REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

Real-Time PPE Monitoring And Demographic Analysis At Construction Sites Using Yolo

Yash Brahmbhatt[1*], Vashisth Zatakia[2], Nupur Upadhyay[3], Sweedle Machado[4]

[1*]Department of Information Technology, DJSCE, yashbrahmbhatt26@gmail.com
[2]Department of Information Technology, DJSCE, zatakiavashisth151@gmail.com
[3]Department of Information Technology, DJSCE, nupurupadhyay1292@gmail.com
[4]Department of Information Technology, DJSCE, sweedle.machado@djsce.ac.in

***Abstract*—**The construction industry poses significant challenges in maintaining worker safety and rights. Existing safety measures involve an on-site officials working on the ground to ensure that worker safety measures like wearing safety helmet and power monitoring are undertaken as mandated. However, these methods are often inefficient, as on-ground checking takes significantly more time and allows the construction site authority to plan their excuse. In response, this paper presents a comprehensive novel approach, an AI-enhanced construction site monitoring system that integrates real-time CCTV hazard detection, age verification, and personal protective equipment (PPE) compliance as robust measures to ensure worker safety. Leveraging a React Native application for worker-side interaction and officer-side oversight and integrating real-time location tracking and Firebase-based notifications, the system fosters proactive safety measures and promotes worker rights. Compared to existing solutions, the proposed system boasts superiority with an average accuracy of 92.5%, scalability with the use of Android-based applications and modifiable APIs and AI models, and user-centric features, leading to a safer and more transparent system that surpasses existing solutions.

***Index Terms*—**Construction Safety, Artificial Intelligence, Re- act Native, Real-time Monitoring, Worker Rights, Safety Helmet Detection, PPE Compliance, Hazard Alert, Grievance Reporting.

I. INTRODUCTION

Construction sites are important areas of economic development and face significant challenges in ensuring the safety and health of workers. In an industry where countless injuries and deaths occur each year, safety measures are essential. Traditional manual monitoring and video surveillance, while important, face limitations in providing real-time data and identifying hazards. Given the need to revolutionize construction site safety, this article focuses on the use of artificial intelligence (AI) in developing innovative construction site monitoring systems.

Previous statistics and reports have revealed the scale of the problem, indicating that a large number of accidents and injuries occur in the construction sector each year. There are still construction sites where monitoring procedures are manual and have difficulty coping with the dynamic and high-risk nature of the construction site. As developments progressed, video surveillance gave way to it, but video surveillance often relied on post-event surveillance, which prevented real-time prevention of incidents. These limitations have led researchers and experts to investigate technologies such as artificial intelligence and machine learning to ensure the safety of construction sites.

Previous literature study focuses on the automatic detection of safety helmet on construction sites through video surveillance. Automatic safety helmet detection pin-points need for automated detection of safety protocols [1][2]. Intelligent analysis technology for power monitoring video data expands the scope of safety measures beyond personnel to include environmental factors [3]. The utilization of computer vision applications in offsite construction illustrates the industry's growing recognition of technology-driven solutions [4]. In addition, real-time face and object detection for video surveillance applications showcases the potential of AI in enhancing security [5], coupled with multi-camera video scene graphs for surveillance video indexing and retrieval, laying the foundation for efficient data analysis [6].

The proposed methodology builds upon the insights gained from these studies and seeks to overcome the identified limitations by integrating AI into construction site monitoring practices. By employing computer vision algorithms, real time hazard detection, and proactive communication through a mobile application, this research aims to provide a holistic solution that not only ensures compliance with safety protocols but also actively prevents accidents. The introduction concludes by outlining the organization of the paper, which includes sections on the literature survey, proposed methodology, system implementation, evaluation, and future work. The ensuing sections will elaborate on how this AI enhanced construction site monitoring system addresses the shortcomings of existing approaches and paves the way for a safer and more efficient construction industry.

II. LITERATURE SURVEY

In the high-risk environment of construction sites, ensuring the safety of workers is a paramount concern. Despite advancements in safety protocols, construction sites are considered as one of the most hazardous and risky industries, where many fatalities and casualties occur. This literature survey gives the overview of how computer vision, machine learning and deep learning technologies has enhanced safety measures.

