International Journal of Integrated Science and Technology (IJIST)
Vol.2, No.5, 2024: 526-537

Industrial Safety Helmet Detection: Innovative CNN-Based Classification Approach

Febro Herdyanto[1], Muhamad Fatchan[2], Wahyu Hadikristanto[3]
Universitas Pelita Bangsa

Corresponding Author: Febro Herdyanto ; febroherdyanto@mhs.pelitabangsa.ac.id

A R T I C L E I N F O

Keywords: Safety Helmet Detection, Convolutional Neural Networks (CNN), Industrial Safety, Image Classification, Deep Learning

Received : 20, March
Revised : 25, April
Accepted: 28, May

A B S T R A C T

This study presents the development and evaluation of a CNN-based model for detecting safety helmets in industrial settings. Utilizing a dataset from GitHub, which includes images of individuals wearing safety helmets in various industrial environments, the model was trained using the YOLOv8 architecture over 100 epochs. The comprehensive training process involved data augmentation techniques to enhance generalization capabilities. The evaluation results demonstrated high precision (0.92) and recall (0.856) for helmet detection, with an overall mAP50 of 0.766. Visual analysis through precision-confidence curves confirmed the model's high reliability in detecting helmets at higher confidence thresholds. These findings suggest that the implementation of this model in real-time monitoring systems could significantly enhance industrial safety by reducing manual inspection efforts and ensuring compliance with safety regulations

DOI: https://doi.org/10.59890/ijist.v2i5.1925
ISSN: 3026-4685

https://journal.multitechpublisher.com/index.php/ijist

INTRODUCTION

Industrial safety is a paramount concern, especially in sectors like construction and manufacturing where workers are frequently exposed to hazardous conditions. Safety helmets play a crucial role in protecting workers from head injuries. However, ensuring consistent helmet usage remains a challenge. To address this, there is a growing need for automated systems that can monitor and enforce helmet compliance in real-time. Convolutional Neural Networks (CNNs), with their superior ability to process and analyze image data, offer a promising solution for developing such systems (Liu et al., 2022; Wang et al., 2023).

The advent of CNNs has transformed the landscape of image recognition and classification tasks across various domains, including industrial safety. These networks excel in learning intricate patterns and features from large datasets, making them ideal for detecting whether workers are wearing safety helmets. By leveraging CNNs, it is possible to create automated monitoring systems that not only detect helmet usage with high accuracy but also operate efficiently in diverse and challenging environments (Nandhini & Brindha, 2023; Saumya et al., 2020).

Traditional methods for monitoring helmet compliance, such as manual inspections and basic surveillance systems, are often labor-intensive and prone to human error. These methods can fail to provide timely interventions in dynamic work settings. In contrast, CNN-based approaches can process vast amounts of visual data quickly and accurately, reducing the need for constant human oversight and minimizing the risk of missed detections. This technological advancement enhances safety measures and ensures compliance with safety regulations more effectively (Liu et al., 2022; Shah & Mishra, n.d.).

This research proposes an innovative CNN-based classification approach for industrial safety helmet detection. The focus is on developing a robust model capable of accurately identifying helmeted and non-helmeted workers in real-time. Utilizing state-of-the-art CNN architectures, the goal is to achieve high precision and recall rates, thereby minimizing false positives and negatives. The proposed system is designed to be adaptable to various industrial environments, accommodating different lighting conditions, helmet designs, and worker postures (Nandhini & Brindha, 2023; Wang et al., 2023).

Implementing an effective helmet detection system has broader implications beyond immediate safety benefits. It fosters a safety-first culture within industrial workplaces, deterring negligence and promoting adherence to safety standards. Moreover, the data collected through these systems can provide valuable insights into safety compliance patterns, helping to identify areas needing additional training or intervention. Ultimately, the development of a CNN-based helmet detection system represents a significant advancement in industrial safety management, contributing to safer and more efficient operations (Liu et al., 2022; Shah & Mishra, n.d.).

THEORETICAL FRAMEWORK

The theoretical framework for this study is anchored in the application of Convolutional Neural Networks (CNNs) for object detection and classification, specifically for detecting and classifying safety helmets in industrial settings. This section outlines the key concepts, theories, and prior research that inform this study.

