

Analyzing Safety Risk Variables in Real time at Construction Sites using YOLOv8 Architecture

Danish Khan^{1,†}, Kumar Tejashwa^{2,†}, Sushruta Mishra^{3,†}

^{1,2,3} Kalinga Institute of Industrial Technology, Deemed to be University, India

Abstract

Safety in construction is a vital matter that impacts the lives and welfare of both employees and the public. Although it is a crucial part of construction safety management, hazard detection is frequently hampered by environmental constraints and human factors. In this study, we provide a novel approach that uses the most advanced deep learning model for object detection—YOLOv8 multi-class classifier—to identify and categorize hazards on construction sites. To achieve this, we gathered and created a curated list of construction sites photos with many kinds of dangers, like objects falling from height, Live Electrical Lines, Fire, and workers without proper protective equipment (PPE) kits. This Dataset of photos was analyzed by us and based on this we trained and assessed our model. We were able to detect multiple violations and risks within a single frame. This approach shows promising results in enhancing hazard identification and management, in construction safety practices.

Keywords

YOLOv8, Deep Learning, Hazards, Recall, Personal Protective Equipment, Safety Management

1. Introduction

Ensuring safety at construction sites is a task of great importance as it greatly impacts the health of workers and the community at large. Shocking figures from the International Labour Organization reveal that one fifth of incidents leading to the deaths of over 2 million workers annually due to work related diseases and accidents are attributed to the construction sector.

Minimizing safety risks on construction sites relies heavily on recognizing dangers but this crucial task is often hindered by various obstacles. Factors related to behavior like lack of awareness, inadequate training and limited supervision as well as issues such as fatigue, stress and distractions can significantly impact workers decision making and performance. Adding to these challenges are factors such as the unpredictable nature of construction sites along with external elements like weather conditions, limited lighting and background noise. These combined difficulties make it challenging to identify and address hazards ultimately raising the likelihood of accidents and injuries, in construction settings.

In the face of challenges there is a growing need for methods and tools to improve hazard detection and classification in the construction field.

Securing Next-Generation Systems using Future Artificial Intelligence Technologies, August 08–9, 2024, Dept of CSE, Maharaja Agrasen Institute of Technology, New Delhi, India

* Corresponding author: Sushruta Mishra(sushruta.mishrafcs@kiit.ac.in)

† These authors contributed equally.

Kalinga Institute of Industrial Technology, Bhubaneswar, India

{danishnadeem2012@gmail.com, ktejashwa80@gmail.com, sushruta.mishrafcs@kiit.ac.in}



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Our study introduces a strategy to transform how hazards are identified and categorized on construction sites. We utilize technologies, like a multi class classifier and YOLOv8 a powerful deep learning model known for its object detection capabilities.

To support our approach, we carefully labeled a dataset of construction site images showing various hazards such as falling objects, electrical risks, fires and workers without proper safety gear. Using this dataset, we conducted training and evaluation of our model rigorously comparing it with existing methods. Our research focuses on assessing how well our approach performs in areas like accuracy, recall rates and speed when detecting and classifying hazards found at construction sites.

The significance of our paper goes beyond academic discussions but has a profound impact on construction safety and hazard detection. Our innovative method can accurately spot hazards in an image marking a significant shift in how construction safety is managed. By providing feedback to site managers and workers our approach shows potential in improving the effectiveness of safety measures thus reducing the likelihood of accidents and injuries at construction sites. Through an examination of our method and its implications we pave the way for a future, for everyone involved in the construction industry.

The main goal of our proposed research can be summarized as follows:

- **Developing our dataset:** The quality of the training data significantly influences the performance of a learning model. Therefore, we utilized our custom web crawler to curate a comprising image captured by ourselves and images sourced from existing datasets. [12].
- **Image preprocessing:** To increase the size of the dataset, this research uses image augmentation techniques to produce images at different scales.
- **Object detection model:** Rather than suggesting a new architecture for YOLO, we examine version 8, the most recent iteration of the program, and capitalize on its capabilities.