This paper [1] reflects how construction industry in USA suffers from the highest number of fatalities and casualties

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

among all industries. Given these significant losses, an automated study on use of helmet in construction site safety was developed. This paper has used Image processing and Machine learning/Deep learning techniques to analyze construction footage to determine whether workers wear safety helmets.
In this review paper [2], authors have presented a lightweight framework for detection of safety helmet. It uses YOLOv4 and YOLOv8 for object detection. The experimental results demonstrated that this lightweight model achieves an impressive detection speed of up to 114.26 frames per second (FPS) and a helmet detection accuracy of 90%. These performance metrics indicate that the model is well-suited for deployment in real-world construction environments, offering both high-speed processing and reliable accuracy.
With lot of technological advances in image processing and object detection, this paper [3] presents intelligent analysis method for power field surveillance video data by using CNN model. By combining frequency domain information, HOG algorithms, and color- based analysis, they enhanced worker safety [3].
Many applications of computer vision in on site construction monitoring have been investigated by researchers and have they have proved to provide valuable information in this field. The development of advanced imaging technology has made it possible to perform real-time face recognition and object detection in video surveillance to monitor public spaces, homes, etc. as a safe place. It provides practical applications in the field, such as planning, feature classification and removal, and improving overall image processing and analysis. Computer vision applications in outdoor construction provide a reference point to understand current applications, limitations, and anomaly detection [4].
Furthermore, as deep learning and CNN techniques are proved to achieve high accuracy in face and object detection, it can also detect as age and gender of the humans detected in the frame [5]. The system has applications in enhancing security, crime prevention, and beyond. The literature review comprehensively explores the safety domain at construction sites, emphasizing real-time face and object detection through computer vision and machine learning. Automatic detection of helmets is a central focus, employing various algorithms such as frequency domain information, HOG, and color-based analysis to enhance worker safety. The review also discusses a YOLOv8-based model for safety helmet detection, noting scalability and adaptability limitations [5].
In conclusion, the current techniques for ensuring safety on construction sites, even exhibit significant limitations. Existing literature review based on, YOLOv4, YOLOv8, CNN a, deep learning while effective in detecting safety helmets, often face challenges in scalability and adaptability to diverse construction environments. Furthermore, automatic detection systems for helmet use and other safety measures are often limited in application and often lack real-time analysis capabilities required on construction sites. Current intelligent surveillance technologies, such as those that use neural networks to leverage video data, offer significant improvements, but they still lack robustness, real-world hazard detection, and worker safety monitoring.
This research work brings a notable novelty to this field by integrating advanced AI-powered monitoring systems that not only detect safety helmets and other PPE with high accuracy but also incorporate real-time face and object detection, age and gender prediction, and hazard identification across multiple cameras.

PROPOSED METHODOLOGY

This research work aims to develop an AI-enhanced construction site monitoring system that includes several vital functionalities, mainly focusing on such as safety helmet, safety vest, PPE kit detection by continuously monitoring video surveillance camera deployed at various places of construction site. Additionally, this paper detects age and gender of the individuals at construction site and thus by identifying the gender of workers contributes to employee management and ensures the compliance with legal age. This system also sends real time location of the workers to the in charge official that helps to track the worker incase if he is stuck in the mishap situation as well ensures whether the worker is adhering to all the safety measures.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

