe head injuries. This study aims to develop an automated helmet detection system using the state-of-the-art YOLOv8 deep learning model to enhance real-time safety monitoring. The dataset used for the analysis consists of 16,867 images, with various data augmentation and preprocessing techniques applied to improve the model's robustness. The YOLOv8 model achieved a 96.9% mAP50 score, outperforming other deep learning models in similar studies. The results demonstrate the effectiveness of the YOLOv8 model for accurate and efficient helmet detection in construction sites, paving the way for improved safety monitoring and enforcement in the construction industry.

1. Introduction

The construction industry is one of the high-risk working environments where employee safety is a priority. Construction sites are inherently hazardous, with high accident and injury potential. Strict adherence to safety regulations, including the mandatory use of personal protective equipment (PPE), is crucial to ensuring employee safety and minimizing risks. Helmets are critical in protecting workers from head injuries that can have serious consequences, including permanent disability and even death. The rapid development of information technologies impacts every aspect of our lives and various business processes. Their automatic and quick response in detection, recognition, and decision-making processes has become integral to data management. In this era of rapidly increasing data, deep learning algorithms are essential for decision-making and process management.

Construction sites are inherently hazardous environments with high accident and injury potential [1]. Strict adherence to safety regulations, including the mandatory use of personal protective equipment (PPE), is important to ensure worker safety and minimize risks [2]. Helmets are critical in protecting workers from head injuries that can lead to serious consequences, including permanent disability and death [3]. Inadequate site inspections and low safety awareness among construction workers can lead to accidents. Real-time detection of helmet use is crucial for rapid action and prevention of such incidents [4].

As computer technology has advanced, automatic visual detection has become increasingly prevalent. The growth of deep learning-based computer vision technologies has opened new possibilities for enhancing safety monitoring and enforcement across various sectors, including construction [5]. Deep learning-based object detection algorithms, such as Convolutional Neural Networks (CNN) and YOLO architectures, have demonstrated promising outcomes in diverse areas like traffic monitoring, pedestrian detection, and facial recognition [6], [7]. Numerous studies have focused on detecting helmet usage [8]- [17]. The ongoing progress in deep learning-based computer vision technologies continues to enhance safety

[*]Corresponding author: *ademkorkmaz@bandirma.edu.tr*
Received: 16.05.2023, Accepted: 19.09.2023

monitoring and enforcement within the construction sector and beyond.

In the study by Hayat & Morgado-Dias [3], they employed a benchmark dataset of 5,000 helmet images, divided into 60%, 20%, and 20% portions for training, testing, and validation, respectively. The findings revealed that the YOLOv5x architecture was the top performer, achieving an impressive average accuracy (mAP) of 92.44%. This demonstrates the model's effectiveness in detecting safety helmets, even under challenging low-light conditions.

In a study by Yung et al. [18], the researchers evaluated the performance of three deep learning algorithms (YOLOv5, YOLOv6, and YOLOv7) in detecting safety helmets through a series of three tests. YOLOv6s and YOLOv7 models demonstrated superior performance in low light conditions compared to the YOLOv5s model. However, it was noted that only some models could differentiate between regular and safety helmets. Ultimately, YOLOv7 emerged as the best performer, achieving the highest mAP of 89.6% in detecting protective helmets.

In a study by Otgonbold et al. [19], the researchers developed a helmet detection model using a dataset of six classes: helmet, head, helmeted head, helmeted person, helmetless person, and face. The study employed several algorithms, including YOLOv3 (YOLOv3, YOLOv3-tiny, and YOLOv3-SPP), YOLOv4 (YOLOv4 and YOLOv4pacsp-x-mish), YOLOv5-P5 (YOLOv5s, YOLOv5m, and YOLOv5x), Faster Region-Based Convolutional, and YNOL. The results showed that Faster-RCNN (Region-based Convolutional Neural Network) achieved the lowest mAP of 36.89, while the YOLOR model attained the highest mAP value of 88.28.

In a study by Chen et al. [20], the researchers aimed to achieve real-time and efficient helmet-wearing detection by utilizing a developed YOLOv4 algorithm. The results revealed that the algorithm achieved an accuracy of 92.98%, a model size of 41.88 M, and a detection speed of 43.23 images per second. Compared to the original YOLOv4, there was an increase in accuracy by 0.52%, a reduction in model size by approximately 83%, and an 88% improvement in detection speed.

