CHAPTER 25

Automated Construction Safety Inspection System Leveraging Computer Vision

Chinmay Kavatagimath[1], Shivani R.[2], Tarun K.[1] and Hruthik R.[1]

ABSTRACT

The construction industry stands as one of the most hazardous sectors globally, accounting for over 60,000 fatalities annually. According to Heinrich, a significant portion of accidents (88%) are attributed to unsafe acts, while 10% result from unsafe conditions, and 2% arise from uncontrollable factors. Of these, 98% of accidents can potentially be mitigated through the implementation of safe practices and ensuring proper site conditions. Despite this, past examinations have indicated a deficiency in manual safety compliance mechanisms for enforcing safety protocols on construction sites. The advent of deep learning algorithms and advanced image-capturing technologies has offered safety professionals valuable tools to bolster safety management strategies. However, a comprehensive understanding of these advancements remains lacking within the existing body of knowledge. Therefore, conducting a systematic literature review on this subject is imperative. This study employed various pertinent keywords to explore literature within the Scopus database. Utilizing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach, relevant literature was identified and analyzed. The findings of this review shed light on prominent researchers, institutions, and countries engaged in this domain. Furthermore, it pinpointed both highly focused and overlooked research themes. extracting future research directions from top-tier investigations. This study serves as a foundational resource for future inquiries in this field, offering safety professionals insights into automated methods for ensuring safety compliance on construction projects.

Keywords: Computer Vision; Safety management; Automation; Construction Hazards; Construction safety.

1.0 Introduction

1.1 Background

Construction ranks among the most hazardous industries worldwide, comprising about 7% of the workforce while being responsible for nearly 30% of work-related fatalities (Fang *et al.*, 2020).

[1]*School of Construction, NICMAR University, Pune, Maharashtra, India*
[2]*Corresponding author, School of Construction, NICMAR University, Pune*
(E-mail: P2370179@student.nicmar.ac.in)

DOI: 10.17492/JPI/NICMAR/2507025 *ISBN: 978-93-49790-54-4*

In India, the risk increases because of rapid infrastructure development. The NCRB reports that more than 10,000 workers die each year due to occupational hazards, with construction sites being a significant factor. Traditionally, in India, site safety relies on inconsistent human oversight. With the help of real-time video analytics, computer vision provides a proactive alternative for identifying hazards and guaranteeing PPE compliance. Firms employing this technology have experienced a reduction in accidents by 20–30%, along with decreased medical, legal, and insurance expenses, as well as productivity improvements of up to 15%. The eventual incorporation of IoT and predictive analytics may enhance outcomes even more. Implementation encompasses the preparation of datasets, training of models, and deployment in real-world scenarios. The YOLOv5 algorithm (Li *et al.*, 2023)can accurately identify compliant and non-compliant workers using annotated images of PPE violations

1.2 Need for study

Unlike the automotive or manufacturing sectors, construction contends with distinct safety challenges arising from its ever-evolving work environments and lack of standardized procedures. Real-time monitoring and early hazard detection are made possible by computer vision technology (Li *et al.*, 2023), which greatly enhances site safety and lowers the incidence of accidents (Yang *et al.*,2022).

1.3 Research objectives

This research article aims to achieve the following objective:(Yang *et al.*,2022)

- Examine how widely Computer Vision is currently being implemented in construction safety.
- Develop an object recognition model using CCTV footage, trained with the YOLOv5 algorithm, capable of identifying construction workers along with their safety vests and helmets.

2.0 Literature Review

2.1 Computer vision

Computer vision, a branch of computer science, has extensive applications in speech recognition, natural language processing, and robot control. Traditional machine learning techniques cannot process data in its raw state. (Fang *et al.*, 2020) In the realm of AI and deep learning, (Kompare *et al.*, 1994)characterizes computer vision as "a technology that enables machines to interpret and understand the content of digital images and video." This description emphasizes the development of computer vision, particularly through deep learning methods, which have significantly enhanced the field's capabilities.