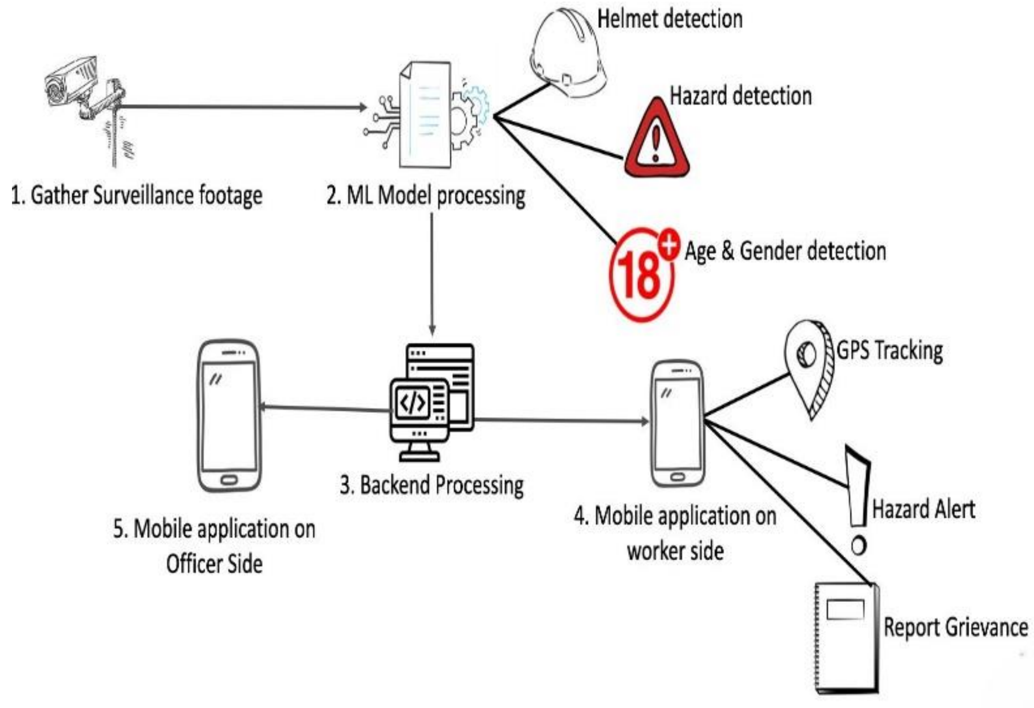


Fig. 1. Proposed System Architecture

The proposed system as shown in Fig. 1 consists of the following key components:

A. Data Acquisition:

The input to this system is real time video footage captured in through surveillance cameras. The cameras are positioned to cover all critical areas, such as high-risk zones and entry points at construction sites to ensure safety monitoring and remote oversite.

B. PPE (Personal Protective Equipment) Detection:

Personal protective equipment (PPE) is mandatory for construction site workers to reduce exposure to hazards. The types of PPE used at construction site include hard hats (helmets), safety glasses and high-visibility clothing such as vest or suits. This research work employs a modified version of the YOLO object detection model for PPE (Personal Protective Equipment) detection, specifically focusing on helmet and vest/suits identification [7][8]. From the video footage, frames are extracted, and objects and people are identified in each frame. Subsequently, it is verified whether the employee on site is wearing the appropriate PPE kit. YOLO's single neural network design allows for real-time face detection in a single pass, incorporating a backbone network based on a variant of EfficientNet [7]. The architecture includes a neck network using convolutional layers for feature fusion and a head network predicting bounding boxes and class probabilities, enhancing spatial resolution and semantic information.

C. Age and Gender Detection Model

After PPE detection, the system utilizes YOLO to identify the age and gender of the workers on the construction site. YOLO first detects faces within the frames, and then a separate model processes these detected faces to classify the age and gender of each worker. The age and gender detection model as shown in Fig. 2 consists of four convolutional layers with 32 filters and a two-size kernel. Max-pooling layers are employed for dimensional reduction, with ReLU activation functions addressing the vanishing gradient problem after each convolutional layer. The model incorporates five dense layers with 128 neurons each, incorporating dropout layers for regularization. The output layer utilizes the softmax activation function for multiclass classification employing categorical cross-entropy loss for assessing similarity between predicted classes probabilities and true class labels. The Adam optimizer is applied to train the age and gender detection models [9]. This approach allows for efficient real-time demographic analysis like age and gender analysis along with PPE monitoring.

