Usage of Safety Helmet Warning System Using Deep Learning Method

Nur Masyita Ayu Ambarwati [1,a] **Dharma Aryani** [2,b] **Asriyadi** [3,c]
[1,2,3] Departement of Electrical Engineering, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia
[a] masyitaasy26@gmail.com (Corresponding Author) [b] dharma.aryani@poliupg.ac.id,
[c] asriyadi@poliupg.ac.id

Abstract*—K3 Helmet Usage Warning System Using Deep Method Learning Algorithm YOLOv8 is a technological innovation that combines the advanced object detection method YOLOv8 (You Only Look Once version 8) with the needs of safety in the work environment, especially the use of K3 (Occupational Health and Safety) helmets. This research aims to improve compliance with the use of K3 helmets through an automation approach using Deep Learning technology.The YOLOv8 algorithm is used to detect whether individuals in the work area are wearing K3 helmets or not. The detection results will be processed by the system to provide automatic warnings if individuals do not wearing an OHS helmet according to safety regulations. The use of Deep Learning technology in this method enables fast and accurate detection, contributing to increased awareness and compliance with workplace safety policies.With proper implementation, the system is expected to help create a safer work environment and improve the safety of workers, making the workplace more compliant with established OHS standards.

Keywords—Safety Helmet; Deep Learning; YOLOv8; Object Detection; Real Time.

I. Introduction

Ensuring worker safety in construction is a complex undertaking. Factors besides human error and technical shortcomings often contribute to a rise in work-related injuries and illnesses, impacting projects significantly[1]. A critical but overlooked aspect in improving overall well-being and worker productivity is the current state of knowledge, awareness, and responsible behavior among the workforce regarding Occupational Helath and Safety (OHS) practices[2].

Disturbing statistics reveal a high number of construction accidents every year, causing fatalities, property damage, and production halts. Data from Jamsostek in 2021 suggests a concerning number of cases 7,298 resulting in a staggering 9,224 worker casualties[3]. This alarming trend underscores the urgent need for a robust, integrated occupational health and safety (OHS) management system in construction projects.

Strict adherence to regulations encompassing engineering practices, safety protocols, worker health, environmental protection, and energy usage is crucial for ensuring orderly construction activities. Personal Protective Equipment (PPE) plays a vital role in mitigating work-related risks[4].

The growing emphasis on personal protective equipment detection underlines its critical importance for construction worker safety and productivity. Advancements in persnoal protective equipment detection technology offer a promising solution to this persistent problem[5].

Advancements in AI have led researchers to explore deep learning algorithms for detecting people without helmets on construction site[6]. These automated systems use computer vision and machine learning to create a reliable framework for identifying helmet usage[7].

A 2021 study by Widodo et al. investigated this concept. They developed a helmet detection system using deep learning techniques on a dataset of construction site images. Their system, employing Convolutional Neural Networks (CNNs) and transfer learning, achieved an impressive 93.33% accuracy in real-time helmet detection, processing 45 frames per second.This research showcases the potential of deep learning for revolutionizing construction safety. By automating helmet compliance checks, these systems can significantly improve safety by identifying and addressing violations[8].

This Final Year Project focuses on developing a helmet detection tool for construction workers using the YOLOv8 framework. Deep learning models serve as the backbone of the machine learning process, enabling the tool to recognize human head conditions and effectively detect workers wearing and not wearing helmets within the project area.

II. Research Methodology

A. Occupational Health and Safety

Ensuring Occupational Health and Safety (OHS) in Indonesia necessitates a collaborative effort from all parties involved, including government agencies, businesses, and the workforce itself[9].Indonesia's construction industry is grappling with a severe occupational safety problem. Data from Satudata Kemnaker reveals a staggering 370,747 work accidents in 2023, surpassing the previous year's figures. The informal sector is particularly concerning, with a high number of accidents. Even more alarming, 18 workers tragically lose their lives daily due to work-related mishaps[3].