Table 1. Success rates of Yolo and DL applications for helmet classification

Authors	Dataset	Applied Models	Results (mAP50-%)
[14]	13,000 images	SSD, Faster R-CNN, YOLOv3, and Improved YOLOv3	77.2, 94.3, 82.3 and 93.1
[12]	3261 images	SSD-MobileNet	36.8
[21]	13620 images	AT-YOLO + DIOU	96.5
[22]	1365 images	YOLOv2	98,52
[17]	2580 images	SCM-YOLO	93.19
[23]	5000 images	YOLO	97.12
[24]	7008 images	YOLO	95
[25]	7581 images	YOLOv5	93
[26]	3000 images	Faster R-CNN, SSD, YOLO v3, YOLO v4 and YOLO v4-HelMask	70.62, 89.72, 90.54, 93.19 and 95.51

This recent literature review underscores the effectiveness of various object detection models across diverse datasets. Studies [14, 12, 21, 22, 17, 23, 24, 25, 26] employed a range of models from SSD and Faster R-CNN to multiple versions of YOLO, including standard and modified variants such as AT-YOLO, YOLOv2, SCM-YOLO, and YOLOv4-HelMask. Despite the varying dataset sizes (ranging from 1,365 to 13,620 images), the results in terms of mAP50%, a standard metric indicating the model's precision, were generally high. YOLO and its variations consistently performed well, with a notably high mAP50% of 98.52 achieved by YOLOv2 on a dataset of 1,365 images [22]. SSD-MobileNet yielded the lowest mAP50% of 36.8 [12], possibly due to various factors, including dataset size and task complexity. These findings underscore the potency of YOLO and its variants in object detection tasks and suggest that considerations of dataset size, data quality, model selection, and task complexity are critical to optimizing model performance.

Conventional manual monitoring methods can be labor-intensive, subject to human error, and often inadequate for covering extensive construction sites. To address these drawbacks, this study investigates the use of deep learning approaches to create an automatic helmet detection system. This system aims to detect workers wearing helmets effectively and accurately in real-time, using cameras at construction sites.

2. Material and Method

In this study, deep learning approaches and image processing methods are used to ensure human safety by detecting whether a person is wearing a helmet. The deep learning-based image processing method is the YOLOv8 (You Only Look Once) model, a state-of-the-art object detection algorithm known for its real-time processing capabilities and accuracy. The weights of the YOLOv8 model were pre-trained using millions of images from the ImageNet dataset, providing a solid foundation for transfer learning and fine-tuning the model for the specific helmet detection task.

2.1. Data Set

This study aims to identify the presence of helmets and verify if individuals are wearing them to enhance safety on construction sites. A dataset of 7036 images was used to accomplish this objective, featuring categories such as humans, human heads, and helmets [27]. These images were gathered from the Mendeley [27] websites and underwent preprocessing to ensure they were appropriate for use in the research. An example view of the dataset is given in Figure 1.

Figure 1. Sample images of the dataset.

Features of the dataset used include:

• A total of 7036 images, upscaled to 16867 images using various data augmentation techniques.

• Images accessible to everyone were selected for the data set.

• The target classes of images in the dataset were selected from different lighting conditions and environments to provide a more robust model.

• Various preprocessing techniques were applied to each image in the dataset, focusing on data augmentation, resulting in a threefold increase in dataset size.

• The dataset is divided into 14674 training images, 1349 validation images, and 844 test images for use in experimental studies.

Data set augmentation studies:

• Set to 3 printouts per training sample. In other words, three images were obtained from each image.

• Horizontal directional flip operation has been performed.

• Crop: 0% Minimum Zoom, 20% Maximum Zoom applied.

• Rotation: Applied from -10° to +10°.

• Grayscale: Applied to 10% of images.

• Hue: Applied between -20° and +20°.

• Saturation: -25% to +25% applied.

• Brightness: -20% to +20% applied.

• Exposure: -20% to +20% applied.

• Blur: Applied up to 1 pixel.

• Cutting: 6 boxes of 3% size each were created.