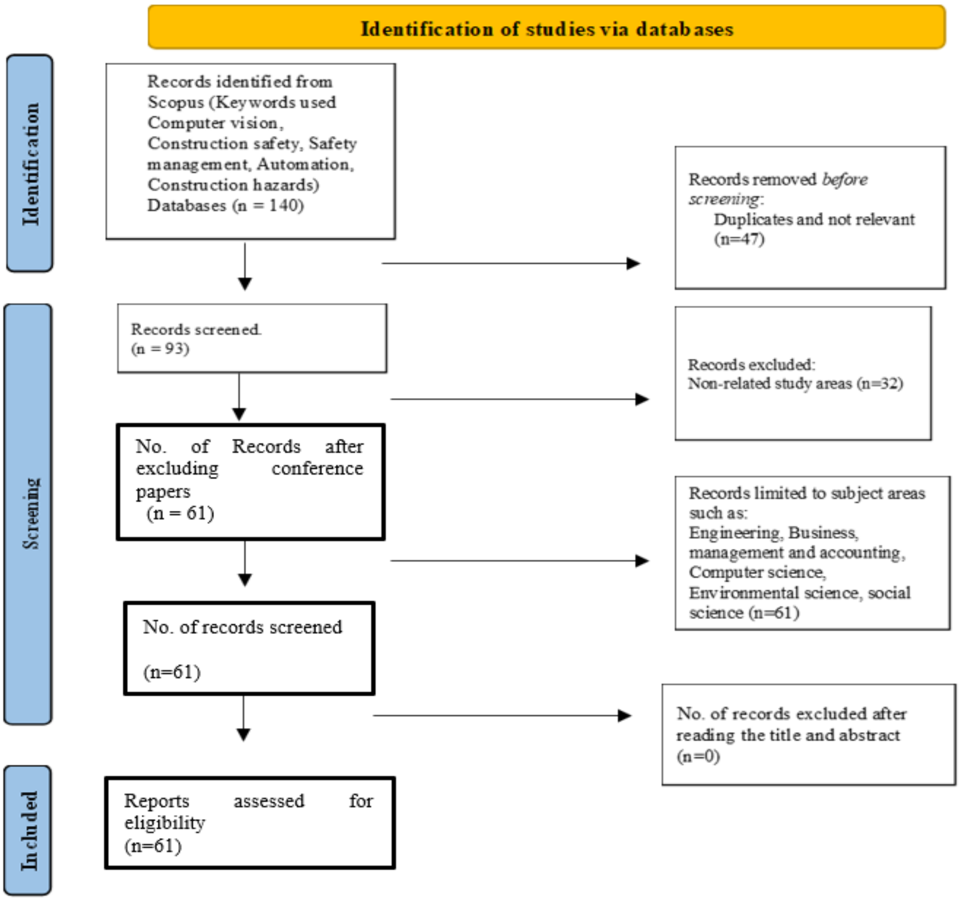
2.2 Construction safety

Construction safety, as outlined by (Zhang *et al.*, 2019), refers to "the identification, assessment, and management of risks linked to construction activities in order to safeguard

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

workers and reduce incidents." As stated by OSHA (2021), it encompasses "policies and procedures to reduce injury risks and ensure workers have access to safety training and equipment."

Figure 1: Identification of Relevant Papers using PRISMA



2.3 Systematic Literature Review (SLR) using "PRISMA"

This study investigates the relationship between Construction Safety and Computer Vision through the use of CNN and the PRISMA methodology (Sharma & Laishram, 2024)

Stage I delineates the SLR process, which includes identifying keywords and selecting databases with Boolean operators to establish the review scope.

Stage II encompasses the identification, screening, and selection of pertinent studies.

