PREDICTIVE SAFETY MONITORING FOR LIFTING OPERATIONS WITH VISION-BASED CRANE-WORKER CONFLICT PREDICTION

Peter Kok-Yiu Wong, Synge C.P. Lam, Isabel Y.N. Lee, Felix C.L. Ting & Jack C.P. Cheng
Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, China

Pak Him Leung
AutoSafe Limited, Hong Kong, China

ABSTRACT: *Construction industry has reported among the highest accident and fatality rates over the past decade. In particular, crane lifting is a notably hazardous operation on construction sites, causing fatal accidents like workers being struck by the boom or objects fallen from tower cranes. Manual monitoring by on-site safety officers is labour-intensive and error-prone, while incorporating computer vision techniques into surveillance cameras would enable more automatic and continuous monitoring of construction site operations. However, existing studies for lifting safety mainly detect the presence of individual objects (e.g. workers, crane components), while a methodology is needed to predict their potential collision more proactively before accidents happen. This paper develops a vision-based framework for predictive lifting safety monitoring, including three modules: (1) object detection and classification: targeting at hook and lifting materials to enable danger zone estimation, along with workers and their personal protective equipment; (2) worker movement tracking and prediction: analyzing the historical moving trajectory of each unique worker to foresee his/her future movement in certain period ahead; (3) multi-level safety assessment: issuing predictive warning in real-time upon any crane-worker conflict foreseen. The proposed framework is applicable to real-time site video processing and enables end-to-end lifting safety monitoring with instant alerting upon unsafe scenarios observed. Importantly, the proposed framework predicts the future movement of workers to proactively identify potential site hazard, in order to trigger earlier safety alert for more timely decision-making. With a large video dataset capturing tower crane operations, the proposed framework demonstrates competitive accuracy and computational efficiency in crane-worker conflict prediction, validating its practicality for real-time lifting safety monitoring.*

KEYWORDS: *Computer Vision; Construction Safety Monitoring; Crane-Worker Conflict Prediction; Deep Learning; Predictive Safety Assessment; Trajectory Tracking.*

1. INTRODUCTION

The construction industry has been plagued for long by a high frequency of accidents and fatalities. According to statistics from the Hong Kong Labour Department (2018), the industry accounted for 76% of occupational fatalities in 2017, making it the most dangerous sector in Hong Kong. Similarly, the U.S. Bureau of Labour Statistics (2018) reported an average rate of 2.6 deaths per day, resulting in 949 deaths for the year. With reference to an overview of the Hong Kong construction industry (Shafique & Rafiq, 2019), there were on average 3597 occupational injuries and 20 occupational fatalities per year between 2011 and 2017. The U.S.A Department of Labour (2022) indicated that the estimated cost of employers' direct compensation to construction accidents is up to US$1 billion per week. These statistics suggest the urgent need for improved construction safety measures to protect the lives of workers and mitigate the financial burden that accidents impose on employers and the economy.

To address this critical issue, governments have established safety guidelines and regulations to standardize the industrial practices of construction safety monitoring. Lifting operations using tower cranes are a crucial aspect of construction work that requires particular attention, as they involve dynamic interactions between workers and machines. Traditionally, safety monitoring relies heavily on manual inspection by on-site safety officers. However, this method is prone to errors due to human fatigue, which can result in overlooked incidents. In recent years, advancements in artificial intelligence have led to the development of computer vision (CV) methods that can automate construction safety monitoring. These methods enable real-time object identification, improving the accuracy and efficiency of safety monitoring. However, there are two research gaps: (1) Previous approaches have focused on analyzing individual objects, such as workers and machines, separately, without a more comprehensive framework that considers their spatial interaction in real-time. (2) Previous studies have primarily focused on analyzing current scenarios/activities on sites, while a more predictive mechanism is needed to proactively identify and prevent potential accidents ahead of time. Therefore, this study develops a predictive safety assessment framework that monitors potential crane-worker conflicts and enable proactive incident prevention, ultimately reducing the number of accidents and fatalities in the construction industry.