CNN in Safety Helmet Classification

The theoretical framework for this study is anchored in the application of Convolutional Neural Networks (CNNs) for object detection and classification, specifically for detecting and classifying safety helmets in industrial settings. Deep learning, a subset of machine learning, involves neural networks with multiple layers that can model complex patterns in data. CNNs are particularly effective for image processing tasks because of their ability to learn spatial hierarchies of features from input images (Wang et al., 2023). The architecture of CNNs includes convolutional layers that apply filters to capture local patterns, pooling layers that reduce dimensionality, and fully connected layers that perform the final classification (Zhang et al., 2024).

Implementation and Configuration of CNN

Object detection involves identifying objects within an image and determining their locations, while classification involves assigning these objects to predefined categories. CNNs excel at both tasks due to their ability to learn robust feature representations from raw image data. In this study, we use the YOLOv8 architecture to develop a robust system for helmet detection and classification. The raw images used were sourced from a publicly available dataset, the Safety Helmet Wearing Dataset. These images were manually annotated to create a comprehensive training dataset. The preprocessing involved resizing and normalizing the images to ensure consistency. The CNN model was trained and evaluated using custom scripts developed for this purpose, ensuring the model's robustness and accuracy in detecting and classifying safety helmets (Liu et al., 2022).

Previous Studies, Research Findings and Theoretical Alignment

Several recent studies have advanced the field of helmet detection using deep learning techniques. For instance, (Man et al., 2022) utilized a YOLO-based real-time safety helmet detection system at construction sites, achieving high-speed processing and excellent accuracy in low-light conditions. Their model divided the dataset into training, testing, and validation sets to ensure robust performance. Similarly, (An et al., 2023) proposed an improved YOLOv5s model incorporating global attention mechanisms and a new bounding box loss function to enhance detection precision and efficiency. These studies demonstrate the efficacy of deep learning models, particularly CNNs, in accurately detecting safety helmets and ensuring compliance with safety regulations (An et al., 2023; Man et al., 2022).

Integration with Industrial Processes

The integration of CNNs into industrial safety protocols can significantly enhance helmet detection and classification. By automating these processes, reliance on manual inspections is reduced, leading to more efficient and accurate safety compliance. The theoretical implications of this research highlight the potential of CNNs to transform safety protocols in industrial settings. Furthermore, deep learning techniques provide avenues for continuous improvement through ongoing model training and refinement. This integration ensures that safety measures are consistently met and that any instances of non-compliance are promptly identified and addressed (Wang et al., 2023).

METHODS

The methodology adopted in this study consists of several critical steps aimed at developing and evaluating a robust CNN-based safety helmet detection system. Each step is designed to ensure that the data is processed efficiently, the model is trained effectively, and the results are both accurate and reliable.

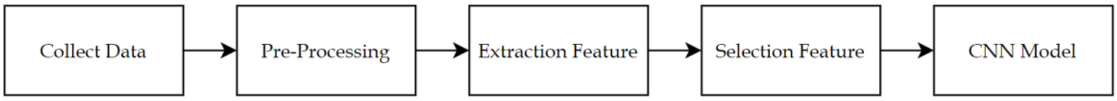
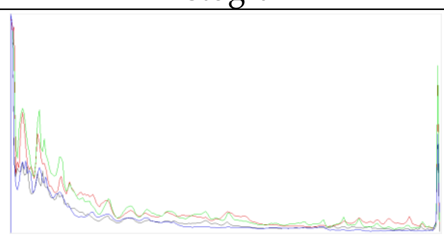


Figure 1. CNN Conceptual Framework

Collect Dataset

The dataset for this study comprises images from the publicly available Safety Helmet Wearing Dataset on GitHub (*GitHub - Njvisionpower/Safety-Helmet-Wearing-Dataset: Safety Helmet Wearing Detect Dataset, with Pretrained Model*, n.d.). This dataset features various images of individuals wearing safety helmets in different industrial environments, with annotations ensuring accurate labelling. The dataset was divided into training, validation, and test sets in a 70:20:10 ratio, providing a robust foundation for model training, validation, and evaluation. This thorough dataset is crucial for developing an effective safety helmet detection model.