This crisis sytems from a confluence of factors: a lack of awareness and ingrained culture of occupational health and safety (OHS), inadequate supervision, unsafe work environments with faulty equipment, and limited OSH training programs.

Helmet usage enforcement in construction zones is crucial for ensuring adherence to safety regulations and preventing accidents[10].Thankfully, advancements in technology offer solutions. Various technologies, including sensors, machine learning, and deep learning, have been developed specifically for helmet detection[11].

The latest deep learning-based object detection methods are gaining traction in the field of helmet detection research. Wang et al[5]. employed different YOLO architectures to successfully detect not only helmets, but also people and vests, in various colors. Their research showed that YOLOv5x offered the highest accuracy, while YOLOv5s provided the fastest processing speed. Similarly, Geng et al[12]. Utilized YOLO-based architectures for helmet detection. Geng et al.'s approach involved using the YOLOv3 architecture to tackle the challenge of imbalanced datasets in helmet detection. They employed a Gaussian blurring method to improve the accuracy of YOLOv3 for this specific task. Wu et al[13]. Proposed a single-shot CNN model for automatic object detection, including helmets.

Future research efforts can focus on enhancing helmet detection accuracy under challenging conditions, such as extremely low light or situations where objects are obscured[14]Additionally, researchers can explore developing systems that not only detect helmets but also identify different helmet types and provide additional worker information, such as their identity and location on the construction site.

B. Deep Learning

Deep learning constitutes a subfield within artificial intelligence (AI) and machine learning (ML). It leverages the development of multi-layered neural networks to attain superior accuracy in a variety of tasks, including object detection, speech recognition, and language translation[15].

Inspired by the structure and function of the human brain, deep learning, often abbreviated as DNNs (deep neural networks), represents a distinct area of study within artificial intelligence and machine learning. In contrast to traditional machine learning models that employ only one or two neural network layers, deep learning utilizes intricate, multi-layered architectures. This enables deep learning to progressively extract patterns and representations from data in a hierarchical fashion.

Convolutional Neural Networks (CNNs) and YOLO (You Only Look Once) represent advancements in technology driven by the evolution of deep learning from multi-layered neural networks[16].

Convolutional Neural Networks (CNNs) represent a potent form of deep learning architecture, meticulously tailored for image analysis. They replicate the hierarchical information processing of the human brain through intricately interconnected layers of neurons.

Convolusional Neural Networks (CNNs) represent an advanced form of deep learning architecture specifically tailored for processing and analyzing visual data[17]. By utilizing complex layers of interconnected neurons, CNNs are engineered to identify patterns and features in images, mirroring the hierarchical information processing of the human brain. This hierarchical approach empowers CNNs to recognize intricate visual information, making them exceptionally adept at tasks like image recognition, object detection, and image classification.

Convolutional neural networks (CNNs) have been trained on large sets of labeled image data, enabling them to continually learn and improve. Their adeptness at extracting meaningful information from images has driven advancements in fields such as computer vision, autonomous vehicles, medical imaging, and beyond. As research progresses, CNNs are poised to play an increasingly influential role in shaping the future of technology[18].

YOLO is a pioneering approach to object detection, differentiating itself from traditional methods with its single-shot approach that enables real-time processing. It boasts a unified architecture integrating object detection and class prediction for enhanced efficiency[19].

To maximize speed, YOLO divides the image into a grid, with each cell responsible for detecting objects within its designated area[20]. It utilizes convolutional neural networks (CNNs) to extract features, predict bounding boxes, and assign class probabilities to detected objects. Trained on a vast dataset of labeled images, YOLO excels in real-time applications such as robotics and autonomous vehicles. Its speed, accuracy, unified framework, and adaptability make it an invaluable tool for various object detection tasks[21].As deep learning advances, YOLO is poised to play an even more significant role in shaping the future of computer vision[22].