• As a result of magnification, 21108 images were targeted, while 16867 images were obtained. This decrease in the number of images seen is because some images obtained because of processing will not fully serve experimental studies.

2.2. YOLOv8 Model

YOLOv8 is a recently developed, highly effective model using the YOLO (You Only Look Once) architecture. It was developed by Ultralytics, known for its work on the YOLOv8, YOLOv3 and YOLOv5 models. Object detection, sample positioning, and image classification can be made in this model, as in the YOLOv7 and YOLOv6 models. The YOLOv8 model also uses the Pytorch library like YOLOv7-v6. It can run on both CPU and GPU units as working hardware.

YOLOv8 can achieve strong accuracy in COCO object classification. For example, the mid-model YOLOv8m can reach 50.2% MAP when measured in COCO. YOLOv8 scores significantly better than YOLOV5 when evaluated against Roboflow 100, a dataset that evaluates model performance in various areas specific to the desired tasks.

In addition, YOLOv8 includes developer convenience features. Unlike other models that can split into many different Python files in the execution of tasks, YOLOv8 does this with a CLI that makes model training more intuitive. The architecture of the YOLOv8 model is given in Figure 2 below. The date of this article study is the only article study of the YOLOv8 model. The architecture of the YOLUv8 model, shown in Figure 2, was visualized by GitHub website users [28].