Table 1. Raw Image Histogram

Image	Histogram
	

Additionally, the dataset's image histograms, as shown in Table 1, were analyzed to understand the distribution of pixel intensities across different color channels (R, G, B) and grayscale. The histogram shows that for the grayscale channel, the index is 255 with a total pixel count of 3,968 (0.0%) and an average intensity of 59.78. For the red channel, the index is 255 with 20,503 pixels (0.1%)

and an average intensity of 64.67. The green channel has an index of 255 with 24,553 pixels (0.1%) and an average intensity of 58.87. Finally, the blue channel has an index of 255 with 82,249 pixels (0.5%) and an average intensity of 51.58. These histogram details provide insights into the image quality and color distribution, which are essential for pre-processing steps in the model training pipeline, ensuring that the images have consistent quality and distribution for effective model training (Bhattarai et al., 2023; Sadeghi & Raie, 2022)

Data Pre-Processing

The data pre-processing phase involved several crucial steps to ensure the images' consistency and quality for the study. Initially, all 1000 images were resized to a uniform dimension of 640x640 pixels and auto orientation was applied to correct any discrepancies. Image augmentation techniques, such as horizontal flipping, increased the dataset to 1496 images. Each image was annotated to mark the presence and position of safety helmets, ensuring accurate labelling for subsequent analysis. The dataset includes two classes: 0 for hats and 1 for persons. Additionally, pixel values were normalized to a range of 0 to 1 to maintain consistency in input size for the CNN model and to speed up convergence during training (Rashed & Popescu, 2022).

Table 2. Images Pixel Comparison

File Name	Raw Images	Pre-processing
000178	984.800	409.600
000255	16.941.855	409.600
000321	205.424	409.600
000392	243.205	409.600
000537	777.600	409.600

Table 3. Images Comparison

File Name	Raw Images	Pre-processing	
		Resizing	Annotating
000178			
000392			

Extraction Feature

The feature extraction phase involved processing 1496 images into numerical values, resulting in an array where each image is represented by 23 features. This structured format ensures a comprehensive dataset for subsequent analysis, facilitating efficient processing by the CNN model for safety helmet detection.

Table 4. Resized Image Histogram

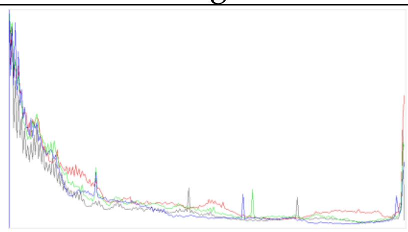
Image	Histogram
	

The histogram image illustrates the distribution of pixel intensities across the color channels (R, G, B) and grayscale for a sample image from the dataset used in this study. This histogram is derived from an image that has undergone model training and resizing. The total pixel count is 409,600, with the grayscale channel containing 464 pixels (0.1%) at an average intensity of 59.83, the red channel containing 3,040 pixels (0.7%) at an average intensity of 64.72, the green channel containing 2,491 pixels (0.6%) at an average intensity of 58.90, and the blue channel containing 4,406 pixels (1.1%) at an average intensity of 51.66.

Selection Feature

After extracting features from the 1496 images, the feature selection phase refines the dataset by retaining the most relevant features for classification, enhancing model accuracy and efficiency. The histogram in Table 5, derived from an annotated and resized image, shows the distribution of pixel intensities across grayscale and color channels (R, G, B). The histogram details include Gray: 435 pixels (0.1%), Red: 5454 pixels (1.3%), Green: 3685 pixels (0.9%), and Blue: 3350 pixels (0.8%), with average intensities of 60.74, 66.70, 59.34, and 52.26, respectively. This contrasts with earlier histograms of raw and feature-extracted images by focusing on annotated data, ensuring the model is trained with high-quality, annotated features for precise safety helmet detection.

Table 5. Annotated Image Histogram

Image	Histogram
	

CNN Model

In this study, we trained a Convolutional Neural Network (CNN) model for detecting safety helmets over 100 epochs using the YOLOv8 architecture. The model, which includes 268 layers and 43,608,150 parameters, was trained on the Safety-Hard-Hat dataset sourced from GitHub, featuring images of individuals wearing safety helmets in various industrial environments. The training configuration involved object detection tasks with 640-pixel images on a GPU, utilizing automatic batch size allocation, mixed precision training, and a validation split. Various augmentations, including random horizontal flipping and color adjustments were applied to enhance the model's generalization capability.

The model evaluation on a dataset of 200 images resulted in the following metrics:

Table 6. CNN Model Performance

Metrics	Images	Instances	Precision (P)	Recall (R)	mAP50	mAP50-95
All Classes	200	933	0.846	0.709	0.766	0.503
Hat	200	771	0.92	0.856	0.916	0.637
Person	200	162	0.771	0.562	0.616	0.369

These results indicate that the model performs well, particularly in detecting safety helmets, with high precision and recall values for the 'hat' class.