2. Literature Review

In recent years there has been a rise in research focused on using advanced technologies to improve safety practices within the construction sector. These investigations explore approaches, such as vision based surveillance and machine learning methods with the goal of boosting safety standards which help avert mishaps and detect injuries. M. Z. Shanti et al.[1] created a construction safety inspection protocol using AI for the UAE. Their study focuses on enhancing safety, at construction sites by utilizing intelligence. The protocol provides advantages like improved efficiency, accuracy and effectiveness in hazard detection ultimately raising safety levels, in the UAEs construction sector. Jeon et al. [2] proposed to categorize types of risks at construction sites and analyze conditions through the use of wearable EEG devices. This innovative method aims to enhance safety on construction sites by providing workers with insights into their alertness levels and their ability to identify and respond to dangers. The research conducted, contributes to the adoption of safety practices in the construction sector potentially reducing incidents at worksites. M. Alkaissy et al. [3] research project delves into enhancing safety, within the construction sector by utilizing machine learning techniques to classify types of injuries. Through the application of different approaches, the study seeks to improve the identification and prevention of injuries, on construction sites. This research adds to the ever-growing collection of knowledge dedicated to minimizing workplace incidents and fostering a culture of safety in the construction industry. S.Yin

et al. [4] created a method to assess the safety practices of construction workers using machine learning. Its goal is to improve safety measures, in construction sites by understanding worker behaviour and pinpointing risks associated with it. J. C. P. Cheng et al. [5] suggested to concentrate on using vision based methods to oversee safety adherence, at construction sites. Specifically, the research involves categorizing gear and identifying workers to enhance safety supervision through the use of computer vision technology this method strives to enhance the real time surveillance of safety protocols ensuring compliance with safety standards and minimizing the likelihood of accidents. J. Lee et al. [6] explored the use of computer vision and learning techniques to enhance safety protocols at construction sites. The aim is to boost real time surveillance and detection of safety risks through cutting edge technologies reducing the risk of accidents and promoting a safe and secure workplace environment. Hou et al. [7] research delves into the application of learning techniques in enhancing safety protocols within the Architecture Engineering and Construction (AEC) industry. It discusses the potential of learning to revolutionize safety measures and mitigate risks in construction projects through case studies. The study offers insights to enhance safety management practices in the AEC sector by evaluating the advantages and limitations of integrating deep learning technology in this domain. H.T.T.L. Pham et al.[8] conducted a research to explore the utilization of learning in enhancing safety protocols within the construction sector both in its state and anticipated future advancements. The study delves into the implementation of deep learning methods highlighting their effectiveness while also discussing other areas for additional research and integration into safety management practices in construction. Y. Zhao et al. [9] study explored the utilization of deep learning methods to improve safety protocols at construction sites. By leveraging algorithms, the research endeavours to identify hazards and monitor the whereabouts of objects or people, within construction zones. Through techniques like object recognition and predicting trajectories the investigation aims to reduce incidents and bolster safety standards on construction sites. In essence by tackling concerns related to risk assessment and safety surveillance in this industry this initiative propels the field of intelligence, in ensuring building safety. W. Fang et al. [10] research delved into enhancing safety protocols, in the construction sector through the integration of learning and computer vision technologies. Its primary objective is to utilize algorithms to link images capturing practices with established safety standards. By harnessing these tools, the study strives to enhance safety surveillance at construction sites by identifying behaviours and upholding safety protocols. Overall, the investigation leverages intelligence resources to enhance safety practices and mitigate accidents, within the construction field.. K. Kim et al. [11] conducted a research project to investigate how YOLO v5 and v8 advanced object detection algorithms can be utilized to detect safety concerns, in construction settings. The goal is to improve safety supervision and reduce risks at construction sites using these algorithms promoting risk management strategies. M. Ferdous et al. [12] published a study detailing a YOLO based approach aimed at recognizing PPE kits at construction sites to enhance safety supervision and minimize hazards. The research underscores the importance of technology driven strategies, in encouraging compliance, with safety guidelines and ensuring the welfare of construction personnel. Roboflow Universe Projects et al.[13] "Data, on Safety at Construction Sites", from Roboflow Universe dated August 2023.

3. Dataset Preparation

In this study a collection of data known as the CSS (Construction Site Safety) dataset has been developed to recognize individuals, vests, masks, helmets and various equipment. Stringent guidelines have been implemented to ensure the model's effectiveness in real world applications. The assessment includes the following factors:

1. Background in related construction

2. Gestures made by humans
3. Object angles and distances
4. Number of classes.

Images with multiple class instances were preferred. The images sourced are from Google Images where they were collected using a custom web-crawler using BeautifulSoup4 Python. For picture pre-processing, the [12] Roboflow platform was utilized. Each image's pertinent regions had bounding boxes carefully drawn by the researcher, who also gave the appropriate class.

