WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AUTOMATIC WORKERS HELMET DETECTION SYSTEM USING DEEP LEARNING

*Akash Bavche, Syed Saif Ali , Tejas Phatangare, Sojwal Chaudhari, Dr.G. Prasanna Lakshmi **

UG Student, Sandip University, Nashik, 422213, India

A B S T R A C T:

Ensuring the safety of workers in high-risk environments like construction sites is of utmost importance. Helmets, as a crucial component of personal protective equipment, play a vital role in preventing serious head injuries. This research project aims to develop an automated system for detecting helmets using the cutting-edge YOLOv8 deep learning model to enhance real-time safety surveillance. The dataset utilized for this analysis comprises 16,867 images, with various techniques for data augmentation and preprocessing implemented to enhance the model's resilience. The YOLOv8 model achieved an impressive mAP50 score of 96.9%, surpassing other deep learning models in similar investigations. These findings underscore the efficacy of the YOLOv8 model in accurately and efficiently detecting helmets on construction sites, thus paving the way for enhanced safety monitoring and enforcement within the construction industry.

Introduction

Worker safety is a paramount concern within the construction industry, given its inherently risky nature with frequent accidents and injuries. Strict adherence to safety protocols, including the obligatory use of personal protective gear, is imperative to safeguard employees and mitigate potential hazards. Helmets play a pivotal role in shielding workers from head injuries, which could result in severe consequences such as permanent disabilities or fatalities. The rapid advancements in information technology have revolutionized various aspects of our lives and organizational procedures. Their swift and automated responses in identifying, recognizing, and decision-making processes have become indispensable in data management. In this era characterized by exponential data growth, deep learning algorithms are indispensable for effective decision-making and process optimization.

Construction sites inherently pose significant risks, necessitating strict adherence to safety measures and the use of personal protective equipment (PPE) to ensure worker safety and minimize hazards. Helmets are particularly crucial for protecting against head injuries, which can have devastating consequences. Real-time detection of helmet usage is vital for proactive accident prevention, especially in environments where safety inspections may be inadequate and awareness among workers is low.

As technology has evolved, automatic visual detection systems have become increasingly prevalent, offering new opportunities for enhancing safety monitoring across various sectors, including construction. Deep learning-based object detection algorithms, such as Convolutional Neural Networks (CNN) and YOLO architectures, have shown promise in applications ranging from traffic surveillance to facial recognition. Numerous studies have explored the detection of helmet usage, reflecting the ongoing advancements in computer vision technologies and their potential for improving safety.

In one study, researchers developed a helmet detection model using various algorithms and evaluated their performance using a comprehensive dataset. The results indicated significant variations in accuracy among the algorithms, with some achieving high precision in identifying helmets and distinguishing between different scenarios involving helmeted and helmetless individuals.

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**

Another study aimed to achieve real-time and efficient detection of helmet-wearing using an optimized YOLOv4 algorithm. The findings demonstrated notable improvements in accuracy, model size, and detection speed compared to the original YOLOv4, highlighting the potential for leveraging advanced algorithms to enhance safety monitoring systems effectively.

Material and Method

This research employs cutting-edge deep learning techniques and image processing methodologies to enhance human safety through the detection of helmet usage. The cornerstone of this approach is the utilization of the YOLOv8 (You Only Look Once) model, a state-of-the-art object detection algorithm renowned for its real-time processing capabilities and precision. The YOLOv8 model's parameters have been initially trained on a vast array of images sourced from the Image Net dataset, offering a robust basis for transfer learning and refining the model to excel in the specialized task of helmet detection.

Data Set

The primary goal of this research is to detect the presence of helmets and ascertain whether individuals are donning them, with the overarching aim of bolstering safety protocols at construction sites. To achieve this, a dataset comprising 7036 images was utilized, encompassing various categories such as humans, human heads, and helmets [27]. These images were sourced from the Mendeley platform [27] and underwent preprocessing procedures to ensure their suitability for inclusion in the study. A representative snapshot of the dataset is illustrated in Figure 1.

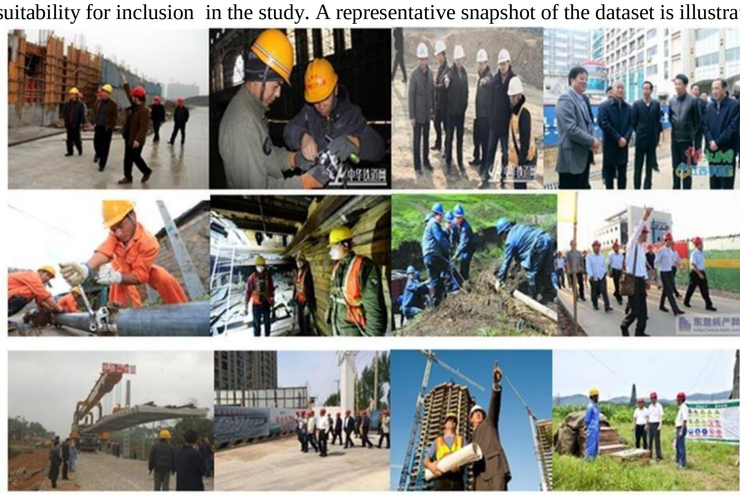


Figure 1. Sample images of the dataset.