Referee List (DOI: 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Peter Kok-Yiu Wong, Synge C. P. Lam, Isabel Y. N. Lee, Felix C. L. Ting, Jack C. P. Cheng, Pak Him Leung, *Predictive Safety Monitoring for Lifting Operations with Vision-Based Crane-Worker Conflict Prediction,* pp. 648-656, © 2023 Author(s), CC BY NC 4.0, DOI 10.36253/979-12-215-0289-3.64

2. RELATED WORK

Collision between workers and construction equipment happens regularly in complex and distracting construction environment that is overcrowded with workers. Close contact between construction machines and workers are one of the major causes of collision event that lead to injuries and deaths. Sensors such as GPS and RFID have been explored in prior studies. With the help of sensors, real-time spatial-temporal information could be provided for proximity measurements. As a result, a spatial-temporal relationship can be detected, and an early warning can be sent out to prevent the accident from happening (Liu et al., 2021). However, to obtain enough information to safeguard the construction site, numerous sensors have to be installed. The heavy financial burden will be caused by purchasing and hiring a professional individual to install and maintain the sensors (Zhang et al., 2020).

CV-based object tracking is a superior alternative to sensors since it lowers the cost and requires fewer resources to set up, therefore, more appealing to the industry. Previous research has trained YOLOv3 deep learning model for 2D positioning various construction site entities on 2D images captured from Aerial vehicles. Several studies developed convolutional neural networks to detect personal protective equipment (PPE), such as helmet and reflective vest (Cheng et al., 2022, Fang et al. 2018, Nath et al. 2020). Besides object detection, several studies developed human tracking algorithms to analyze behavior of each person more continuously (Kim et al., 2019, Wong et al., 2021). Other studies attempted to predict the future action of construction machines like excavators, based on their historical motion patterns (Luo et al., 2021), and also semantic segmentation that fine-grains the detected objects at pixel level to allow better positioning (Jeelani et al., 2021).

While previous studies can perform real-time object detection and tracking, a more comprehensive framework beyond developing those algorithms is needed for practical construction safety monitoring. An automatic safety evaluation system shall be established to enable effective intervention mechanisms to prevent the accident from happening. Previous studies have proposed some distance-based hazard evaluation criteria (Son et al., 2019). Some researchers have taken the velocity of construction equipment and workers into consideration as there is an association between larger velocity and collision accidents (Golovina et al., 2016). A previous study has attempted to determine the dynamics direct fall zone of a crane load using a mounted tower crane camera with computer vision (Chian et al., 2022). These studies enhance construction safety monitoring with the ability to predict the direct fall zone, where workers can be proactively prevented from entering danger zones.

3. PROPOSED METHODOLOGY

To facilitate tower crane safety monitoring, a vision-based framework is developed which comprehensively supports end-to-end CCTV analytics for real-time safety assessment. The overall procedure and information flow are summarised in **Figure 1**, with three major functional modules: (1) **object detection and classification**: interested objects in each video frame are detected and classified into three categories (i.e. workers, hook and lifting materials); (2) **worker movement tracking and prediction**: analyzing the historical moving trajectory of each unique worker and predict his/her possible location in certain period ahead; (3) **multi-level safety assessment**: issuing predictive warning in real-time upon any unsafe crane-worker conflict observed.