We resize images to 640x640 resolution, this helps standardize the input and maintain consistency in the dataset. This in turn simplifies processing and boosts accuracy.

We also Normalize Pixel Values of our 8-bit images by using the eq (1).

$$\text{Normalized Value} = \frac{\text{Original Pixel Value}}{255} \quad (1)$$

Employing this technique stabilizes the learning process, equalizing influence by preventing larger values to dominate the learning process. Also, avoids saturation of activation functions. Another Data-Preprocessing technique we use is Data-Augmentation, used to artificially increase the size and diversity of our dataset by rotation, flipping, changing saturation and contrast of images. This helps avoid overfitting and helps the model to better generalize. We have a total of 2801 images of resolution 640x640 and 10 different detection classes (Vehicle, Machinery, Safety Vest, Safety Cone, Person, NO-safety vest, NO-Mask, NO- Hardhat, Mask, Hardhat). The Test, Validation and Train split is Test: 2605 (93%), Valid: 114 (4%), Test: 82 (3%). The data split between different classes is as follows in Table 1.

Table 1

Image data split between all 10 classes.

Mode	Train	Valid	Test
Hardhat	1314	42	30
Mask	1096	19	16
NO-Hardhat	1380	37	25
NO-Mask	1531	44	30
NO-safety vest	1864	56	36
Person	2526	84	59
Safety Cone	631	13	8
Safety Vest	1319	28	22
Machinery	2101	26	22
Vehicle	744	16	15
Data Volume	2605	114	82

4. Proposed Model

This research's main objective is to increase the accuracy of danger identification by merging pre-processing approaches for data with already-existing models. The framework of YOLOv8, which is used as the main model for PPE detection in this work, is displayed in Figure 2. The remarkable

network structure of YOLOv8 converts target detection into a regression problem. In order to provide quick detection and satisfy real-time demands, it imminently offers bounding box coordinates and categorizes targets across various locations in the picture. The proposed project's methodology comprises the subsequent primary steps:

- The training and validation datasets are constructed using annotated image data.
- The object detection algorithm is the Yolo-v8 algorithm.
- Our datasets are utilized to train the model to identify PPE and heavy equipment.
- Hazards are recognized according to the equipment or machinery in their immediate vicinity and the workers' state (i.e., whether they are wearing the proper PPE).
- Any infraction of safety regulations or potential hazards is identified and reported.