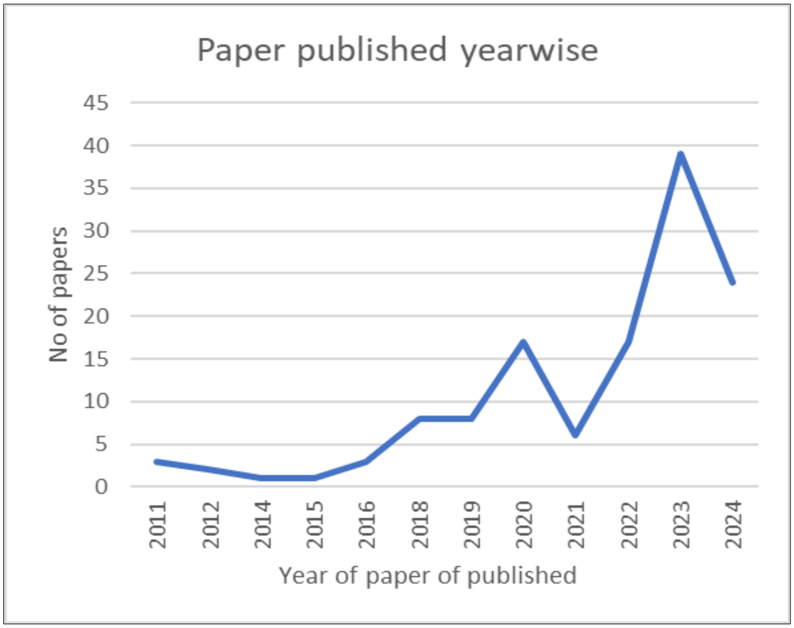
Stage III encompasses the analysis and reporting of findings, with an emphasis on theory, methods, and study locations to identify research gaps and future directions in construction safety.

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

Table 1: Paper Published Year-wise

Year	No. of Paper	Year	No. of Paper	Year	No. of Paper
2010	0	2015	1	2020	17
2011	3	2016	3	2021	6
2012	2	2017	0	2022	17
2013	0	2018	8	2023	39
2014	1	2019	8	2024	24

Figure 2: Analysis of Papers Published



2.4 Literature review using the "TCCM" method

Theories-characteristics-contexts-methods (TCCM) framework for a systematic search and analysis of articles, with the goal of addressing the following questions: Which theories, variables, and contexts are being applied by researchers? What are the options and future pathways? (Ip & Chark, 2025)

Technology (T): Employ computer vision grounded in YOLOv5 for the real-time identification of personnel and PPE in CCTV footage, ensuring compliance with safety regulations and detection of risks.

Context (C): Focus on Indian construction sites, addressing safety oversights, human errors, and delays in hazard identification due to environmental factors.

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

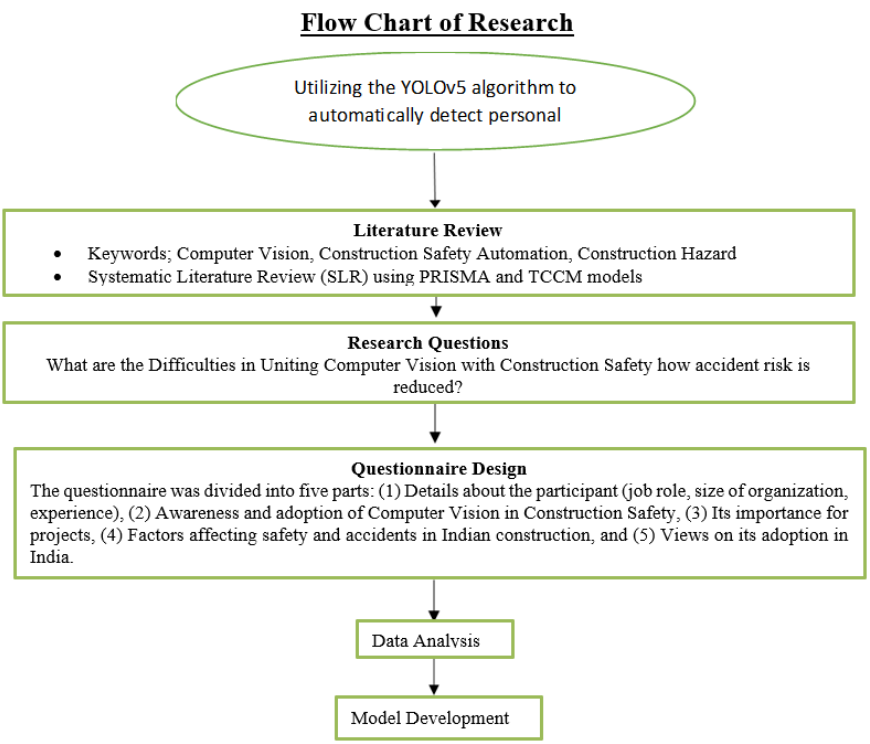
Content (C): Utilize computer vision alongside safety measures to automate hazard identification and notifications, improving reaction times and site safety.