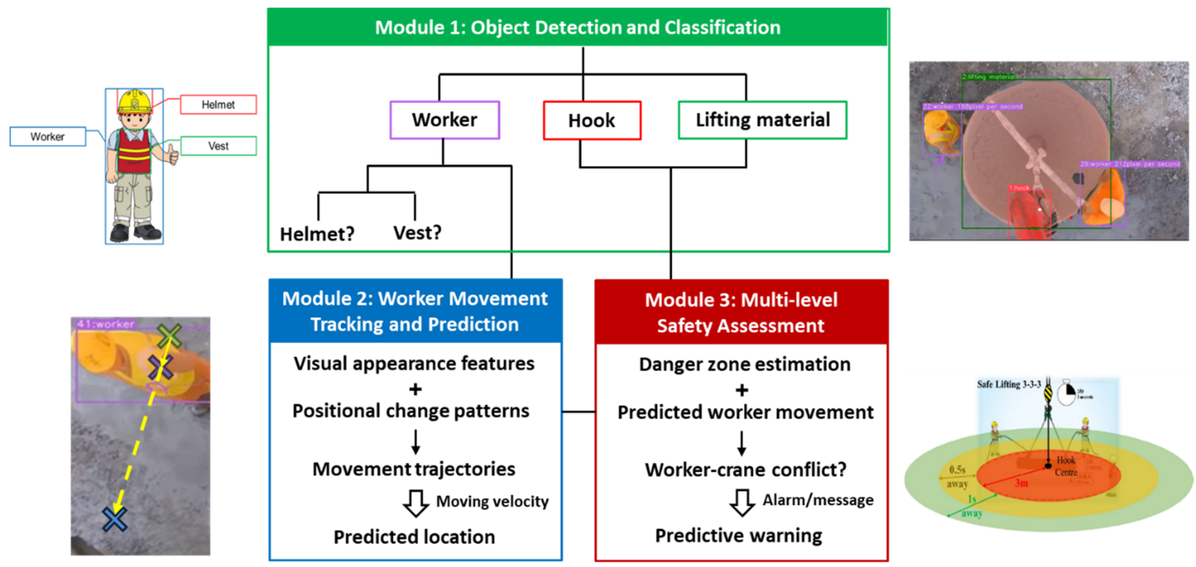

Figure 1: Overall information flow of the proposed crane safety monitoring framework

3.1 Object detection and classification

Upon receiving the raw videos from multiple cameras, the objects of interest are identified in Step 1. This is a crucial step that demands an automated process and accurately bounded objects (i.e. cropping the portion of image around each object with minimal background clutter). There have been numerous studies specialized in construction object detection (Fang et al., 2018; Luo et al., 2019; Memarzadeh et al., 2013), as well as comprehensive surveys of various state-of-the-art object detection methods (Brunetti et al., 2018; Huang et al., 2017). Hence, this paper adopts a competitive detection model for the object detection step. In particular, the YOLOv8 algorithm is used in view of its detection accuracy and inference efficiency revealed in recent studies.

More specifically, three types of construction objects are targeted, i.e. construction workers, crane hook and lifting materials during crane operations. With videos collected from construction sites, each object-of-interest is detected and cropped as a rectangular bounding box. Subsequently, a classification module outputs the corresponding class index associated with each bounding box. **Figure 2** illustrates a sample output of detection and classification (worker bounded by a purple box, hook by red and lifting material by green).