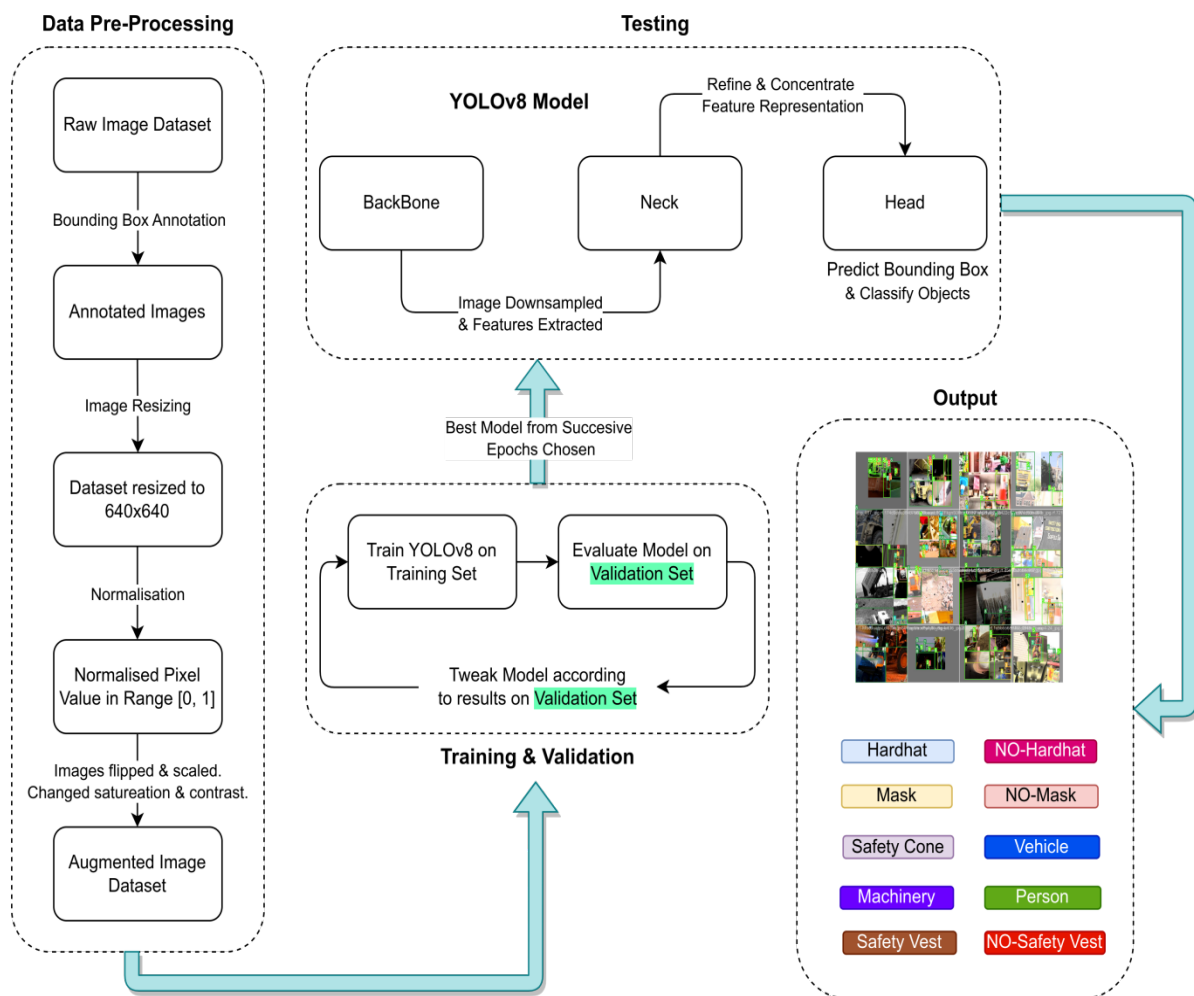


Figure 2: Proposed YOLOv8 framework for PPE detection.

The YOLOv8 model is not completely trained from scratch, rather we use a technique called Transfer learning. Where the pre-trained weights (parameters) are used which were acquired on the previous dataset (COCO in our case) as a foundation. The broad knowledge of object detecting features is encoded by these pre-trained weights. We add additional layers tailored to our own dataset to replace the pre-trained YOLOv8 model's final layers (head). These additional layers are trained using annotations for the objects that we are interested in.

YOLO is an object detector that works in a single step and is implemented as a supervised learning algorithm. By using CNN once rather than repeatedly, this one-stage method is shown in Figure 3. It accomplishes localization and classification concurrently—hence the name "one-stage". Cells are created on the image by forming an $S \times S$ grid. Each cell's localization and classification are handled by a CNN, which produces many potential bounding boxes and evaluates their degree of confidence. For the detection of multiple objects with their positions in a single image, the Non-Maximum Suppression technique eliminates all the bounding boxes except the one which has the highest confidence for every type of object.

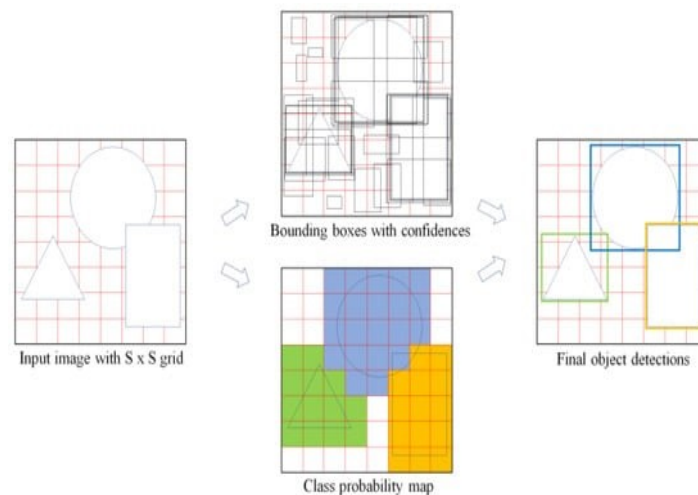


Figure 3: Diagrammatic representation of One-Stage Object Detection

Anchor-Free detection, a ground-breaking method that automatically predicts bounding boxes at an object's center, is incorporated into YOLOv8. Because preset anchor boxes are no longer required, the model is more durable and flexible enough to accommodate a wider range of item sizes and forms. It guarantees accurate detection results by improving object localization accuracy.