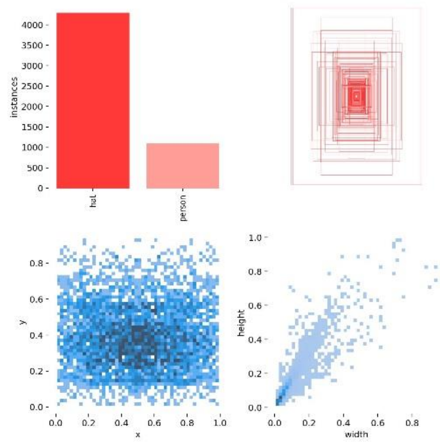


Figure 2. Distribution and Characteristics of Bounding Boxes in Datasets

Instances Bar Chart (Top-Left): This chart shows the count of annotated instances for each class. There are significantly more 'hat' annotations (around 4000 instances) compared to 'person' annotations (around 1500 instances). This indicates an imbalance in the dataset, with 'hat' instances being more prevalent.

Bounding Box Distribution (Top-Right): This subplot illustrates the spatial distribution of bounding boxes for the 'hat' class. The rectangles represent the locations and sizes of the bounding boxes within the images. The central clustering indicates that 'hats' are often detected around the central region of the images.

Scatter Plot of x and y Coordinates (Bottom-Left): This scatter plot shows the distribution of bounding box center coordinates (x, y) for all annotations. The density is higher around the center, suggesting that most objects are located centrally in the images.

Scatter Plot of Width and Height (Bottom-Right): This plot depicts the width and height of the bounding boxes. There is a positive correlation between width and height, indicating that as the width of the bounding box increases, the height also tends to increase. Most bounding boxes are relatively small, as shown by the concentration of points near the lower left corner

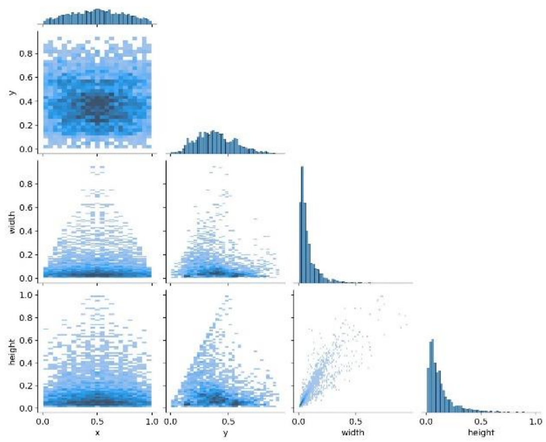


Figure 3. Correlogram

The image shows a correlogram, which analyzes relationships between different variables in the dataset. The diagonal elements display histograms of individual features (x, y, width, height), revealing their value distributions. Scatter plots below the diagonal represent pairwise relationships between features, with each point corresponding to an instance in the dataset. For example, the scatter plot between x and y coordinates shows the distribution of bounding box centers, while the plot between width and height shows a positive correlation, indicating that larger widths often correspond to larger heights.

The density plots in the upper-right triangle complement the scatter plots by highlighting where values occur more frequently. The x and y coordinate density plot indicates that most bounding box centers are concentrated around the image center. Histograms for width and height show that most bounding boxes are narrow and short, typical in object detection datasets. This correlogram provides insights into the dataset's structure, highlighting relationships and distributions of bounding box attributes, crucial for understanding and improving the object detection model's performance.