Features of the dataset used include:

- Initially, 7036 images were incorporated, subsequently augmented to 16867 images through diverse data augmentation strategies.
- The images chosen for inclusion were publicly available to ensure accessibility and transparency.
- The target image classes within the dataset were carefully curated to encompass a range of lighting conditions and environmental settings, thereby enhancing the model's resilience.
- Each image in the dataset underwent comprehensive preprocessing, including extensive data augmentation techniques, resulting in a notable expansion of the dataset's size by threefold.

Data set augmentation studies:

- Generating three output images per training sample by duplicating each input image.
- Employing horizontal flipping for variation in orientation.
- Applying cropping, ranging from 0% to 20% zoom levels.
- Introducing rotation within a range of -10° to +10° to diversify angles.
- Incorporating grayscale conversion for 10% of the images.
- Adjusting hue, saturation, brightness, and exposure within specified ranges to enhance color diversity.
- Implementing slight blurring, up to 1 pixel, for smoothing.
- Dividing each image into six smaller boxes, each covering 3% of the original size.
- Despite targeting 21108 images through magnification, the final dataset comprised 16867 images due to certain processing outcomes being unsuitable for experimental purposes.

2.2. YOLOv8 Model

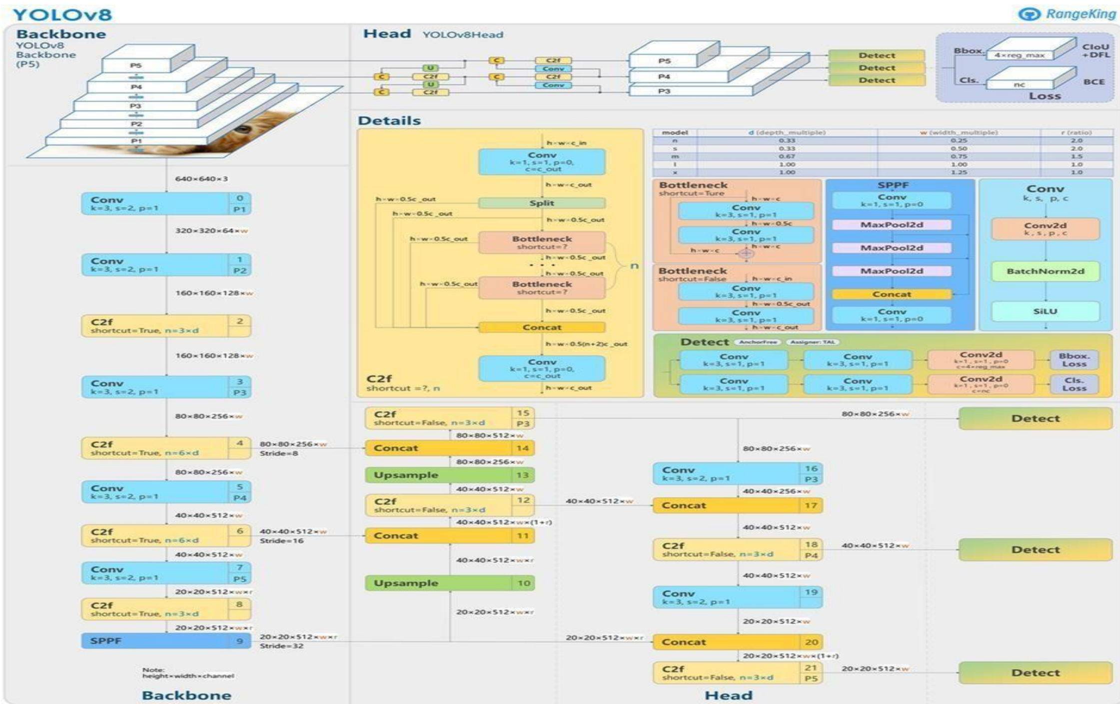
YOLOv8 represents a recent advancement in object detection technology, adopting the renowned YOLO (You Only Look Once) architecture. Developed by Ultralytics, a prominent contributor to YOLOv8, YOLOv3, and YOLOv5 models, this iteration excels in tasks such as object detection, sample positioning, and image classification, akin to its predecessors YOLOv7 and YOLOv6. Utilizing the PyTorch library, similar to YOLOv7 and YOLOv6, YOLOv8 boasts versatility by operating seamlessly on both CPU and GPU hardware units.