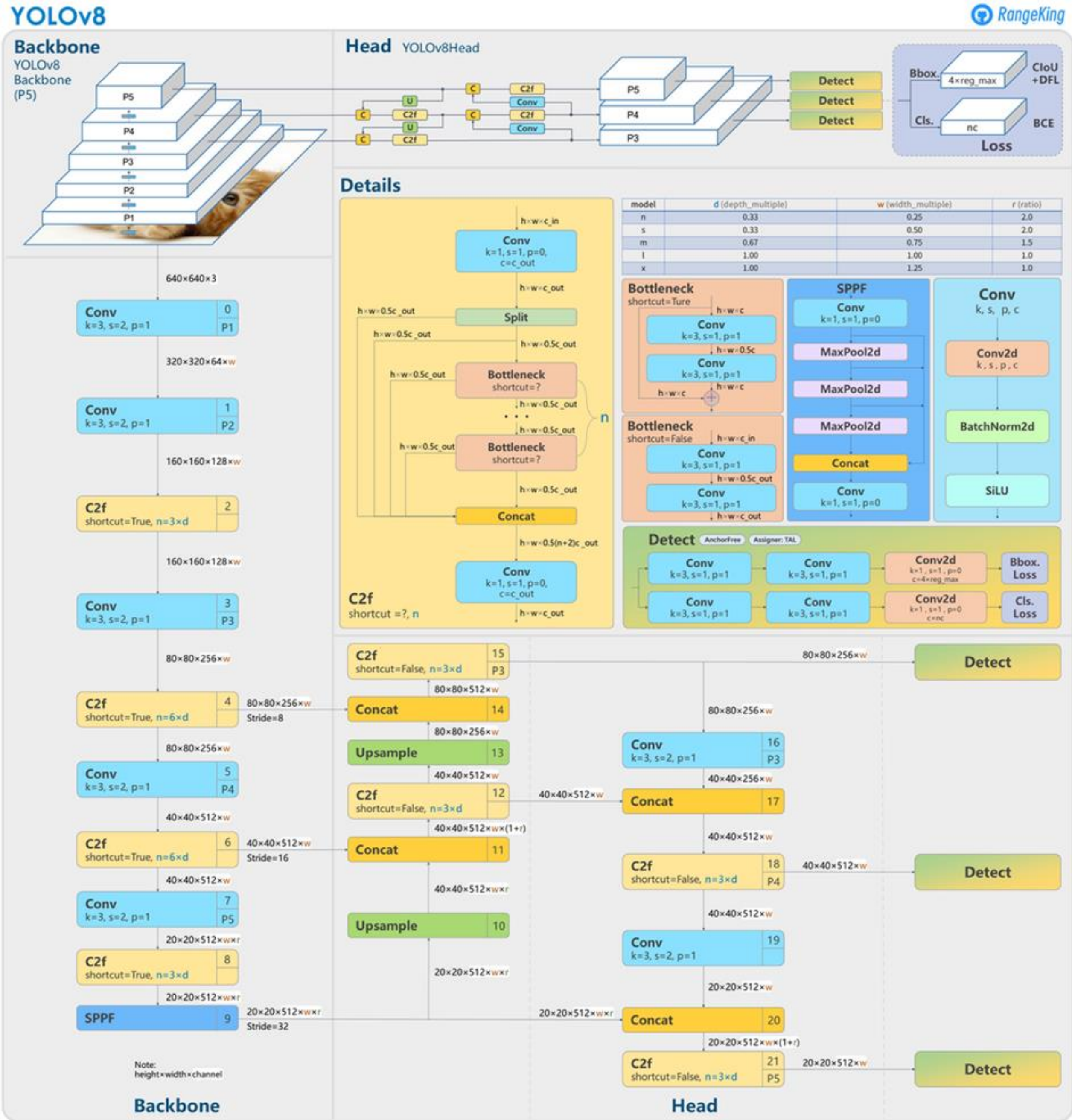

Figure 2. YOLOv8 Architecture [28]

3. Results and Discussion

Experimental studies were conducted with the study's deep learning-based YOLOv8 model. In experimental studies, it was aimed to determine three classes with image processing techniques. These classes are designated as helmet detection (helmet), head detection (head), and both head and helmet detection (all). In the working model, the batch value is eight, and the epoch value is 50 as a parameter. The numerical results of the experimental study are given in Table 2.

Table 2. Numerical results of the experimental study.

Class	Precision	Recall	mAP50	mAP50-95
All	0,938	0,933	0,969	0,642
Head	0,923	0,918	0,956	0,639
Helmet	0,952	0,947	0,971	0,646

Table 2 provides metrics evaluating the model's performance in terms of precision, recall, mAP50, and mAP50-95 for two categories: 'Head' and 'Helmet.'

Precision refers to the proportion of true positive predictions (correctly identified heads or helmets) among all positive predictions made by the model. An accuracy of 0.923 for 'Head' and 0.952 for 'Helmet' means the model is highly accurate when it predicts the presence of a head or helmet in an image.

Recall measures the proportion of actual positives (real heads or helmets in images) that the model correctly identified. A recall of 0.918 for 'Head' and 0.947 for 'Helmet' suggests the model is proficient at detecting most instances of heads or helmets when they are present.

The mAP50 (mean average precision at 50% Intersection over Union - IoU) is a commonly used metric for object detection tasks. It considers both precision and recall calculating an overall performance score. A score of 0.956 for 'Head' and 0.971 for 'Helmet' indicates excellent performance, with the model correctly identifying and accurately placing bounding boxes around heads and helmets in most cases.

The mAP50-95 is another version of the mAP score, but it averages scores over a range of IoU thresholds from 0.5 to 0.95. This stricter metric can provide a more comprehensive view of the model's performance. The scores of 0.639 for 'Head' and 0.646 for 'Helmet' are significantly lower than the mAP50 scores, suggesting the model's performance decreases at higher IoU thresholds.

The ’All’ category provides the average of the metrics across both the 'Head' and 'Helmet' categories. The overall mAP50 score of 0.969 indicates that the model performed very well across all classes in the dataset.

The confusion matrix outputs are given in Figure 3.

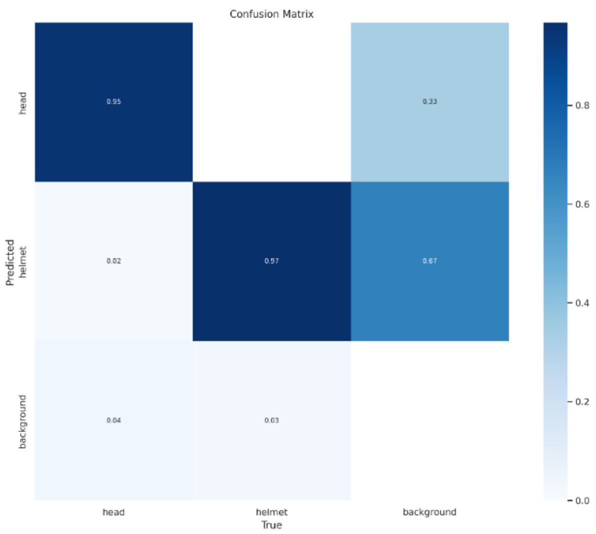

Figure 3. Confusion Matrix Output of the Study.