In summary, this research compares the two methods to evaluate the better neural network method for use in the K3 Helmet detection system

C. System Design Model

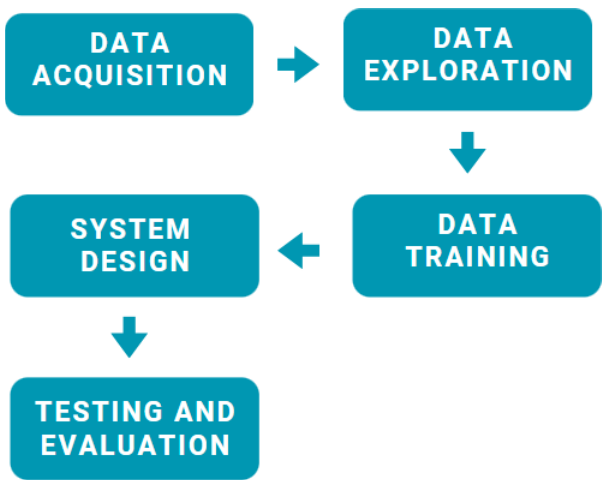


Figure 1. System Design Model

This paper proposes the development of a warning system that will contribute to the orderly implementation of Occupational Health and Safety (OHS) protocols in construction projects. The system will specifically focus on helmet detection for workers within construction areas. A general overview of the workflow for creating this OHS helmet detection warning system is presented in Figure 1.

1) Data Acquisition

In this phase, our research will direct its attention towards acquiring an appropriate dataset for the purposes of training and evaluation. The desired data format encompasses annotated images featuring safety helmets and human heads. Public online repositories will serve as the principal resource for exploring datasets. Our initial investigative platform will be Kaggle, focusing specifically on datasets comprising images of construction workers donning safety helmets.

2) Data Exploration

The dataset comprises a total of 2,000 images segregated into two distinct classes: "head" and "helmet." It will undergo partitioning into two subsets, with 70% earmarked for training and 30% for testing purposes.

3) Data Training

It is important to note that the training and testing of datasets are fundamental processes in the development and evaluation of Machine Learning models. This stage involves the utilization of the dataset to train the model, which allows it to acquire the capability to make precise predictions.

4) System Design

In the context of developing a Machine Learning-based Project Helmet Warning Detection Device, the usage of ESP32CAM and a Buzzer is an integral part of the developmental process, as illustrated in Figure 2.

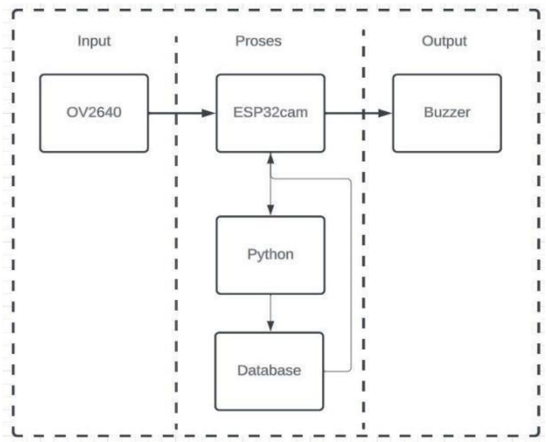


Figure 2. Block Diagram

5) Testing and Evaluation

During this phase, a thorough review of the process is conducted, and an evaluation is performed based on the outcomes obtained from a system or model. Subsequently, the results are analyzed to confirm that the system or model performs as expected, and to understand how these findings can be leveraged to resolve challenges or make well-informed decisions.

III. Results and Discussion

In this research, the results are categorized into five parts: LSTM , sentiment analysis, network analysis, website design, and testing. Each part is detailed and explained below.

A. Hardware Appearance

Figure 3. Hardware Appearance

Figure 3 provides a visual representation of the K3 Helmet Usage Warning System's design. This system is composed of several key components, including ESP32Cam, FT232RL, Buzzer, BoardCH340, and interconnected cables via USB.