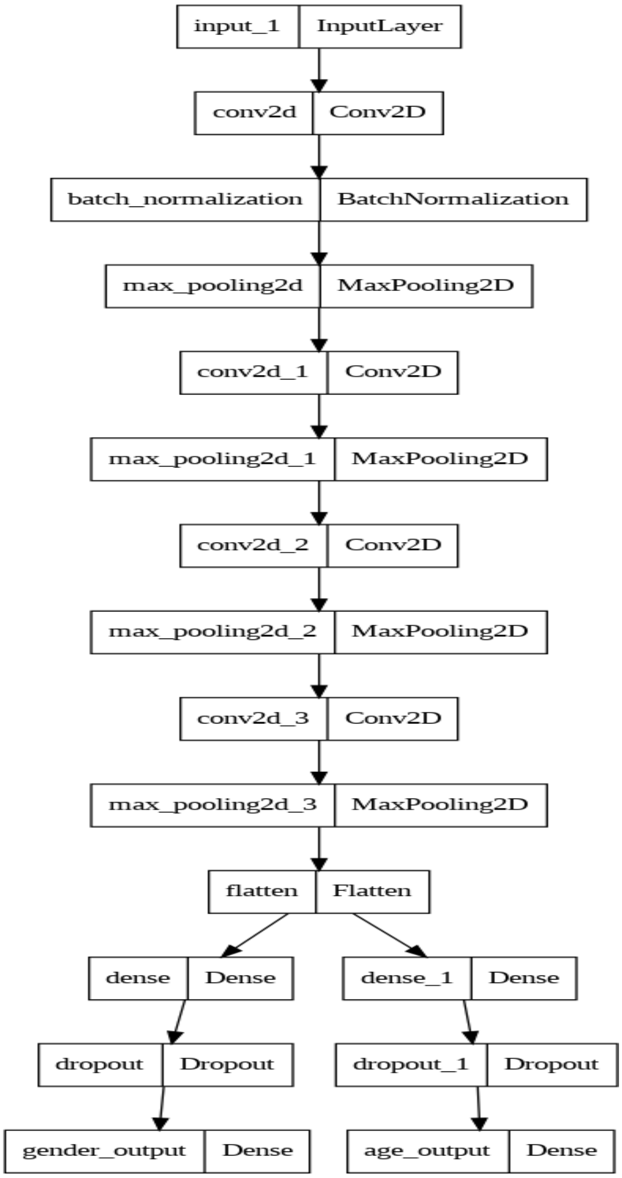


Fig. 2. Age and Gender Convolutional Neural Network Model

D. Real-Time Alerts and GPS Tracking:

The proposed system sends real time alerts to high authority or the in charge official via SMS when a worker fails to adhere to safety measures, such as not wearing safety helmet or safety vests, or when workers below legal age is spotted on construction site. Moreover, the proposed system allows construction site workers to enable their GPS, facilitating their location tracking in case of any accidents and mishaps at the site. Grievance reporting and redressal mechanisms are available for workers and accessible to supervising officers.

In conclusion, this research presents a complete solution that includes real-time safety helmet and safety vest detection, PPE compliance, and age and gender identification of workers, all integrated into a mobile application for efficient monitoring and response. The system implemented a modified YOLO model for object detection and age and gender prediction, with real-time alerts and GPS tracking for workers. Additionally, this application also provides grievance reporting and redressal mechanisms for both workers and supervising officers.

III. EXPERIMENTS AND RESULTS

This section gives a brief description of about the datasets on which the system is trained and tested along with the walkthrough of GUI of the system. Section A. describes about the datasets on which the system is tested. Section B reflects the implementation details.

A. Dataset Used:

For training and validating the proposed system, experiments were conducted on the following datasets. Below is a brief

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

summarization of each data set's characteristics:
1. **UTKFace dataset [10]**: The UTK (UTKFace) dataset is a facial image dataset that includes face images with annotated age and gender. This dataset includes a total of 25 thousand images, each labeled with the subject's age and gender.
2. **Custom Object Detection Dataset:** This dataset was developed specifically for this research work consisting roughly 6,000 photos which was used for training, and 100 images were used for validation. The images in the dataset were manually annotated for better model performance. The object classes used in the dataset include:

- Hardhat: Detects the presence or absence of a hardhat on a construction worker's head.
- Mask: Detects the presence or absence of a mask on a worker's face.
- NO-Hardhat: Identifies workers who are not wearing a hardhat.
- NO-Mask: Identifies workers who are not wearing a mask.
- Safety Vest: Detects the presence or absence of a high- visibility safety vest.
- NO-Safety Vest: Identifies workers who are not wearing a safety vest.

This dataset provides valuable information for monitoring PPE compliance, identifying potential safety hazards, and optimizing worksite safety protocols [9]. By understanding the specific detection classes and their significance, you can gain deeper insights into the data and develop effective interventions for improving worker safety on construction sites.

A. Implementation Details:

The proposed framework is implemented with the wide range of libraries available through Open-cv, Media-Pipe, React Native, Tensor Flow, React Native and Google Colab. The snapshots of core modules of proposed framework are explained below. It is designed to cater to two types of users: workers and officers. In Officer's section there are option to view worker's profile, a Grievance section which contains all the problems addressed by the workers, Alert Section which is about any workers raised alerts or the alerts regarding not wearing helmets or under age detection as shown in Fig. 3(a) and (b). Furthermore, Worker section contains an options like viewing their own profile, report any grievance they have to the officer, to raise an alert as shown in Fig. 4(a) and 4(b).