Methodology (M): Train YOLOv5 using current and historical CCTV footage. Evaluate performance by comparing detection accuracy, accident rates, and response efficiency with traditional methods

2.5 Research gap

Lack of Comprehensive Studies: While many studies explore the use of computer vision in construction, most focus on aspects like object detection, PPE compliance, or hazard recognition. No comprehensive study exists that brings together various applications and evaluates their effectiveness in real-world scenarios.

Figure 3: Flow Chart of Research



3.0 Methodology

Literature review & industry analysis: Examination of scholarly articles, case studies, and industry reports to assess the role of computer vision in construction safety in India concerning implementation levels, challenges, and benefits.

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

Survey for identifying crucial parameters: Questionnaire survey targeting site managers and safety officers to identify critical activities and primary focus on PPE monitoring. This will establish the focus of the object recognition model.

Baseline safety assessment: Collection of pre-implementation data through CCTV surveillance to investigate adherence to PPE and unsafe actions, establishing a reference point for comparison. YOLOv5 Object Recognition Model Development – Creating a computer vision model to detect workers and PPE using labeled data, iterative refinements, and accuracy checks.

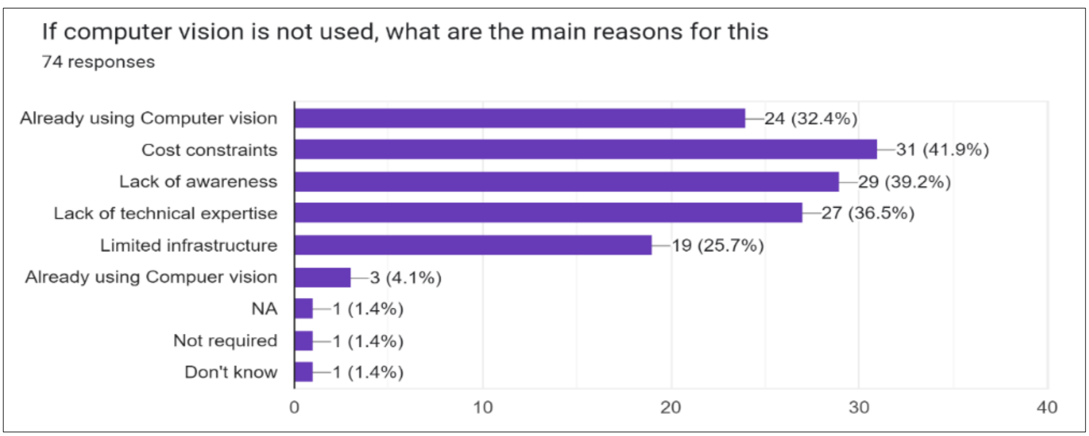
4.0 Data Analysis

This study analyzes responses from 74 construction professionals to assess key factors affecting Computer Vision (CV) and Construction Safety (CS) integration.