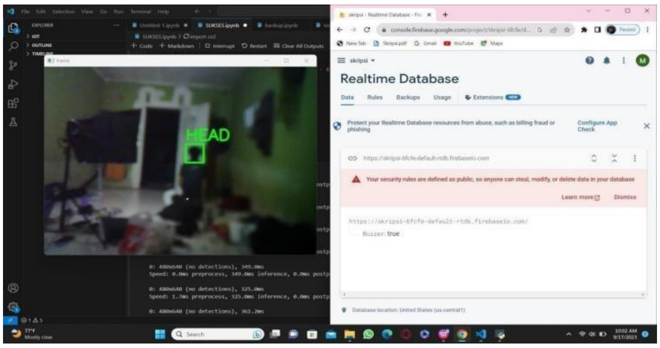


Figure 4. Detection Display

In Figure 4, the diagram illustrates the system's process for object detection and image pre-processing. The subsequent step involves object recognition, followed by the classification of the object as either a head or a helmet. Upon detecting a head, the system transmits the data to Firebase with a value of 'true,' which triggers the buzzer to sound. Conversely, if a helmet is detected, the system sends 'false' to Firebase, and the buzzer remains silent.

B. Modelling

This stage of the neural network model creation process produces two models: CNN and YoloV8.

1) CNN

The model is designed using a sequential model, enabling the systematic addition of layers in a predefined sequence. Initially, a convolutional layer with 64 filters, a 3x3 kernel size, and an input image size of 128x128 pixels is integrated. Following this, a max-pooling layer is introduced to reduce the image dimension. Subsequently, a batch normalization layer is added, and a rectified linear unit (ReLU) activation function is applied to expedite training and enhance model stability. [25]

This process involves the step-by-step addition of convolutional layers, max-pooling layers, batch normalization layers, and ReLU layers, which constitutes a common pattern in CNNs for progressively extracting features. After feature extraction, a flattened layer is included to convert the 2D feature matrix into a one-dimensional vector, enabling connection to the dense layer.

The dense layer consists of 64 neurons, determined through empirical testing. Subsequently, the ReLU activation function is applied once more, along with a dropout layer to mitigate overfitting by randomly eliminating 50% of connections during training.[26]

2) YOLOV8

The YOLOv8 model, which consists of 168 layers and 11,126,745 parameters, underwent evaluation using a validation dataset comprising 180 images free of background interference and data corruption. The evaluation encompassed 12 object classes and employed Precision (P), Recall (R), and Average Precision (AP) metrics at Intersection over Union (IoU) 0.50 (mAP50) and Average Precision across the IoU range of 0.50 to 0.95 (mAP50-95).[27]

The model achieved a mAP50 score of 0.638 and a mAP50-95 score of 0.403, indicative of robust object detection capabilities. Noteworthy are the "helmet" class (mAP50 of 0.946) and the "head" class (mAP50 of 0.917) which exhibited exemplary performance.

However, the model encountered challenges detecting "person," resulting in a notably low mAP50 score of 0.0513. Regarding speed, the model's inference time is 425.9 ms per image, accompanied by additional time allowances for pre-processing (3.8 ms), loss calculation (0.0 ms), and post-processing (2.0 ms) per image.

C. *YoloV8 Testing*

The YOLO method was applied to conduct testing aimed at the detection of heads and helmets. This testing involved 50 iterations or epochs to obtain accurate recall values for the detection process.

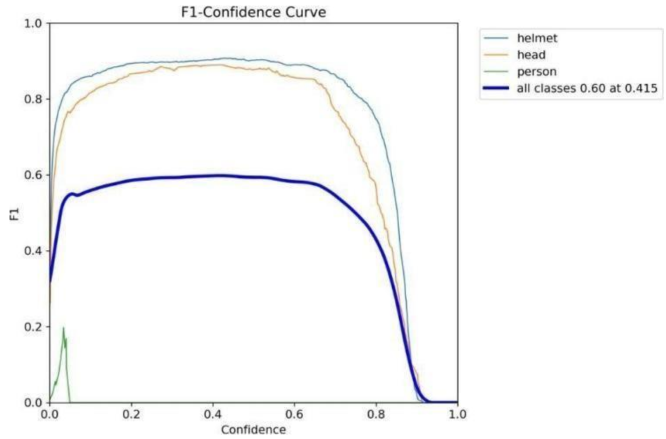


Figure 5. F1-Score

The F1-score curve in Figure 5 depicts the model's performance after training at epoch 50, displaying the F1-score values for each class and the overall training performance.