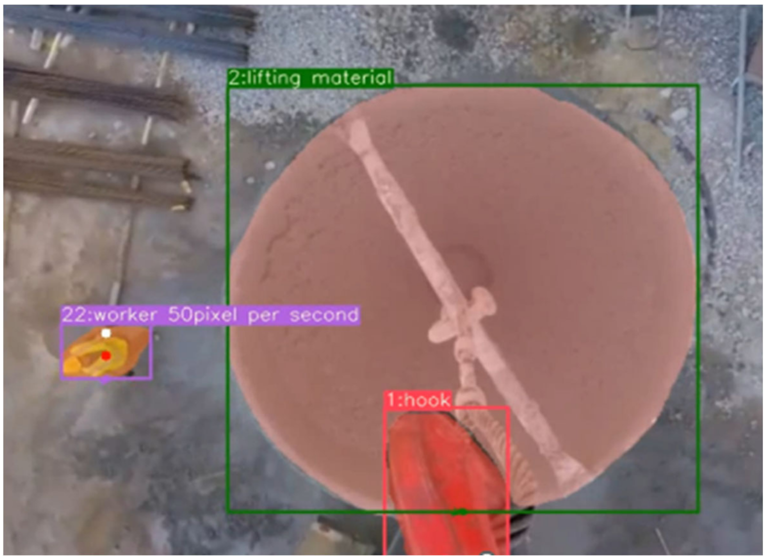

Figure 2: Illustration of object detection and classification results

For those detected boxes labeled as workers, a more fine-grained classification regime is defined to further analyze whether each worker wears necessary PPE, i.e. helmet and vest. As illustrated in **Figure 3**, two sub-categories are output by the classification model to determine the presence of helmet and vest respectively in their corresponding part of a body. To make the methodology more practical, the model is trained with both confirming and disconfirming classes, e.g. the head part is marked even no helmet exists around there. This approach renders the PPE inspection more accurate, because it avoids improper behavior, e.g. hand-carrying a helmet without properly wearing on the head. In that case, our method can correctly report that PPE is not properly worn, while ordinary detection method only identifies the presence of PPE in hand, which indeed violates PPE compliance.

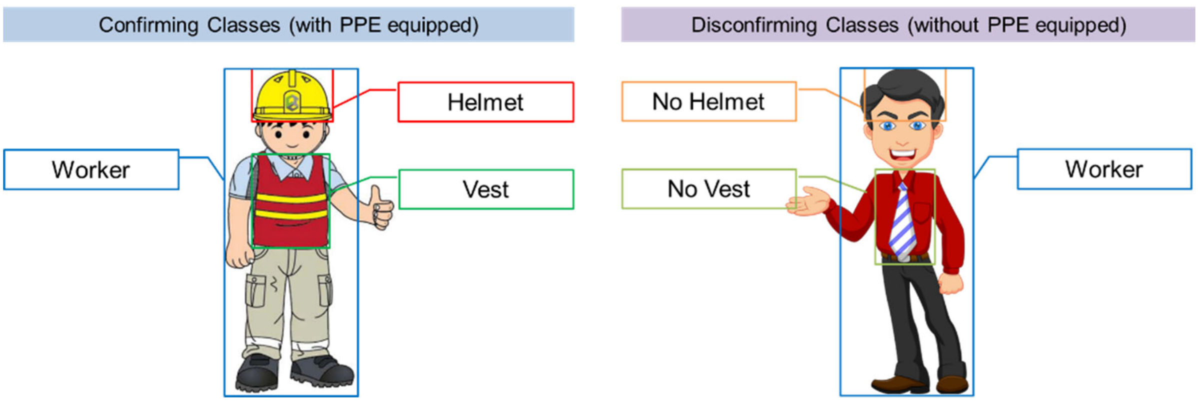

Figure 3: Definition of PPE statuses of a worker

3.2 Worker movement tracking and prediction

On top of the object detection results, worker trajectory tracking is performed to support worker behavioral analysis based on movement pattern. In this study, the method DeepSORT (Wojke et al., 2017) is utilized as a baseline to perform worker trajectory tracking over video frames, acquiring a complete trajectory of individual construction worker. The set of bounding boxes classified as 'worker' in Step 1 are further processed by DeepSORT, which subsequently analyzes the appearance features extracted from each worker and the positional change of the bounding boxes, in order to map unique identities to individual worker. **Figure 4** illustrates the assignment of unique identities to individual workers (22 to the left-sided worker, 29 to the right one).

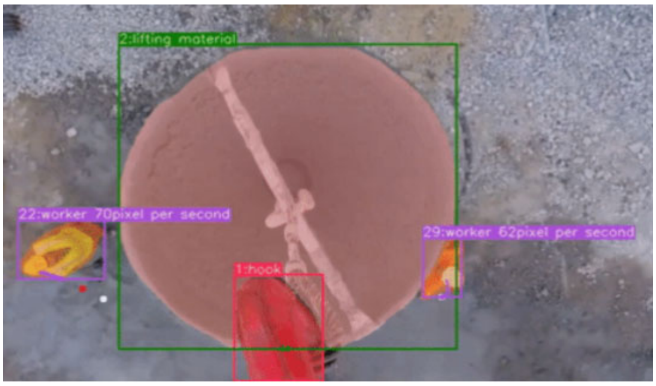

Figure 4: Illustration of worker tracking with unique identity assigned to different workers