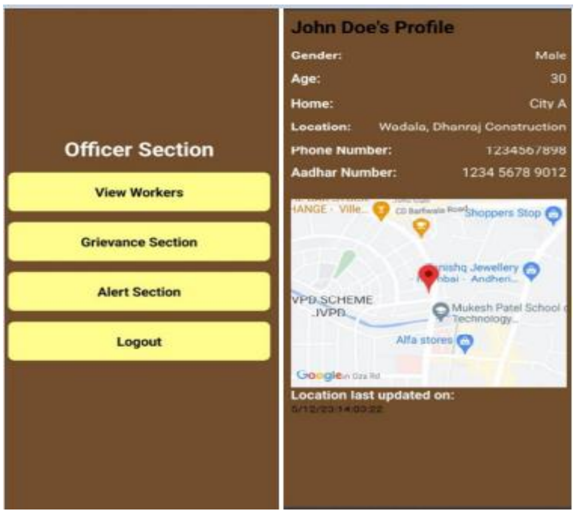


Fig. 3(a). Officer Section Page Fig. 3(b). Worker Profile Page

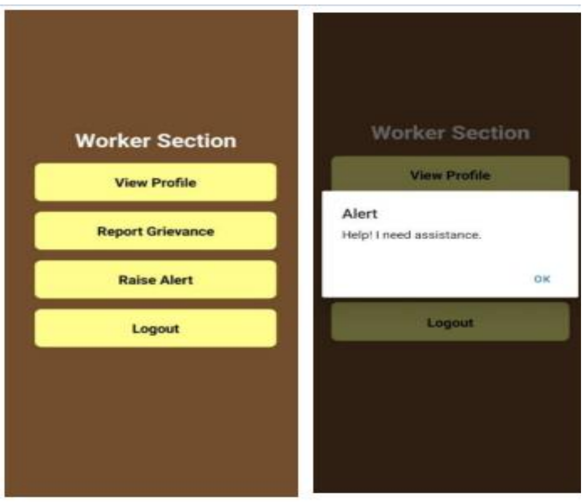


Fig. 4(a). Worker Section Page Fig. 4(b). Alert Page

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

Fig. 5, Fig. 6, Fig.7 shows PPE Output representing model scores of various detected parameters, Legal Age Detection using Age and Gender Model, Under Age Detection using Age and Gender Model as shown below.

Fig. 5. PPE Output representing model scores of various detected parameters

Fig. 6. Legal Age Detection using Age and Gender Model

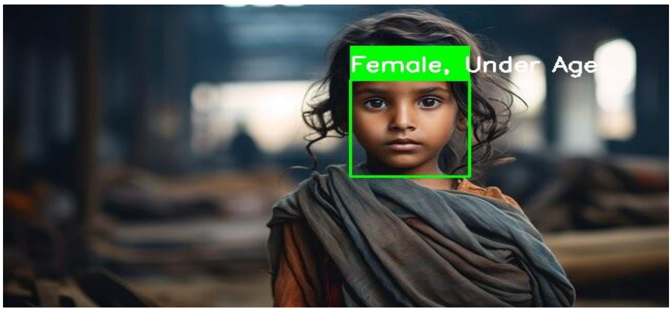
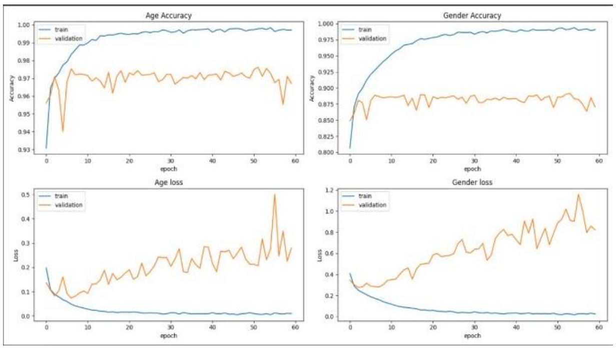


Fig. 7. Under Age Detection using Age and Gender Model

V. RESULTS

The CNN model was designed to predict the age and gender of workers. The age prediction model achieved an average validation accuracy of 97%, with a minimum accuracy of 94% as shown in Fig. 8. For this model, the average loss was recorded at 0.22. In contrast, the gender prediction model attained a peak accuracy of 90% and an average accuracy of 88%, with an average loss of 0.4 as shown in Fig. 8. These metrics were computed based on a 75-25 train-test split ratio. The dataset utilized is the UTKFace Dataset, which originally contains over 25,000 images. After augmentation, the dataset size expanded to approximately 100,000 images.



REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

Fig. 8. Age and Gender Model Accuracy and Loss

The "epoch" column signifies the training duration, with accompanying columns such as "train/box_loss," "train/cls _loss," and "train/dfl_loss" detailing the respective losses for bounding boxes, classification, and deformable convolutional networks during the training phase. Performance metrics like precision, recall, mean average precision at 50%(mAP50), and mAP from 50% to 95% (mAP50-95) for bounding boxes are also included under "metrics/precision(B)", "metrics/recall(B)", "metrics/mAP50(B)" and "metrics/mAP50-95(B)". The validation losses are similarly recorded in columns like "val/box_loss", "val/cls_loss" and "val/dfl_loss". Learning rates for distinct parameter groups are specified under "lr/pg0" "lr/pg1" and "lr/pg2" as shown in Fig. 9, Fig. 10, Fig. 11 and Fig. 12 .

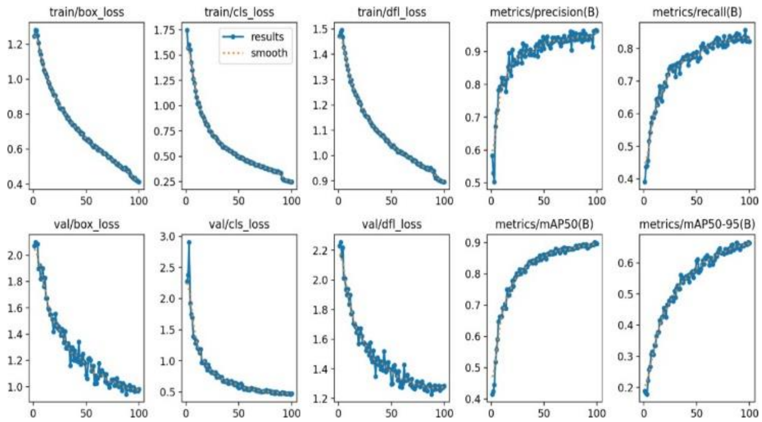


Fig. 9. Graphical Representation of YoloV8x Performance Metrics

Fig. 10. YOLOv8x Performance Metrics for every 10 epochs

Epoch	Train Box Loss	Train Cls Loss	Train DFL Loss	Precision	Recall	mAP50	mAP50-95	Val Box Loss	Val Cls Loss	Val DFL Loss	LR/pg0	LR/pg1	LR/pg2
0-9	1.2446	1.7485	1.4718	0.58345	0.39075	0.41533	0.18834	2.0716	2.2757	2.2303	0.00023654	0.00023654	0.00023654
10-19	1.0384	1.0203	1.2909	0.80563	0.60632	0.69009	0.33454	1.8281	1.2068	1.8987	0.00065038	0.00065038	0.00065038
20-29	0.90524	0.79912	1.203	0.85872	0.68381	0.76081	0.42588	1.5527	0.93644	1.6728	0.0005797	0.0005797	0.0005797
30-39	0.79287	0.64429	1.1259	0.89568	0.7493	0.81217	0.51877	1.2934	0.75283	1.4684	0.00050901	0.00050901	0.00050901
40-49	0.71514	0.55842	1.0783	0.88826	0.78054	0.83875	0.55147	1.2383	0.64676	1.4471	0.00043832	0.00043832	0.00043832
50-59	0.64912	0.48611	1.0354	0.93964	0.76107	0.85878	0.61097	1.062	0.57401	1.2991	0.00036764	0.00036764	0.00036764
60-69	0.59616	0.43654	1.0071	0.94003	0.80918	0.88066	0.61717	1.0653	0.54943	1.3019	0.00029695	0.00029695	0.00029695
70-79	0.55265	0.40023	0.97948	0.93202	0.82904	0.88095	0.64142	1.0233	0.5027	1.2752	0.00022627	0.00022627	0.00022627
80-89	0.50595	0.36532	0.95313	0.93951	0.83852	0.8869	0.64653	0.99931	0.48802	1.2612	0.00015558	0.00015558	0.00015558
90-99	0.46604	0.28479	0.92292	0.93618	0.83043	0.88222	0.64336	1.0109	0.4918	1.2881	0.000084895	0.000084895	0.00008489