This visualization indicates that the overall F1-score for all classes is 79%. Specifically, the helmet class achieves an F1-score of 85%, while the head class achieves 71%. It's important to note that values closer to 1.0 or 100 indicate the model's best performance in accurately detecting all objects in images and maintaining a good balance in its predictions. Conversely, values closer to 0 indicate the poorest performance in object detection.

Furthermore, in Figure 4.4, the model trained at epoch 50 accurately predicts the helmet class with 87% accuracy and the head class with 83% accuracy. This demonstrates that a higher True Positive (TP) rate in object prediction results in better overall prediction performance by the model.

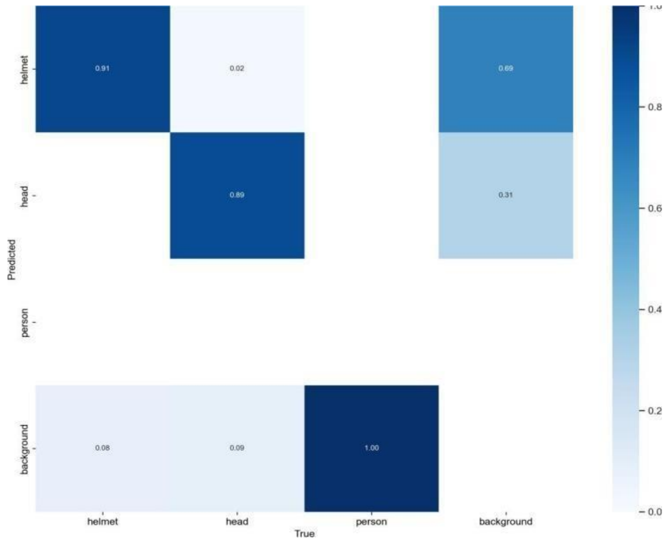


Figure 6.Confusion Matrix

This study evaluated the performance of the trained model using two key metrics: F1-score and Mean Average Precision at 0.5 Intersection over Union (mAP@0.5 IoU). F1-score provided a comprehensive assessment of the model's detection capabilities, balancing precision (the proportion of true positives among all detections) and recall (the proportion of actual positives correctly identified) for object detection tasks. Conversely, mAP@0.5 IoU evaluated the model's ability to accurately localize objects. F1-score emphasized the model's capacity to comprehensively identify all objects within an image, while mAP@0.5 IoU focused on the precision of object localization. The training process employed a learning rate of 1e-4 and spanned 150 epochs.

D. *Testing*

Table 1 presents a average of the system evaluation results for a range of detection distances (2m, 4m, 6m) and lighting conditions (morning, afternoon, night) in both indoor and outdoor scenarios. The tabulated values represent the detection accuracy of safety helmets.

Table 1. Table of Average Test Accuracies

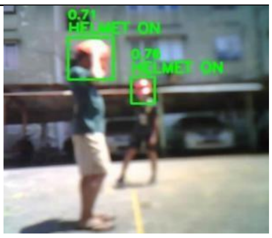
Distance	Time	Indoor	Outdoor
2 meters	Morning	0.79	0.81
	Afternoon	0.82	0.85
	Night	0.81	-
4 meters	Morning	0.71	0.76
	Afternoon	0.81	0.82
	Night	0.68	-
6 meters	Morning	0.62	0.67
	Afternoon	0.67	0.61
	Night	0.63	-

Table 2 presents the average detection delay of the deep learning-based safety helmet warning system under various lighting conditions (morning, afternoon, night) and detection distances (2m, 4m, 6m). A shorter delay indicates a more responsive system, which is crucial for ensuring worker safety. These results complement the accuracy findings in Table 1, providing a more comprehensive overview of the system's performance.