Based on the trajectories of individual construction workers, their potential movement is then predict to foreseen whether their moving trajectories will potentially coincide with any lifting zone of the tower cranes nearby. This will allow dispatching warning signals in a more timely manner before workers actually enter the lifting zones. The prediction of future movement of each worker is defined in **Equation (1)**, which computes the image coordinates of the predicted worker location *t* timesteps later based on his/her observed velocity *v*.

$$d_2 = d_1 + v \times t \tag{1}$$

where,
d_1 = coordinates of the current location,
d_2 = coordinates of the predicted location,
v = velocity along corresponding direction,
t = time (measured by number of frames).

3.3 Multi-level safety assessment

By combining the output from object detection and worker movement tracking modules, spatial relationship between construction workers and lifting equipment is established. Regarding tower crane operations, the "Safe Lifting 3-3-3" Principle published by Hong Kong government (2020) is an industry standard in lifting operations. As illustrated in **Figure 5**, the 3-3-3 Principle states that workers should keep themselves 3m away from the lifting materials to ensure their safety. Yet, the 3-3-3 Principle only defines a single level of safety distance to be maintained from the lifting zone, while different degree of proximity may imply various levels of safety. Moreover, the standard only considers static behavior of workers (i.e. current location), while a more predictive safety monitoring regime is needed to consider the possible movement of each worker in certain period ahead.

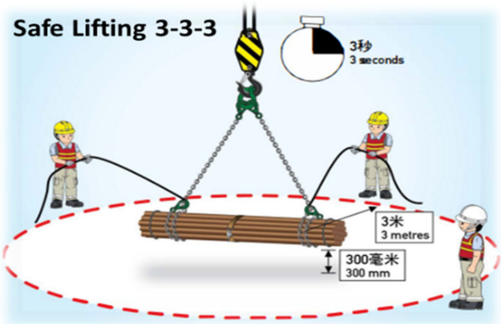

Figure 5: Definition of Safe Lifting 3-3-3 Principles

On top of the 3-3-3 Principle, the 3m operating region is defined as the danger zone around the lifting material detected by the object detection module. With the bounding box of lifting material generated, a danger zone of radius 3m around the bounding box centre is estimated. Different warning signals are then sent according to the corresponding risk scenarios and the predicted trajectory of the workers. As summarized in **Table 1**, different scenarios implies corresponding level of response. If a worker is already inside a fatal zone, 'Action' is issued to urge for immediate handling. If he/she is predicted to enter a fatal zone in 3 or 5 seconds, the responses become 'Alarm' and 'Alert' respectively which are less severe. Such a multi-level warning mechanism enables more flexible and predictive safety assessment.

Table 1: Proposed three-level mechanism of lifting safety assessment

Response	Scenario
Action	Worker inside fatal zone
Alarm	Worker will enter fatal zone in 3 seconds
Alert	Worker will enter fatal zone in 5 seconds

To alert both the workers and the residential site safety personnel to the potential safety hazards, an instant warning system is developed with a series of if-else loops, and connected with an external chatbot API. When the model detects the worker's tendency to enter the defined lifting zone, warning messages are issued to inform safety officers of the incident detected. The corresponding frames is also captured, with descriptive text about the unsafe scenario, and sent to registered stakeholders via an instant messaging platform (e.g. Telegram) for remedial actions.

4. EXPERIMENT

4.1 Experimental setup

To prepare a rich dataset for validation, CCTV videos taken in different angles were collected, including those taken by at-grade cameras and mounted-on-crane cameras. **Table 2** summarizes the attributes and sources of the videos solicited, which can be referred to in future studies for tower crane safety monitoring.