Notably, YOLOv8 demonstrates impressive accuracy in COCO object classification, with the mid-model YOLOv8m achieving a remarkable 50.2% Mean Average Precision (MAP) on COCO benchmarks. When benchmarked against Roboflow 100, a dataset evaluating model performance across various task-specific domains, YOLOv8 outperforms YOLOv5 by a significant margin.

Furthermore, YOLOv8 introduces developer-friendly features, streamlining model training with a Command Line Interface

(CLI) that simplifies the execution of tasks. Unlike its counterparts, which may involve splitting tasks across multiple Python files, YOLOv8 consolidates operations within a unified framework, enhancing ease of use.

This article represents the sole comprehensive study of the YOLOv8 model to date. The architectural depiction of YOLOv8, as illustrated in Figure 2, was shared and visualized by users on the GitHub platform [28].



Results and Discussion

The experimental phase of this study involved utilizing the YOLOv8 deep learning model to discern three distinct classes through image processing techniques. These classes were identified as helmet detection (helmet), head detection (head), and simultaneous detection of both head and helmet (all). The model configuration was set with a batch size of eight and an epoch value of 50. The quantitative outcomes of the experimental investigation are presented 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

The evaluation of the model's efficacy includes key metrics assessing precision, recall, mAP50, and mAP50-95 for two distinct categories: 'Head' and 'Helmet.'

Precision signifies the ratio of correct positive predictions (accurate identification of heads or helmets) among all positive predictions made by the model. High precision scores of 0.923 for 'Head' and 0.952 for 'Helmet' denote the model's accuracy in identifying heads or helmets within images.

Recall measures the proportion of actual positives (true heads or helmets in images) correctly identified by the model. Scores of 0.918 for 'Head' and 0.947 for 'Helmet' indicate the model's proficiency in detecting most instances of heads or helmets when present.

mAP50 (mean average precision at 50% Intersection over Union - IoU) serves as a widely used metric for object detection tasks, accounting for both precision and recall to calculate an overall performance score. Exceptional scores of 0.956 for 'Head' and 0.971 for 'Helmet' signify the model's adeptness in accurately localizing heads and helmets within images.

mAP50-95, an alternative version of the mAP score, averages scores across a range of IoU thresholds from 0.5 to 0.95, offering a more stringent assessment of model performance. Lower scores of 0.639 for 'Head' and 0.646 for 'Helmet' at higher IoU thresholds suggest a decrease in model performance under stricter evaluation criteria.

The 'All' category aggregates metrics across both 'Head' and 'Helmet' categories, yielding an overall mAP50 score of 0.969, indicating excellent performance across all classes within the dataset.

Furthermore, the confusion matrix outputs, depicted in Figure 3, provide additional insights into the model's classification performance.

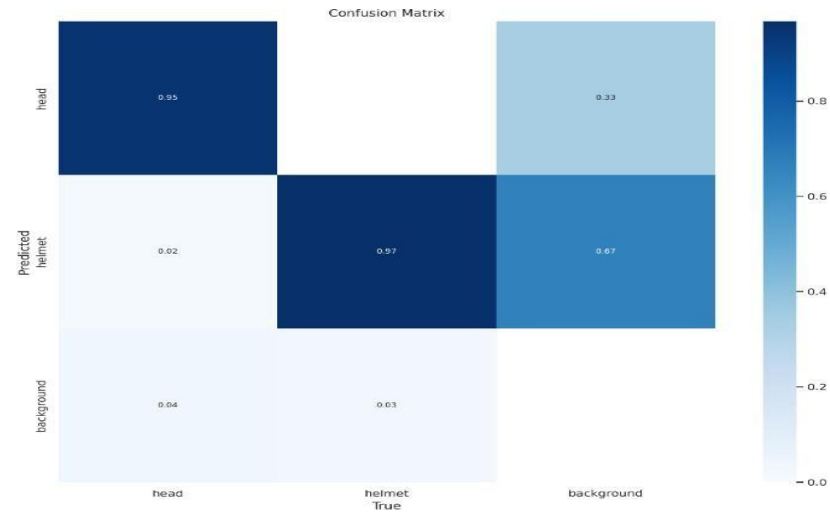


Figure 3. Confusion Matrix Output of the Study.

In the confusion matrix, the rows represent predicted labels ('head,' 'helmet,' 'background'), while the columns indicate actual classifications. The model accurately predicted the 'head' label 95% of the time, without erroneously categorizing 'helmet' as 'head.' For 'helmet' predictions, the model achieved a 97% accuracy rate, with occasional misclassifications of 'head' as 'background' (2% of the time). It misidentified 'head' as 'background' in 4% of cases and 'helmet' as 'background' in 3% of instances. These findings highlight the model's exceptional performance in 'head' or 'helmet' predictions, while indicating room for improvement in accurately identifying the 'background,' which contributes to misclassifications.