Predicted labels (head, helmet, background) are marked in rows, while actual titles are listed in columns. When the model predicted the "head" label, it was correct 95% of the time. He never mistakenly defined "helmet" as "head.” For the "helmet" predictions, the model was right 97% of the time, sometimes misclassifying "head" as "head" (2% of the time). Misclassified "head" as "background" 4% of the time and "helmet" as "background" 3% of the time. It is clear from these results that the model performs exceptionally well in the 'head' or 'helmet' prediction but needs help in accurately identifying the 'background,’ leading to a high rate of misclassification.

In Figure 4, the output values of the study are given graphically.

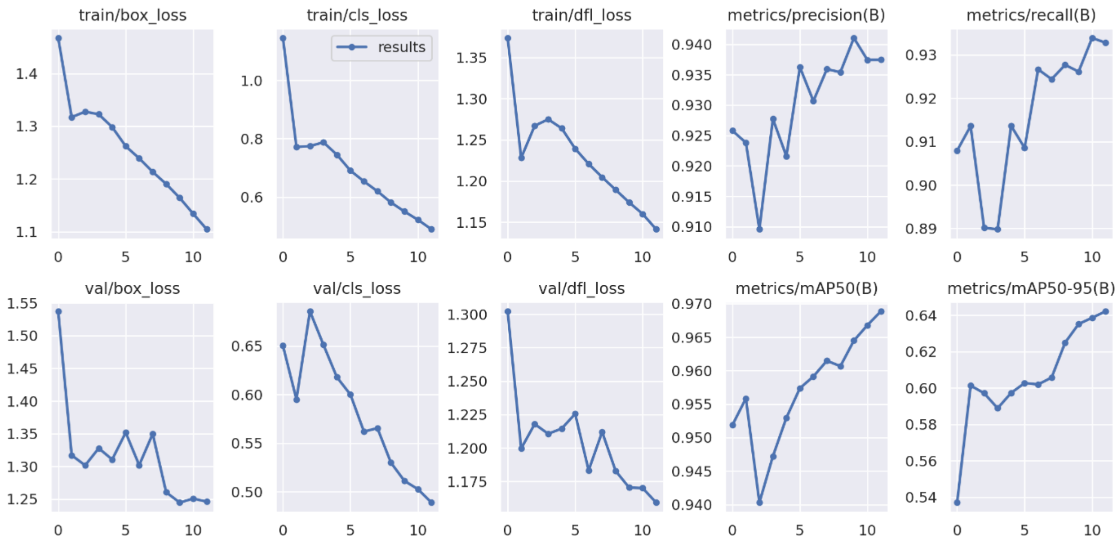

Figure 4. Graphical Representation of Experimental Study Analysis.

The test samples of the fixation system obtained from the experimental studies are given in Figure 5.

Figure 5. Example Helmet/Head Detection Ratio.

The data set used has been the subject of other scientific studies before. The table below gives this study's comparative results with similar data sets and similar studies.

Table 3. Comparison of experimental studies on automatic helmet detection with similar. Datasets.

Author	Model	mAP50 (%)
[14]	SSD, Faster R-CNN, YOLOv3, and Improved YOLOv3	77.2, 94.3, 82.3 and 93.1
[12]		36.8
[29]	SSD	96.0
[11]		68.5
[30]	YOLOv5s, YOLOv5m, YOLOv51 and YOLOv5x	93.6, 94.3, 94.4 and 94.7
[4]	YOLOv5 and Improved YOLOv5	92.1 and 95.7
[18]	YOLOv5s, YOLOv6s and YOLOv7	83.7, 83.5 and 89.6

Our Study	YOLOv8	96.9

Table 3 showcases a variety of studies, each utilizing different models for helmet detection and reporting the corresponding mAP50 scores achieved by each model. This comparison demonstrates the general improvement in the accuracy of helmet detection tasks with the evolution of the YOLO model from version 3 to version 8, keeping in mind that the different studies may have used different datasets and evaluation methods. The YOLOv8 model from this study achieved the highest mAP50 score of 96.9%, outperforming all other models tested in similar studies. This indicates that YOLOv8 is highly effective in helmet detection compared to other deep-learning models.