Table 2: Statistics of the image dataset collected for model evaluation

Angle	Length	Types	Sources
At-grade	2 min 50 sec	PPE wearing	https://youtu.be/zmVjnWEX_5c
At-grade	24 min 36 sec	Crane operations	https://youtu.be/Ag5yV8qZKMQ
At-grade	15 min 56 sec	Worker behavior	https://youtu.be/3AbhT6TLf60
Top-down	3 min	Crane operations	https://youtu.be/IlaEJgq0aEw
Top-down	4 min 18 sec	Crane operations	https://youtu.be/Vg6SOcPviDs
Top-down	4 min 35 sec	Crane operations	https://youtu.be/lrhQHX3r-pM
Top-down	59 sec	Crane operations	https://youtu.be/viBcyF2H_1A

A total of 5575 images were generated by extracting frames out of the collected videos, with manual inspection and sampling of high-quality frames, i.e. capturing diverse details of worker / crane operations. A detailed statistics of the dataset generated is summarized in **Table 3**.

Table 3: Statistics of the image dataset collected for model evaluation

Set	No. of images
Training	4889
Validation	458
Testing	228
Total	5575

The collected data then underwent a series of augmentations to maximize the generalization capability of the model being trained. The types of pre-processing include image resizing, rotation by EXIF orientation values and grayscale conversion. The images were also augmented by horizontal and vertical flipping, hue and saturation adjustment. Afterwards, the dataset was split into training, validation and testing sets. The training set was fed into different variants of object detection models, including YOLOv8-Large, YOLOv8-Small and YOLOv8-Nano, which consist of different degrees of model complexity in terms of neural network architecture.

As summarized in **Equations (2)-(4)**, the evaluation metrics of for object detection and classification include recall, precision and average precision (AP) score. Moreover, the accuracy of worker trajectory tracking is evaluated by multi-object tracking accuracy (MOTA), as defined in **Equation (5)**, where $IDSW$ denotes the frequency of identity switching among workers detected. In addition, the computational speed of the proposed method is also evaluated (in frame-per-second), which validates the practicality of our framework in real-time CCTV processing for construction site monitoring.

$$Recall = \frac{TP}{TP + FN} \quad \textbf{(2)}$$

$$Precision = \frac{TP}{TP\ + FP} \quad \textbf{(3)}$$

$$AP\ score = \frac{TP + TN}{TP\ + FP + TN + FN} \quad \textbf{(4)}$$

$$MOTA = 1 - \frac{FN + FP + IDSW}{TP + TN} \quad \textbf{(5)}$$

4.2 Results and discussion

Table 4 summarizes the AP scores of the object detection module of the proposed framework. Overall, a mean AP of 97.0% is achieved among all the three object classes, with 99.5% AP score for both the classes 'hook' and 'material'. A slightly lower AP score of 92.0% is obtained for 'worker', because of the significant variation of worker sizes in the images, which capture both top-down and at-grade angles from largely varying distances.

Table 4: Evaluation results of object detection

	Class	**AP score**
Class-wise	Worker	92.0%
	Crane hook	99.5%
	Lifting materials	99.5%
Overall	Recall	98.0%
	Precision	96.0%
	Mean AP	97.0%

Table 5 summarizes the AP scores of the worker PPE classification module of the proposed framework. Overall, the mean AP improves from 96.9% to 99.5% when training the PPE classification module with both confirming and dis-confirming cases. Such an approach also boosts the class-wise AP scores, from 96.7% to 99.3% ('helmet') and from 97.0% to 99.5% ('vest'). The effect of incorporating the dis-confirming cases is that the classification model has learnt more distinctive features of those PPE from the negative samples. For instance, by seeing ordinary cloths without vest, the model intrinsically learns better how a vest should look like and hence more accurately classifies whether a person is properly wearing a vest.