Table 2. Table of Average Delay Test Results

Distance	Time	Indoor	Outdoor
2 meters	Morning	5.3	5.4
	Afternoon	5.5	5.1
	Night	7.1	-
4 meters	Morning	6.7	5.6
	Afternoon	7.1	6.3
	Night	7.4	-
6 meters	Morning	8.2	7.6
	Afternoon	8.5	8.6
	Night	11.1	-

Table 3. Testing Result

Distance (m)	State	Dataset
1-6	All workers are required to wear helmets while maintaining an equal distance from one another.	
1-6	Two workers noted, one with proper head protection and another lacking appropriate PPE.	
1-6	The worksite layout allows for proper head protection for all personnel, regardless of location	
1-6	Discrepancy identified in PPE utilization among on-site personnel. One worker was noted to be without a safety helmet.	
1-6	Non-compliance with safety regulations was observed. Three workers were identified not wearing the necessary head protection gear	
1-6	A discrepancy in PPE (Personal Protective Equipment) utilization was noted. Two workers were compliant with head protection requirements, but one individual was not wearing a helmet and was seen holding it in their hand.	
1-6	Non-compliance with safety regulations was evident at varying distances across the worksite. One worker was observed wearing a helmet, while another worker was not wearing a helmet.	

IV. Conclusion

1. The developed system integrates a helmet usage warning system using the YOLOv8 Deep Learning algorithm. The system is composed of several integral components, including ESP32Cam, FT232RL, Buzzer BoardCH340, and USB cables.
2. The system has been developed to incorporate a helmet usage warning mechanism using the YOLOv8 Deep Learning algorithm. It comprises various essential components, including ESP32Cam, FT232RL, Buzzer BoardCH340, and USB cables.
3. As per the research findings, YOLOv8 can perform quick inference, which is particularly crucial in security applications where real-time object detection is essential. Faster object identification can be advantageous in situations requiring prompt responses.
4. The adaptability to changes in lighting is a crucial consideration in the practical application of this system. YOLOv8 has demonstrated strong object detection capabilities across various lighting conditions, including low-light environments.
5. A primary application derived from the results section is the development of an automated system capable of detecting helmet usage through image or video analysis. This system can be integrated into various devices, such as CCTV cameras, to provide real-time alerts within industrial area, thus can minimizing workplace accidents.

Acknowledgement

We are grateful to Mrs Dharma Aryani, S.T., M.T., PhD. and Mr Asriyadi, S.ST., M.Eng. for their guidance, support and technical assistance.. Finally, we thank our families and friends for their unwa

References

[1] A. Alsharef, A. Albert, I. Awolusi, and E. Jaselskis, "Severe injuries among construction workers: Insights from OSHA's new severe injury reporting program," *Saf Sci*, vol. 163, p. 106126, Jul. 2023, doi: 10.1016/j.ssci.2023.106126.

[2] A. S. B. Putra, E. D. Kusumawati, and D. Kartikasari, "Unpacking the Roots and Impact of Workplace Well-being: A Literature Review," *International Journal of Multidisciplinary Approach Research and Science*, vol. 2, no. 01, pp. 312–321, Dec. 2023, doi: 10.59653/ijmars.v2i01.433.

[3] Y. Adiratna *et al.*, *Profil Keselamatan dan Kesehatan Kerja Nasional Indonesia Tahun 2022*, Cetakan pertama. Jakarta Selatan: Kementerian Ketenagakerjaan Republik Indonesia, 2022.

[4] J. Wang and S. Razavi, "Network-Based Safety Leading Indicators for Safety Risk Analysis in Construction," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 63, no. 1, pp. 1787–1791, Nov. 2019, doi: 10.1177/1071181319631272.