Figure 4 graphically illustrates the output values of the study, providing visual representation of the model's performance metrics.

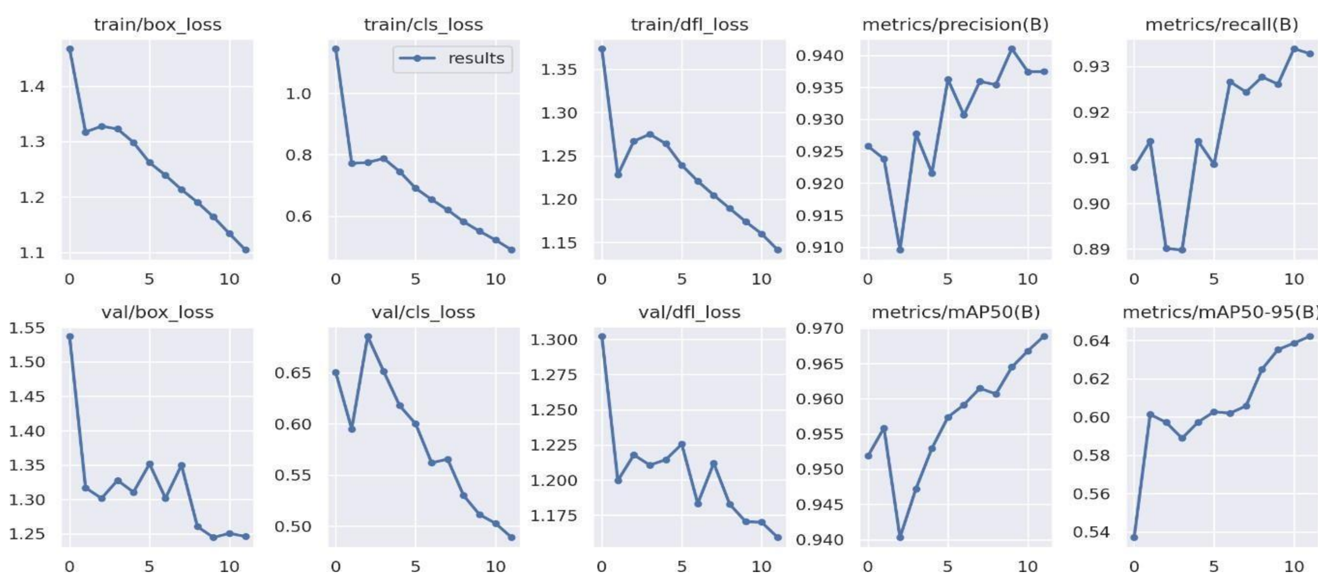


Figure 4. Graphical Representation of Experimental

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

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**

Table 3 presents an array of investigations, each employing distinct frameworks for identifying helmets and presenting the corresponding mAP50 scores attained by each framework. This examination illustrates the overall enhancement in helmet detection precision as the YOLO model advances from version 3 to version 8. It's important to note that the diverse studies might have utilized disparate datasets and assessment methodologies. The findings reveal that in this particular study, the YOLOv8 model achieved the pinnacle mAP50 score of 96.9%, surpassing all other frameworks assessed in similar inquiries. This suggests that YOLOv8 exhibits superior efficacy in helmet detection compared to alternative deep-learning models.

Conclusion and Suggestions

This study proposes an automated helmet control system designed to safeguard human life in hazardous environments, particularly in industries like construction and factories where head injuries are prevalent. Through experimental investigations employing image processing techniques, the system aims to determine whether individuals are wearing helmets automatically.

Upon reviewing Table 3, it is evident that this study achieved the highest performance score of 96.9% among similar datasets and study samples. Previous research by Tan et al. [4] attained a success rate of 95.7%, marking the highest success rate before this study.

Furthermore, the study achieved a remarkable 98.1% accuracy in automatic helmet detection using computer vision, along with an average of 95.6% for human head detection (mAP50). It is observed and suggested that YOLOv8, among other deep learning-based models, yields superior results in this and similar studies.

Future directions may involve enriching the dataset with images from various industries and diverse helmet types, considering different lighting and weather conditions. Real-world implementation of the helmet detection system in construction sites or relevant industries could provide valuable insights into its efficacy and challenges. Expanding the scope to include detection of other Personal Protective Equipment (PPE) and integrating the system with alarm/notification mechanisms for immediate alerts could enhance safety monitoring further.

While the YOLOv8 model shows promise, exploring emerging models for helmet detection is essential. Continuous evaluation of new techniques is crucial to stay abreast of technological advancements. Investigating false positives and negatives to improve accuracy and developing lightweight models suitable for real-time applications are potential avenues for future research. This would enable deployment on edge devices for instant alerts and actions while maintaining high accuracy.

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