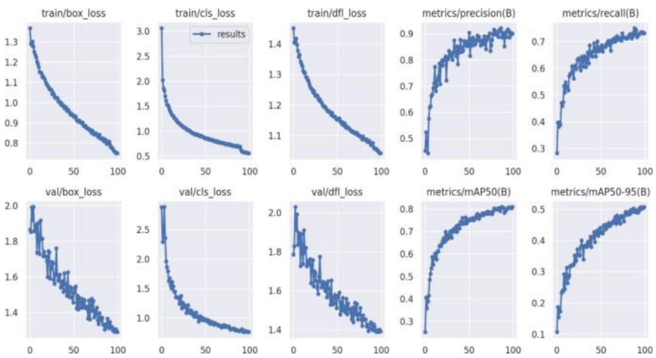


Fig. 11. Graphical Representation of YoloV8n Performance Metrics

Epoch	Train Box Loss	Train Cls Loss	Train DFL Loss	Precision	Recall	mAP50	mAP50-95	Val Box Loss	Val Cls Loss	Val DFL Loss	LR/pg0	LR/pg1	LR/pg2
0-9	1.2922	2.0789	1.3946	0.5701	0.392	0.4054	0.1869	1.8527	2.2875	1.8273	0.0702	0.0033	0.0033
10-19	1.1916	1.4223	1.3339	0.711	0.548	0.5856	0.2842	1.802	1.5603	1.8099	0.0091	0.0091	0.0091
20-29	1.043	1.0752	1.2302	0.807	0.631	0.6842	0.3522	1.683	1.2334	1.6888	0.0076	0.0076	0.0076
30-39	0.9906	0.9639	1.1951	0.8143	0.659	0.7224	0.3955	1.5416	1.0619	1.6169	0.0066	0.0066	0.0066
40-49	0.9361	0.8632	1.1531	0.8632	0.676	0.7516	0.4407	1.4565	0.9607	1.5105	0.0052	0.0052	0.0052
50-59	0.8745	0.7763	1.1137	0.89	0.705	0.778	0.4744	1.4662	0.8636	1.5125	0.0035	0.0035	0.0035
60-69	0.8468	0.7396	1.0985	0.8559	0.716	0.7861	0.4755	1.3907	0.8505	1.4762	0.0027	0.0027	0.0027
70-79	0.8172	0.7057	1.0833	0.9211	0.72	0.8062	0.4976	1.3341	0.7937	1.4058	0.0016	0.0016	0.0016
80-89	0.811	0.7097	1.0876	0.8967	0.726	0.8032	0.5018	1.3074	0.7668	1.3994	0.0005	0.0005	0.0005
90-99	0.7495	0.559	1.0426	0.8999	0.731	0.8088	0.5071	1.291	0.7616	1.3917	0.0003	0.0003	0.0003

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No.2 (2024)
http://www.veterinaria.org
Article Received: Revised: Published:

Fig. 12. YOLOv8n Performance Metrics for every 10 epochs

Confusion Matrix is a type of visualization used in machine learning to understand the performance of a classification model. The matrix shown in Fig.13. includes various labels such as “Hardhat”, “Mask”, “NO-Hardhat”, “NO-Mask”, “NO Safety Vest”, “Per- son”, “Safety Cone”, “Safety Vest”, “machinery”, “vehicle”, and “background”. Each cell in the matrix represents the inter- section of predicted and true labels with a specific value. The values are normalized and color-coded, with darker shades of blue indicating higher values. This matrix helps in identifying the classes that the model is getting confused with and can guide further improvement of the model.