[5] Z. Wang, Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao, "Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches," *Sensors*, vol. 21, no. 10, p. 3478, May 2021, doi: 10.3390/s21103478.

[6] A. Hayat and F. Morgado-Dias, "Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety," *Applied Sciences*, vol. 12, no. 16, p. 8268, Aug. 2022, doi: 10.3390/app12168268.

[7] A. Moohialdin, F. Lamari, M. Marc, and B. Trigunarsyah, "A Real-Time Computer Vision System for Workers' PPE and Posture Detection in Actual Construction Site Environment," 2021, pp. 2169–2181. doi: 10.1007/978-981-15-8079-6_199.

[8] B. Widodo, H. A. Armanto, and E. Setyati, "Deteksi Pemakaian Helm Proyek Dengan Metode Convolutional Neural Network," *Journal of Intelligent System and Computation*, vol. 3, no. 1, pp. 23–29, Apr. 2021, doi: 10.52985/insyst.v3i1.157.

[9] H. Nugraha, "ANALISIS PELAKSANAAN PROGRAM KESELAMATAN DAN KESEHATAN KERJA DALAM UPAYA MEMINIMALKAN KECELAKAAN KERJA PADA PEGAWAI PT. KERETA API INDONESIA (PERSERO)," *Coopetition : Jurnal Ilmiah Manajemen*, vol. 10, no. 2, Nov. 2019, doi: 10.32670/coopetition.v10i2.43.

[10] A. KORKMAZ and M. T. AĞDAŞ, "Deep Learning-Based Automatic Helmet Detection System in Construction Site Cameras," *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, vol. 12, no. 3, pp. 773–782, Sep. 2023, doi: 10.17798/bitlisfen.1297952.

[11] J. Lee and S. Lee, "Construction Site Safety Management: A Computer Vision and Deep Learning Approach," *Sensors*, vol. 23, no. 2, p. 944, Jan. 2023, doi: 10.3390/s23020944.

[12] R. Geng, Y. Ma, and W. Huang, "An improved helmet detection method for YOLOv3 on an unbalanced dataset," in *2021 3rd International Conference on Advances in Computer Technology,Information Science and Communication (CTISC)*, IEEE, Apr. 2021, pp. 328–332. doi: 10.1109/CTISC52352.2021.00066.

[13] J. Wu, N. Cai, W. Chen, H. Wang, and G. Wang, "Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset," *Autom Constr*, vol. 106, p. 102894, Oct. 2019, doi: 10.1016/j.autcon.2019.102894.

[14] M. Jakubec, E. Lieskovska, A. Brezani, and J. Tothova, "Deep Learning-Based Automatic Helmet Recognition for Two-Wheeled Road Safety," *Transportation Research Procedia*, vol. 74, pp. 1171–1178, 2023, doi: 10.1016/j.trpro.2023.11.258.

[15] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[16] J. Kang, S. Tariq, H. Oh, and S. S. Woo, "A Survey of Deep Learning-Based Object Detection Methods and Datasets for Overhead Imagery," *IEEE Access*, vol. 10, pp. 20118–20134, 2022, doi: 10.1109/ACCESS.2022.3149052.

[17] M. Browne and S. S. Ghidary, "Convolutional Neural Networks for Image Processing: An Application in Robot Vision," 2003, pp. 641–652. doi: 10.1007/978-3-540-24581-0_55.

[18] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[19] A. Chandio *et al.*, "Precise Single-stage Detector," 2022.

[20] Y. Yin, H. Li, and W. Fu, "Faster-YOLO: An accurate and faster object detection method," *Digit Signal Process*, vol. 102, p. 102756, Jul. 2020, doi: 10.1016/j.dsp.2020.102756.

[21] X. Lei, "OBJECT DETECTION FOR PERCEPTUALLY-DEGRADED ENVIRONMENTS," California State Polytechnic University, Pomona, 2020.

[22] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach Learn Knowl Extr*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023, doi: 10.3390/make5040083.