An important problem with previous YOLO models was that anchor boxes frequently showed the distribution of boxes in the benchmark dataset instead of the particular distribution of boxes in a custom dataset. A variety of circumstances exist where the workers' safety is at risk. Computer vision could be of great use to us in automating the monitoring of such actions, ensuring building site safety overall. Safety conditions, which are considered as follows:

- **Personal Protective Equipment (PPE):** In Figure 6, PPE, which is short for personal protective equipment, is equipment used to guard against illnesses and injuries sustained at work. These can be applied to mitigate a range of occupational risks. The several PPE components are depicted in the diagram below.
- **Fall Protection:** Utilizing protective equipment such as body harnesses and body netting to avert injuries sustained during a fall. CCTV systems can be used to keep an eye on dangerous areas and look for hazards
- **Body Posture:** Employee body posture is essential for preventing accidents, particularly in high-risk areas like scaffolding frameworks. An increased risk of events can result from improper posture and body position.
- **Trench Collapse:** Each year, a large number of people die as a result of trench collapse. Wearing safety gear when descending into large, deep ditches can help prevent mishaps.
- **Quality of Materials:** To protect both newly constructed structures and workers, all construction and transport tools and equipment must be well-made, in excellent condition, and free from flaws.

5. Experimentation And Result

The libraries used for this study are os, glob, pandas, numpy, matplotlib, seaborn, PIL, tqdm and Ultralytics. This experiment's primary goal is to detect things in real time and with great accuracy. Our datasets in this work were trained using the YOLO-v8 model. Table 2 shows the system needs.

Table 2
System Requirements

Component	Requirement
Processor	AMD Ryzen 7 7800X3D Processor 8 core 4.2GHz
RAM	32 GB
GPU	20GB AMD Radeon RX 7900 XT
Hard Disk	500 GB NVME SSD

A single trip through the complete training dataset for learning is known as epoch. A low number of epochs leads to underfitting since there is not enough learning. Excessive epoch counts result in longer training times and the possibility of overfitting to the training set, which deteriorates test outcomes using fresh, unused data. In order to accomplish adequate training and have the best test results, the right epoch value must be chosen. But until the dataset is applied, this value cannot be determined. Consequently, the YOLO instruction suggests a 300-epoch start. Reducing the number of epochs if overfitting is observed is advised; if not, increasing it to 600, 1200, and so on when calculating the proper epoch value by assessing the learning level. The learning rate (Lr) helps determine and adjust the ratio for improving the parameter values by using the degree of loss in every learning session. Given that the loss increases with each learning cycle, this value shouldn't be set too high to avoid divergence without learning. Too low of a value causes learning to proceed very slowly. As a result, when determining an appropriate Lr, the data quality should be considered. Depending on the optimizer being used, the recommended starting Lr in YOLO varies. The default optimizer in YOLO, which is Stochastic Gradient Descent (SGD) has the value set to 0.01. However, for Adam, AdamW, and RMSProp, it is 0.001. We used the following hyper parameters Table 3. to train our model.

Table 3
Hyper parameters used to train Yolov8

Variable	Description	Value
lr	Learning Rate	0.01
batch_size	Batch Size	16
optimizer	Optimizer	SGD

patience	Early Stopping	20
imgsz	Image Size	640
epochs	Epoch to train For	100

After training our model for 100 epochs, the result we got are as follows in Figure 4.