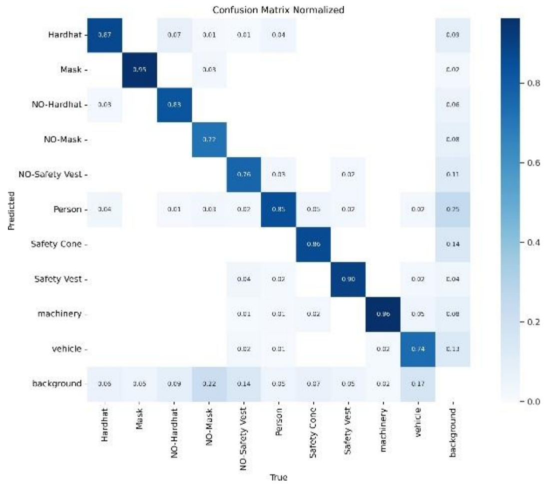


Fig. 13. Confusion Matrix for analyzing performance of PPE Detection Model

IV.CONCLUSIONS AND FUTURE WORK

The AI-enhanced construction site monitoring system presented in this paper represents a significant leap forward in mitigating safety challenges within dynamic construction environments. The integration of a sophisticated AI model, leveraging YOLOv8 for real-time hazard detection, along with a user-friendly React Native mobile application and a robust communication system, forms a holistic approach to enhance worker safety and overall site management. Implementation and testing have proven that the system’s accuracy in detecting vulnerabilities and its rapid response to alerts and complaints demonstrate its superiority over existing solutions. Looking ahead, future work could focus on expanding the system’s functionality. Environmental monitoring, including factors such as temperature and noise, could create a safer environment. Increasing the adaptability of AI models to a wide range of construction and equipment applications would improve their performance in many areas. Research into automated hazard prevention methods and the inclusion of aspects related to worker health and well-being are promising possibilities for future developments. Continuous training and improvement of the AI model provides effective and adaptable solutions to create a safe environment, while remaining up-to-date over time. In summary, the system not only addresses existing safety issues, but also lays the foundations for further progress by promoting a culture of safety, health and construction work.

REFERENCES

[1] A. H. M. Rubaiyat et al., “Automatic Detection of Helmet Usage for Construction Safety”, in Proceedings of the 2016 IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW), Omaha, NE, USA, 2016, pp. 135-142. DOI: 10.1109/WIW.2016.045.

[2] D. Ren, T. Sun, C. Yu, and C. Zhou, “Safety Helmet Detection for Construction Site: A Research Study”, in Proceedings of the 2021 International Conference on Computer Information Science and Arti- ficial Intelligence (CISAI), Kunming, China, 2021, pp. 186-189. DOI: 10.1109/CISAI54367.2021.00042.

[3] W. Xiaoyan, C. Pei, L. Rui, Y. Bo, and N. Wenzhao, “Intelligent Analysis Technology of Power Monitoring Video Data Based on Convolutional Neural Network: A Research Perspective”, in Proceedings of the 2020 5th International Conference on Mechanical, Control, and Computer Engineering (ICMCCE), Harbin, China, 2020, pp. 2461-2464. DOI: 10.1109/ICMCCE51767.2020.00531.

[4] F. Alsakka, S. Assaf, I. El-Chami, and M. Al-Hussein, “Computer Vision Applications in Offsite Construction”, Automation in Construction, vol. 154, p. 104803, 2023. DOI: 10.1016/j.autcon.2023.104980.

[5] Mrs. Jain, Sakshi Chaturvedi, Hetvi Patel, and Rajeshwari Jaiswal, “Real-time Face and Object Detection with Age and Gender Prediction for Video Surveillance Applications”, International Journal for Research in Applied Science and Engineering Technology, vol. 11, pp. 2358-2363, 2023. DOI: 10.22214/ijraset.2023.49987.

[6] T. Patel, A. Y. H. Yao, Y. Qiang, W. T. Ooi, and R. Zimmermann, “Multi- Camera Video Scene Graphs for Surveillance Videos Indexing and Retrieval”, in Proceedings of the 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 2021, pp. 2383- 2387. DOI: 10.1109/ICIP42928.2021.9506713.

[7] Ultralytics, "YOLO: You Only Look One-level", [Online]. Available: https://www.ultralytics.com/yolo-vision/. [Accessed: Oct. 20, 2024].
[8] Hinepo, "YOLOv8 Finetuning for PPE Detection", Kaggle, [Online]. Available: https://www.kaggle.com/code/hinepo/ yolov8-finetuning-for-ppe-detection/. [Accessed: Sep. 29, 2023].
[9] N. Yan, H. Cheng, M. Wang, Q. Huang and F. Chen, "DP-Face: Privacy- Preserving Face Recognition Using Siamese Network", 2021 20th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Nanning, China, 2021, pp. 92-95, doi: 10.1109/DCABES52998.2021.00030.
[10] Susanqq, "UTKFace Dataset", [Online]. Available: https://susanqq. github.io/UTKFace/. [Accessed: Sep. 29, 2023].