Table 5: Evaluation results of worker PPE classification

		Case 1 – trained with confirming classes only	**Case 2 – trained with confirming & dis-confirming classes**
Class-wise AP scores	Helmet	96.7%	99.3% ↑
	No helmet	/	99.2%
	Vest	97.0%	99.5% ↑
	No vest	/	99.6%
Overall scores	Recall	96.9%	98.8% ↑
	Precision	97.1%	98.8% ↑
	Mean AP	96.9%	99.5% ↑

Table 6 summarizes the MOTA (for worker tracking) and computational speeds when combining DeepSORT with different YOLOv8 variants. Regarding worker tracking accuracy, YOLOv8-Large outperforms the other two models with the highest MOTA of 90.1%, while having slower computational speed than the other two (2.7 frames per second). YOLOv8n-Nano shows the fastest inferencing (13.4 frames per second), while its MOTA is 81.8% which may be due to the increased chance of missing detections. Hence, YOLOv8-Small achieves the most balanced performance (85.2% MOTA and 7.9 frames per second).

Table 6: Evaluation results of worker trajectory tracking and overall computational speed

Model variant	**MOTA**	**Computational speed (frame-per-second)**
YOLOv8-Large+DeepSORT	90.1%	2.7
YOLOv8-Small+DeepSORT	85.2%	7.9
YOLOv8-Nano+DeepSORT	81.8%	13.4

Figure 6 illustrates the predictive warning mechanism of the proposed framework. The developed modules process a complete video and identifies that construction workers are working within the danger zone during lifting operations. By detecting the location of the lifting equipment and tracking the movement of individual construction workers, warning signals and recommended actions are dispatched via a Telegram chatbot upon identifying the unsafe scenarios. The spatial relationship among the equipment and workers is accurately established, which then informs on-site safety managers of the workers' risk statuses, urging for immediate actions more timely. Hence, our proposed framework enables more predictive safety monitoring of crane operations.