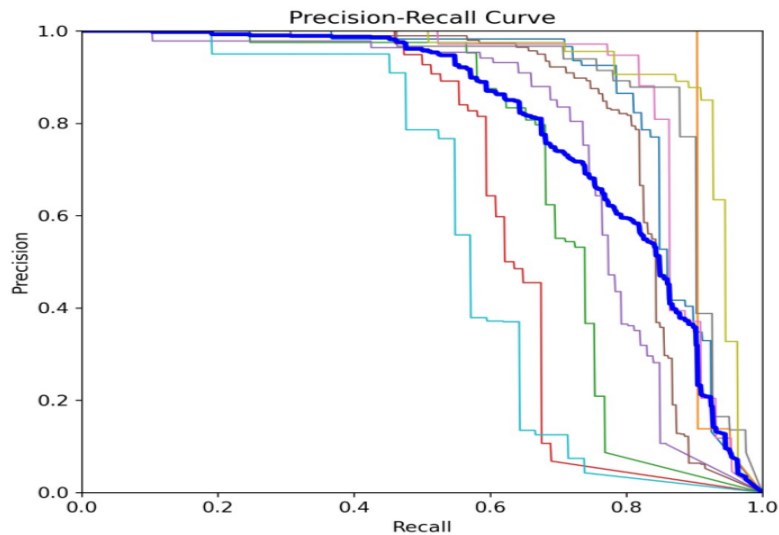


Figure 4: Precision-Recall Curve

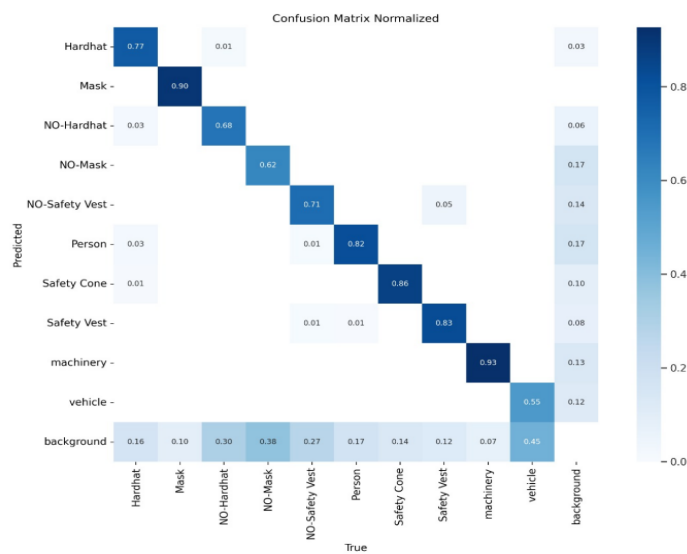


Figure 5: Normalized Confusion Matrix

The confusion matrix for the above implementation is shown in figure 5 where the predicted and true value is mapped onto the normalized matrix elements. Our training and validation loss can be seen in figure 6.

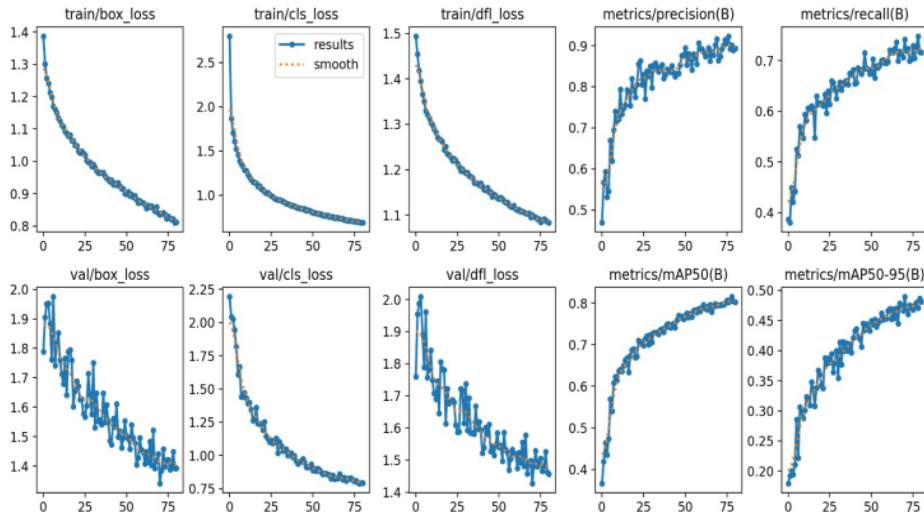


Figure 6: Training and Validation Loss Graphs.