4. Conclusion and Suggestions

In this study, a system that performs automatic helmet control in areas where human life is in danger in common working and living environments such as construction and factories, especially in areas where there is a possibility of falling off an object harmful to the head, has been proposed. Studies have been done. In experimental studies, it has been possible to determine whether people automatically wear helmets on their heads with image processing techniques.

When Table 3 is examined, it is seen that the highest performance score among similar data sets and similar study samples is 96.9% with this study. The previous research with the highest success rate was Tan et al. [4]. It obtained a success value of 95.7%.

In addition, 98.1% for automatic helmet detection with computer vision and an average of 95.6% for human head detection (mAP50) were obtained in the study. In this and similar studies, it has been observed and suggested that YOLOv8, one of the deep learning-based models, gives more successful results than other models.

Potential directions for future work include enhancing the variety of the dataset by incorporating images from different industries where helmets are used, capturing various types of helmets, and considering diverse lighting and weather conditions. Real-time implementation of the helmet detection system in an actual construction site or other relevant industry could offer invaluable insights into its real-world effectiveness and the challenges that might arise in such a context. The scope of the study could also be extended to include the detection of other forms of Personal Protective Equipment (PPE), such as safety vests, gloves, and safety glasses, contributing to a more comprehensive safety monitoring system. Additionally, integrating the helmet detection system with an alarm or notification system could serve as an immediate alert mechanism for supervisors or safety officers when a worker is detected without a helmet. While the YOLOv8 model has demonstrated promising results, other emerging models should be explored for helmet detection, ensuring continuous evaluation of new models and techniques to remain at the forefront of technology. A deeper investigation into the causes of false positives and negatives in the current model could lead to enhanced accuracy, involving a detailed analysis of the cases where the model fails and implementing strategies to rectify these inaccuracies. Lastly, developing lightweight models suitable for real-time applications, which minimally compromise accuracy but significantly reduce computational requirements, could be a key direction for future research. This would enable on-site deployment on edge devices for instant alerts and actions.

Contributions of the authors

Corresponding author: writing (original draft, review & editing) and investigation.
Coauthor: Methodology and experimental study.

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The study is complied with research and publication ethics

References

[1] X. Huang and J. Hinze, “Analysis of construction worker fall accidents,” *J. Constr. Eng. Manag.*, vol. 129, no. 3, pp. 262–271, 2003.

[2] R. A. Haslam et al., “Contributing factors in construction accidents,” *Appl. Ergon.*, vol. 36, no. 4, pp. 401–415, 2005.

[3] A. Hayat and F. Morgado-Dias, “Deep learning-based automatic safety helmet detection system for construction safety,” *Appl. Sci. (Basel),* vol. 12, no. 16, p. 8268, 2022. https://doi.org/10.3390/app12168268

[4] S. Tan, G. Lu, Z. Jiang, and L. Huang, “Improved YOLOv5 network model and application in safety helmet detection,” *in 2021 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, 2021. https://doi.org/10.1109/ISR50024.2021.9419561

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature,* vol. 521, no. 7553, pp. 436–444, 2015. doi:10.1038/nature14539

[6] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv [cs.CV], 2018.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *in 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[8] P. Doungmala and K. Klubsuwan, “Helmet wearing detection in Thailand using Haar like feature and circle Hough transform on image processing,” *in 2016 IEEE International Conference on Computer and Information Technology (CIT),* 2016.

[9] W. Zhang, C.-F. Yang, F. Jiang, X.-Z. Gao, and X. Zhang, “Safety helmet wearing detection based on image processing and deep learning,” *in 2020 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 2020.

[10] S. H. Kim, C. Wang, S. D. Min, and S. H. Lee, “Safety helmet wearing management system for construction workers using three-axis accelerometer sensor,” *Appl. Sci. (Basel),* vol. 8, no. 12, p. 2400, 2018.

[11] X. Long, W. Cui, and Z. Zheng, “Safety helmet wearing detection based on deep learning,” *in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2019.

[12] Y. Li, H. Wei, Z. Han, J. Huang, and W. Wang, “Deep learning-based safety helmet detection in engineering management based on convolutional neural networks,” *Advances in Civil Engineering,* vol. 2020, pp. 1–10, 2020.