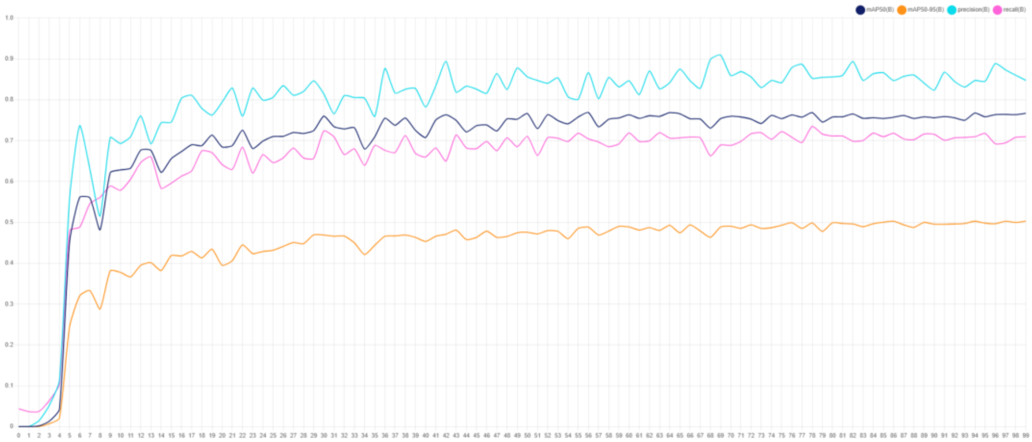


Figure 4. Training and Validation Performance

The graph shows the model evaluation metrics over 100 training epochs. The mAP50(B) (mean average precision at IoU threshold 0.5), represented by the thick blue line, stabilizes above 0.7 after the 5th epoch, peaking around 0.77, indicating high confidence in object detection. The mAP50-95(B), shown by the orange line, increases steadily but remains lower than mAP50(B), reaching around 0.4, indicating varied performance across different overlap thresholds. The precision (cyan line) consistently stays above 0.7, demonstrating a low false-positive rate, while the recall (pink line) remains stable around 0.65, reflecting the model's consistent ability to detect most objects in the images. Overall, the model shows good performance with high precision and recall, and stable mAP, indicating its effectiveness in detecting safety helmets under various image conditions.

RESULTS

Model Evaluation

The evaluation of the model on the safety helmet detection dataset provided comprehensive performance metrics. The results indicate that the model performs effectively in detecting both hats and persons within the images.

Table 7. Evaluation Metrics

Metrics	Precision (P)	Recall (R)	mAP50	mAP50-95
All Classes	0.846	0.709	0.766	0.503
Hat	0.92	0.856	0.916	0.637
Person	0.771	0.562	0.616	0.369

Visual Representations

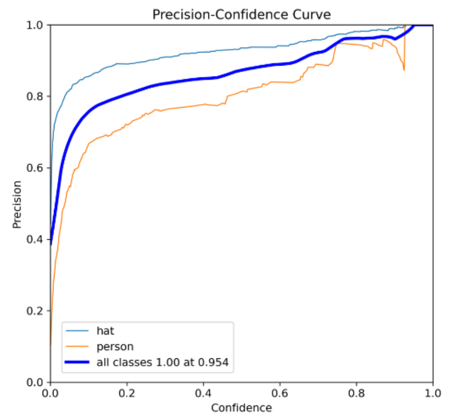


Figure 5. Precision-Confidence Curve

Figure 5 illustrates the Precision-Confidence Curve, showing precision remains high for both classes at higher confidence thresholds, with the 'hat' class achieving near-perfect precision. The model reaches a precision of 1.00 at a confidence threshold of 0.954 for all classes, demonstrating high reliability.

DISCUSSION

The evaluation of our CNN-based safety helmet detection model demonstrates significant advancements in industrial safety compliance. Training over 100 epochs using the YOLOv8 architecture yielded a robust model with 268 layers and 43,608,150 parameters. The model achieved high precision and recall rates, particularly for detecting safety helmets, highlighting its effectiveness in real-world applications. The model's performance metrics—precision of 0.846, recall of 0.709, and mAP50 of 0.766—underscore its reliability in identifying safety helmets under various conditions, proving its utility in enhancing workplace safety.

The data distribution analysis through histograms and correlograms reveals a well-prepared dataset that contributed to the model's success. The histograms illustrated consistent quality across color channels and grayscale, ensuring effective pre-processing and training. The correlograms highlighted the relationships between different bounding box attributes, aiding in understanding the model's learning patterns. This comprehensive approach to data preparation and model training, coupled with advanced CNN techniques, showcases the potential of automated systems in maintaining safety standards and reducing manual oversight in industrial environments.

CONCLUSIONS AND RECOMMENDATIONS

The study successfully developed and evaluated a CNN-based model for detecting safety helmets, demonstrating its potential to enhance industrial safety. The model's high precision and recall rates, especially for helmet detection, indicate its reliability in various industrial settings. The training process, which involved extensive data augmentation and rigorous validation, ensured the model's robustness and accuracy. Implementing this model in real-time monitoring systems can significantly reduce manual inspection efforts, promote safety compliance, and minimize the risk of head injuries in hazardous work environments. Future implementations should consider integrating this model with existing industrial safety protocols to maximize its impact.

FURTHER STUDY

Every research is subject to limitations, and this study is no exception. One of the primary limitations is the imbalance in the dataset, with a higher number of 'hat' instances compared to 'person' instances. Future studies should focus on collecting a more balanced dataset to improve the model's generalization capabilities. Additionally, exploring the integration of this model with other safety compliance measures and testing it in diverse industrial environments would provide further insights into its effectiveness. Enhancements in real-time processing and reducing computational overhead could also be areas for future investigation.