Table 2: Data Analysis

Category	Details	Number of Respondents	Percentage (%)
Gender	Male	60	81.08%
	Female	14	18.91%
Work Experience	1-3 years	56	75.67%
	3-6 years	11	14.86%
	6-9 years	5	6.75%
	More than 9 years	2	2.70%
Company Size	1-50 employees	30	37.80%
	51-200 employees	8	10.80%
	201-500 employee	7	9.50%
	500+ employees	31	41.90%

Figure 4: Response for 'Key Reasons for not Utilizing Computer Vision'

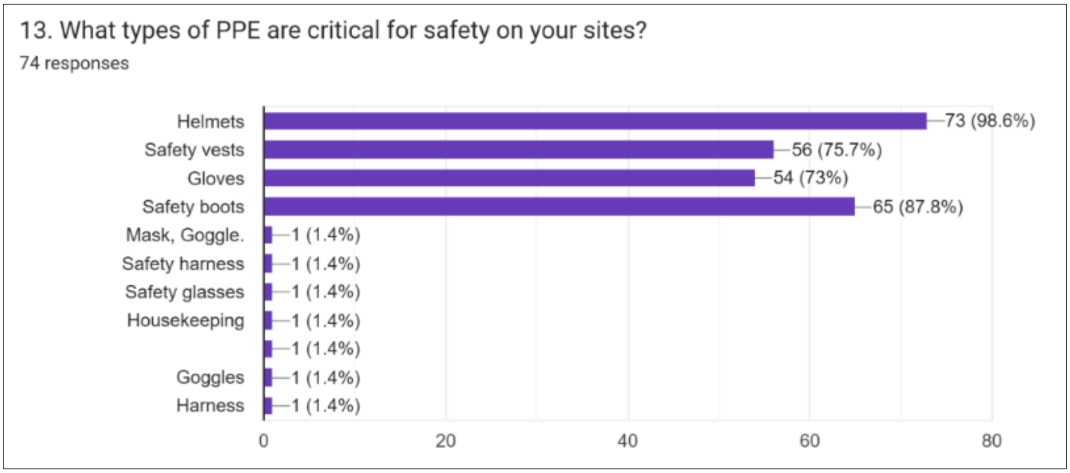


DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

4.1 Inference from survey

According to the survey, the primary obstacles to adopting computer vision in construction are a lack of awareness and technical expertise, with cost concerns following closely behind. All respondents, however, indicated a readiness to embrace the technology after successful pilot tests, acknowledging its potential to improve site safety. In light of these findings, our model based on YOLO will zero in on identifying whether helmets and safety vests—designated as the most vital PPE—are present or absent. An alert system that operates in real-time will be incorporated to inform site managers of any instances of safety non-compliance.

Figure 5: Response for 'Which kind of PPE is Crucial'



5.0 Model Development

5.1 Overview of YOLOv5

YOLOv5s is a fast and efficient deep-learning (Ngoc-Thoan *et al.*, 2023) model capable of quickly recognizing multiple objects within an image. The "s" stands for "small," suggesting that it requires less computational resources while still being equally effective.

5.2 The architecture of YOLOv5

- Backbone employs Cross Stage Partial Network (CSPNet) to reduce processing load and improve learning Employs convolutional and bottleneck layers to derive fundamental visual traits.

DOI: 10.17492/JPI/NICMAR/2507025 *ISBN: 978-93-49790-54-4*

- Neck uses Path Aggregation Network (PANet) for the combination of features from different scales. Assists in recognizing items of different sizes, useful for construction locations where laborers are positioned at varying distances.
- Head Utilizes anchor boxes and bounding box regression for precise object detection. Items are classified and assigned confidence scores (e.g., Helmet, Safety Vest, No-Helmet, No-Vest).

The integration of CSPNet, PANet, and the YOLO head enhances YOLOv5 speed, accuracy, and efficiency.