[13] H. Wang, Z. Hu, Y. Guo, Z. Yang, F. Zhou, and P. Xu, “A real-time safety helmet wearing detection approach based on CSYOLOv3,” *Appl. Sci. (Basel),* vol. 10, no. 19, p. 6732, 2020.

[14] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, “Detection algorithm of safety helmet wearing based on deep learning,” *Concurr. Comput*., vol. 33, no. 13, 2021.

[15] G. Han, M. Zhu, X. Zhao, and H. Gao, “Method based on the cross-layer attention mechanism and multiscale perception for safety helmet-wearing detection,” *Comput. Electr. Eng*., vol. 95, no. 107458, p. 107458, 2021.

[16] K. Han and X. Zeng, "Deep Learning-Based Workers Safety Helmet Wearing Detection on Construction Sites Using Multi-Scale Features," *in IEEE Access*, vol. 10, pp. 718-729, 2022, doi: 10.1109/ACCESS.2021.3138407.

[17] B. Zhang, C.-F. Sun, S.-Q. Fang, Y.-H. Zhao, and S. Su, “Workshop safety helmet wearing detection model based on SCM-YOLO,” *Sensors (Basel),* vol. 22, no. 17, p. 6702, 2022.

[18] N. D. T. Yung, W. K. Wong, F. H. Juwono, and Z. A. Sim, “Safety helmet detection using deep learning: Implementation and comparative study using YOLOv5, YOLOv6, and YOLOv7,” *in 2022 International Conference on Green Energy, Computing and Sustainable Technology (GECOST)*, 2022. doi:10.1109/GECOST55694.2022.10010490

[19] M.-E. Otgonbold et al., “SHEL5K: An extended dataset and benchmarking for Safety HELmet detection,” *Sensors (Basel),* vol. 22, no. 6, p. 2315, 2022. Doi:10.3390/s22062315

[20] J. Chen, S. Deng, P. Wang, X. Huang, and Y. Liu, “Lightweight helmet detection algorithm using an improved YOLOv4,” *Sensors (Basel),* vol. 23, no. 3, p. 1256, 2023. doi:10.3390/s23031256

[21] Q. Zhou, J. Qin, X. Xiang, Y. Tan, and N. N. Xiong, “Algorithm of helmet wearing detection based on AT-YOLO deep mode,” *Comput. Mater. Contin.*, vol. 69, no. 1, pp. 159–174, 2021.

[22] Y. Jamtsho, P. Riyamongkol, and R. Waranusast, “Real-time license plate detection for non-helmeted motorcyclist using YOLO,” *ICT Express,* vol. 7, no. 1, pp. 104–109, 2021.

[23] N. K. Anushkannan, V. R. Kumbhar, S. K. Maddila, C. S. Kolli, B. Vidhya, and R. G. Vidhya, “YOLO algorithm for helmet detection in industries for safety purpose,” *in 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, 2022.

[24] J. Li, Y. Li, J. F. Villaverde, X. Chen, and X. Zhang, "A safety wearing helmet detection method using deep leaning approach," *J. Opt.*, 2023.

[25] J. Fang, X. Lin, F. Zhou, Y. Tian, and M. Zhang, "Safety Helmet Detection Based on Optimized YOLOv5," *in 2023 Prognostics and Health Management Conference (PHM),* IEEE, 2023, pp. 117–121.

[26] Z. Zhang, Y. Tang, Y. Yang, and C. Yan, "Safety Helmet and Mask Detection at Construction Site Based on Deep Learning," *in 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA),* vol. 3, IEEE, 2023, pp. 990–995.

[27] M. Gochoo, "Safety helmet wearing dataset." Mendeley, 2021. doi: 10.17632/9rcv8mm682.1

[28] RangeKing, (2023), Brief Summary of YOLOv8 Model Structure, URL: https://github.com/ultralytics/ultralytics/issues/189 date of access: 01/05/2023.

[29] A. Kamboj and N. Powar, "Safety helmet detection in industrial environment using deep learning," *in 9th International Conference on Information Technology Convergence and Services (ITCSE 2020)*, 2020.

[30] F. Zhou, H. Zhao, and Z. Nie, "Safety helmet detection based on YOLOv5," *in 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*, 2021.