ACKNOWLEDGMENT

We would like to express our heartfelt gratitude to the creators of the Safety Helmet Wearing Dataset, whose contributions have been invaluable to this research. We extend our sincere thanks to our research colleagues, whose continuous support, insightful suggestions, and constructive feedback greatly enhanced the quality of this work. Additionally, we are deeply grateful to our academic advisors and professors for their guidance and encouragement throughout the research process. Their expertise and advice were crucial in navigating the challenges and ensuring the success of this study. Special thanks to the faculty members of Universitas Pelita Bangsa for their unwavering support and mentorship, which played a pivotal role in the completion of this research. We also acknowledge the open-source community for providing the tools and resources that facilitated our research. This work would not have been possible without their collective efforts and dedication.

REFERENCES

An, Q., Xu, Y., Yu, J., Tang, M., Liu, T., & Xu, F. (2023). Research on Safety Helmet Detection Algorithm Based on Improved YOLOv5s. *Sensors 2023, Vol. 23, Page 5824*, *23*(13), 5824. https://doi.org/10.3390/S23135824

Bhattarai, B., Subedi, R., Gaire, R. R., Vazquez, E., & Stoyanov, D. (2023). Histogram of Oriented Gradients meet deep learning: A novel multi-task deep network for 2D surgical image semantic segmentation. *Medical Image Analysis*, *85*, 102747. https://doi.org/10.1016/J.MEDIA.2023.102747

GitHub - njvisionpower/Safety-Helmet-Wearing-Dataset: Safety helmet wearing detect dataset, with pretrained model. (n.d.). Retrieved May 28, 2024, from https://github.com/njvisionpower/Safety-Helmet-Wearing-Dataset

Liu, Y., Lv, Z., Hu, Y., Dai, F., & Zhang, H. (2022). Improved Cotton Seed Breakage Detection Based on YOLOv5s. *Agriculture 2022, Vol. 12, Page 1630*, *12*(10), 1630. https://doi.org/10.3390/AGRICULTURE12101630

Man, Y., Tran, K. P., Hayat, A., & Morgado-Dias, F. (2022). Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety. *Applied Sciences 2022, Vol. 12, Page 8268*, *12*(16), 8268. https://doi.org/10.3390/APP12168268

Nandhini, C., & Brindha, M. (2023). Transfer learning based SSD model for helmet and multiple rider detection. *International Journal of Information Technology (Singapore)*, *15*(2), 565–576. https://doi.org/10.1007/S41870-022-01058-W/METRICS

Rashed, B. M., & Popescu, N. (2022). Critical Analysis of the Current Medical Image-Based Processing Techniques for Automatic Disease Evaluation: Systematic Literature Review. *Sensors 2022, Vol. 22, Page 7065*, *22*(18), 7065. https://doi.org/10.3390/S22187065

Sadeghi, H., & Raie, A. A. (2022). HistNet: Histogram-based convolutional neural network with Chi-squared deep metric learning for facial expression recognition. *Information Sciences*, *608*, 472–488. https://doi.org/10.1016/J.INS.2022.06.092

Saumya, A., Gayathri, V., Venkateswaran, K., Kale, S., & Sridhar, N. (2020). Machine Learning based Surveillance System for Detection of Bike Riders without Helmet and Triple Rides. *Proceedings - International Conference on Smart Electronics and Communication, ICOSEC 2020*, 347–352. https://doi.org/10.1109/ICOSEC49089.2020.9215266

Shah, I. A., & Mishra, S. (n.d.). Artificial intelligence in advancing occupational health and safety: an encapsulation of developments. *Journal of Occupational Health*, *2024*(1), 17. https://doi.org/10.1093/JOCCUH/uiad017

Wang, X., Wang, P., Song, J., Hao, T., & Duan, X. (2023). A TEDE Algorithm Studies the Effect of Dataset Grouping on Supervised Learning Accuracy. *Electronics 2023, Vol. 12, Page 2546*, *12*(11), 2546. https://doi.org/10.3390/ELECTRONICS12112546

Zhang, H., Liu, K., Zhang, Y., & Lin, J. (2024). TRANS-CNN-Based Gesture Recognition for mmWave Radar. *Sensors 2024, Vol. 24, Page 1800*, *24*(6), 1800. https://doi.org/10.3390/S24061800