6. Conclusion

Although the desire for progress frequently takes precedence, advancement still depends on having a workforce that is secure and well. In this study, we employed the YOLO model, version 8 (YOLO-v8), as an object detection model for our dataset to detect personnel, heavy machinery, and personal protective equipment (PPE). The model shows encouraging results in correctly and real-time detecting workers, personal protective equipment, and heavy machinery on construction sites. Partially visible items in the picture and video frames could be identified by the model. Applying the created dataset to YOLO v8 suggests a consistent procedure for raising prediction accuracy through adjustments to epochs, optimizers, and hyperparameters. The testing dataset yielded the following findings for assessment metrics: precision 0.891, recall 0.797, and mAP of 0.84.4, including safety vests in the worker/PPE dataset and loaders and cranes in the heavy equipment dataset. These indicate that the model performed well in identifying underrepresented classes. The final results are competitive enough when compared to test and validation accuracies offered by other research using computer vision in construction safety management. Future work will involve improving the suggested system's ability to detect small objects. Increasing the amount of data in the training datasets will be the primary enhancement. Furthermore, hazardous situations can be foreseen and avoided with the application of spatial-temporal analysis.

References

- [1] IEEE Access, vol. 9, pp. 166603-166616, 2021, doi: 10.1109/ACCESS.2021.3135662; M. Z. Shanti and associates, "A Novel Approach to AI-Powered Smart Construction Safety Inspection Method in the United Arab Emirates."
- [2] "Multi-class classification of construction hazards via cognitive states assessment using wearable EEG," Jeon, J., and Cai, H. (2022). Advanced Engineering Informatics, volume 53, page 101646, 2022. DOI: 10.1016
- [3] Alkaissy, M. et al., "Improving construction safety through machine learning-based injury type classification," Safety Science, vol. 162, pp. 106102, 2023. DOI: 10.1016
- [4] Development of a Classification Framework for Construction Personnel's Safety Behavior Based on Machine Learning, S. Yin, Y. Wu, Y. Shen, and S. Rowlinson, vol. 13, no. 1, p. 43, 2023. DOI: 10.3390

- [5] "Vision-based monitoring of site safety compliance based on worker re-identification and personal protective equipment classification," *Automation in Construction*, vol. 139, pp. 104312, 2022, by J. C. P. Cheng, P. K.-Y. Wong, H. Luo, M. Wang, and P. H. Leung. DOI: 10.1016
- [6] J. Lee and S. Lee, "Construction Site Safety Management: A Computer Vision and Deep Learning Approach," *Sensors*, vol. 23, no. 2, p. 944, 2023. DOI: 10.3390
- [7] H.T.T.L. Pham, M. Rafeizonooz, S. Han, and D.-E. Lee, "Positive and Prospective Developments of Deep Learning Applications for Safety Management xcvxcxxxxxxxxxxxxx Construction," *Sustainability*, vol. 13, no. 24, pp. 13579, 2021. DOI: 10.3390
- [8] Y. Zhao, Q. Chen, W. Cao, J. Yang, J. Xiong, and G. Gui, "Deep Learning for Risk Detection and Trajectory Tracking at Construction Sites," *IEEE Access*, vol. 7, pp. 30905-30912, 2019. DOI: 10.3390
- [9] W. Fang, P. E. D. Love, L. Ding, S. Xu, T. Kong, H. Li, "Computer Vision and Deep Learning to Manage Safety in Construction: Matching Images of Unsafe Behavior and Semantic Rules," In **IEEE Transactions on Engineering Management**, 70(12), Dec. 2023, pp. 4120–4132. DOI: 10.1109
- [10] Kim, K., Kim, S., and Jeong, K. "Application of YOLO v5 and v8 for Recognition of Safety Risk Factors at Construction Sites," *Sustainability*, vol. 15, no. 20, pp. 15179, October 2023.
- [11] M. Ferdous & S. M. M. Ahsan, "YOLO-based architecture to detect personal protective equipment (PPE) for construction sites: PPE detector," *PeerJ Computer Science*, vol. 8, pp. e999, 2022. DOI: 10.7717
- [12] Roboflow Universe Projects, "Construction Site Safety Dataset," Roboflow Universe, Aug. 2023.
- [13] H. Chen, L. Hou, X. Wang, and G. Zhang, "Deep Learning-Based Applications for Safety Management in the AEC Industry: A Review," *Access*, vol. 9, pp. 166603–166616, 2021, doi: 10.1109/ACCESS.2021.3135662.