Table 3: YOLOv5 Performance

Metric	YOLOv5s Performance
Input Resolution	600 × 640 pixels
Inferences Speed	~ 140 FPS (Tesla V100 GPU)
mAP (mean Avg. Precision)	~ 45% - 50% (varies with dataset)
Model Size	~ 14 MB
Training Time	Faster than other YOLO version

5.3 Processes involved

Data Collection and Preparation: Acquire images/videos of employees, both with and without PPE. Augment, preprocess, and label data to improve model training.

Training and Selection of Model: Train the YOLOv5 model for PPE detection using annotated data Make use of transfer learning to improve precision and efficacy. Utilize metrics like mAP and recall validating performance.

Integration and Implementation: Employ OpenCV for real-time video processing. Enhance detection in diverse lighting situations.

Detection and Alert System: Utilize bounding boxes to pinpoint workers who are not compliant. Deliver immediate alerts via text messages, e-mails, or alarms. Provide a dashboard for monitoring compliance.

Testing and Optimization: Evaluate accuracy based on real site measurements. Modify thresholds to reduce the occurrence of false positives. Comparison with SSD and Faster R-CNN models.

5.4 Deployment and monitoring

Integrate with CCTV for instant use. Enable remote monitoring via the cloud.

6.0 Discussion and Results

6.1 Utilized dataset

Images in the dataset show employees with Safety vests and helmets. Among the data

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

sources are Open-source datasets, such as the COCO, Kaggle, and PPE Detection datasets. Actual photos were taken on building sites in a variety of lighting and environmental settings.

Figure 6: Some of the Pictures used to Train the Model

6.2 Detection of helmet and safety vest through YOLO V5 model with co-authors

Figure 7: Before the Application of YOLO

DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

Figure 8: After the Application of YOLO

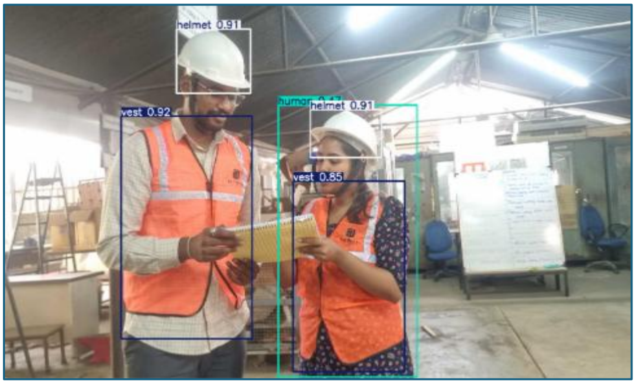
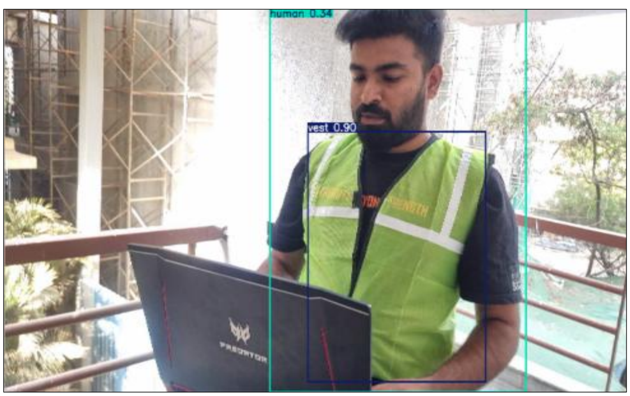


Figure 9: Without Helmet and Safety Vest

Figure 10: With Safety Vest and no Helmet



DOI: 10.17492/JPI/NICMAR/2507025 ISBN: 978-93-49790-54-4

6.3 Testing of YOLO V5 with a group of workers on site

Figure 11: Testing of YOLO V5 with a Group of Workers on Site

Figure 12: Alert Notification Through Telegram of a Missing Element with Picture

DOI: 10.17492/JPI/NICMAR/2507025 *ISBN: 978-93-49790-54-4*

7.0 Conclusion

Construction is one of the most hazardous industries, where frequent accidents arise from human limitations and delays in recognizing risks. As conventional safety measures often fail, there is a need for advanced solutions. Computer vision, and intense learning models like YOLOv5, have proven effective in bolstering safety monitoring by identifying PPE such as helmets and vests in real-time, thus reducing workplace hazards. A literature review highlighted the global efficacy of AI-driven safety solutions but pointed out their slow adoption in India due to factors like high costs, limited awareness, and insufficient regulations.

To address this problem, a model based on YOLOv5 was developed and trained with images of construction workers. When integrated with CCTV systems for real-time monitoring, it demonstrated a high level of accuracy. This model generated automatic alerts for violations, resulting in a marked enhancement of compliance enforcement. A survey of 74 industry professionals identified critical challenges, including a lack of technical expertise, budget constraints, and resistance to change. Nevertheless, the respondents acknowledged that effective pilot projects could ease broader implementation.

References

Fang, W., Ding, L., Love, P. E. D., Luo, H., Li, H., Peña-Mora, F., Zhong, B., & Zhou, C. (2020). Computer vision applications in construction safety assurance. In *Automation in Construction* (Vol. 110). Elsevier B.V. https://doi.org/10.1016/j.autcon.2019.103013

Ip, M. M. H., & Chark, R. (2025). The compass for restaurant menu design experiments: Dissecting restaurant menu design through the TCCM framework. *International Journal of Hospitality Management*, *127*. https://doi.org/10.1016/j.ijhm.2025.104118

Kompare, B., Bratko, I., Steinman, F., & Džeroski, S. (1994). Using machine learning techniques in the construction of models I. Introduction. *Ecological Modelling*, *75–76*(C), 617–628. https://doi.org/10.1016/0304-3800(94)90054-X

Li, P., Wu, F., Xue, S., & Guo, L. (2023). Study on the Interaction Behaviors Identification of Construction Workers Based on ST-GCN and YOLO. *Sensors*, *23*(14). https://doi.org/10.3390/s23146318

Ngoc-Thoan, N., Bui, D. Q. T., Tran, C. N. N., & Tran, D. H. (2024). Improved detection network model based on YOLOv5 for warning safety in construction sites. *International Journal of Construction Management*, *24*(9), 1007–1017. https://doi.org/10.1080/15623599.2023.2171836

DOI: 10.17492/JPI/NICMAR/2507025 *ISBN: 978-93-49790-54-4*

Ngoc-Thoan, N., Bui, D.-Q. T., Tran, C. N. N., & Tran, D.-H. (2023). Improved detection network model based on YOLOv5 for warning safety in construction sites. *International Journal of Construction Management*. https://doi.org/10.1080/15623599.2023.2171836

Seo, J., Han, S., Lee, S., & Kim, H. (2015). Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, *29*(2), 239–251. https://doi.org/10.1016/j.aei.2015.02.001

Sharma, N., & Laishram, B. (2024). Understanding the relevance and impact of the cost of quality in the construction industry: a systematic literature review using PRISMA. *Construction Innovation*. https://doi.org/10.1108/CI-08-2023-0197

Yang, W., Zhou, G. Le, Gu, Z. W., Jiang, X. D., & Lu, Z. M. (2022). Safety helmet wearing detection based on an improved YOLOV3 scheme. *International Journal of Innovative Computing, Information and Control*, *18*(3), 973–988. https://doi.org/10.24507/ijicic.18.03.973

Zhang, M., Cao, T., & Zhao, X. (2019). Using smartphones to detect and identify construction workers' near-miss falls based on ANN. *Journal of Construction Engineering and Management*, *145*(1). https://doi.org/10.1061/(asce)co.1943-7862.0001582

DOI: 10.17492/JPI/NICMAR/2507025 *ISBN: 978-93-49790-54-4*