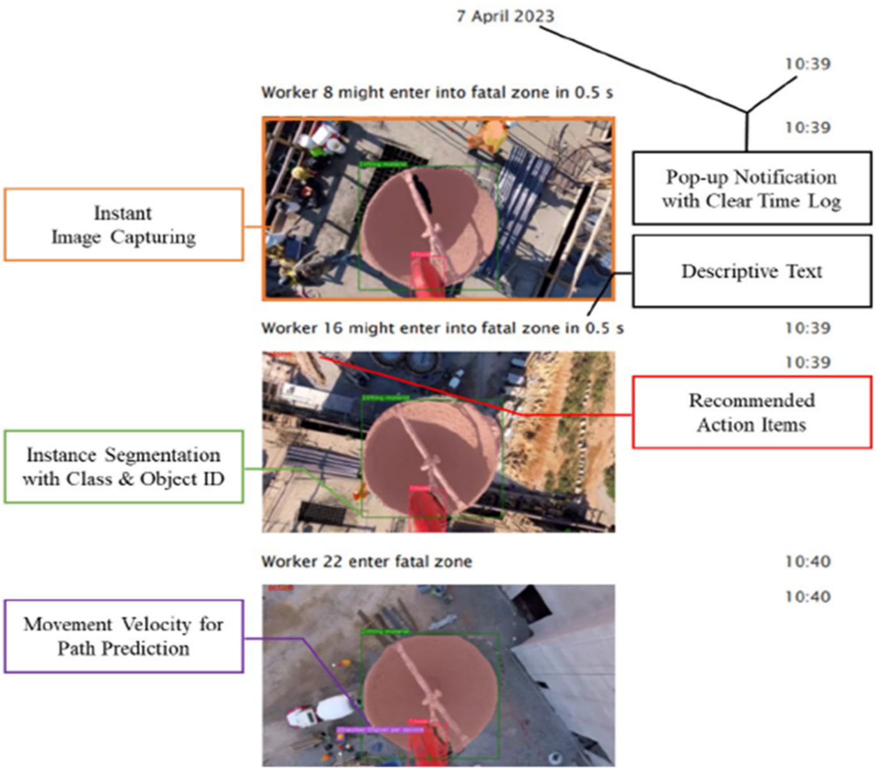

Figure 6: Demonstration of predictive warning mechanism in video processing

5. CONCLUSION AND FUTURE WORK

This paper proposes a vision-based framework for predictive lifting safety monitoring, relieving the tedious and error-prone manual inspection on sites in traditional practices. By analyzing the spatial interaction among essential objects in lifting operations (e.g. predicted movement of workers, danger zone around hook and lifting materials) more predictive incident identification is enabled for timely on-site safety assessment. The competitive accuracy and computational efficiency demonstrated in this study validates the practicality of the proposed framework. Based on the experimental findings, two research directions are suggested for future research: (1) **camera placement optimization in actual deployment**, considering various factors such as view coverage, degree of object occlusion, view angle and distance (implying video quality and hence analytical accuracy), etc. Research effort may be devoted into quantifying these factors into optimization framework formulated for camera placement, including the number of cameras, their position and orientation, etc.; (2) **multi-modal sensor integration**, extending the vision-based methodology to analyze more worker behavior such as injury/fall detection, and possibly also incorporating other kinds of sensors such as temperature sensor for heat-stroke warning monitoring and proximity sensor for worker-equipment conflict. More comprehensive research in the future will contribute to forming a systematic approach for construction safety monitoring.

REFERENCES

1. Brunetti, A., Buongiorno, D., Trotta, G. F., & Bevilacqua, V. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. Neurocomputing, 300, 17-33. https://doi.org/10.1016/j.neucom.2018.01.092
2. Cheng, J. P., Wong, P. K. Y., Luo, H., Wang, M., & Leung, P. H. (2022). Vision-based monitoring of site safety compliance based on worker re-identification and personal protective equipment classification. Automation in Construction, 139, 104312. https://doi.org/10.1016/j.autcon.2022.104312
3. Chian, E. Y. T., Goh, Y. M., Tian, J., & Guo, B. H. (2022). Dynamic identification of crane load fall zone: A computer vision approach. Safety science, 156, 105904. https://doi.org/10.1016/j.ssci.2022.105904
4. Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. Automation in construction, 85, 1-9. https://doi.org/10.1016/j.autcon.2017.09.018
5. Golovina, O., Teizer, J., & Pradhananga, N. (2016). Heat map generation for predictive safety planning: Preventing struck-by and near miss interactions between workers-on-foot and construction equipment. Automation in construction, 71, 99-115. https://doi.org/10.1016/j.autcon.2016.03.008
6. HKSARG Development Bureau (2020). Lifting Safety Handbook. Construction Safety Week 2020. [Online]. Available: https://www.cic.hk/files/page/51/Lifting%20Safety%20Handbook%20%E5%90%8A%E9%81%8B%E5%AE%89%E5%85%A8%E6%89%8B%E5%86%8A.pdf. [Accessed: Oct. 29, 2022]
7. HKSARG Labour Department (2018). Occupational Safety and Health Statistics: Bulletin Issue No. 18. [Online]. Available: https://www.labour.gov.hk/eng/osh/pdf/Bulletin2017_issue18_eng.pdf. [Accessed: 18-Dec-2022]
8. Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer I., Wojna Z., Song Y., Guadarrama S., & Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7310-7311). https://doi.org/10.1109/cvpr.2017.351
9. Jeelani, I., Asadi, K., Ramshankar, H., Han, K., & Albert, A. (2021). Real-time vision-based worker localization & hazard detection for construction. Automation in Construction, 121, 103448. https://doi.org/10.1016/j.autcon.2020.103448
10. Kim D., Liu M., Lee S. H., & Kamat V. R. (2019). Remote proximity monitoring between mobile construction resources using camera-mounted uavs. Automation in Construction, vol. 99, pp. 168–182. https://doi.org/10.1016/j.autcon.2018.12.014
11. Liu, W., Meng, Q., Li, Z., & Hu, X. (2021). Applications of computer vision in monitoring the unsafe behavior of construction workers: Current status and challenges. Buildings, 11(9), 409. https://doi.org/10.3390/buildings11090409
12. Luo, H., Wang, M., Wong, P. K. Y., Tang, J., & Cheng, J. C. (2021). Construction machine pose prediction considering historical motions and activity attributes using gated recurrent unit (GRU). Automation in Construction, 121, 103444. https://doi.org/10.1016/j.autcon.2020.103444
13. Luo, X., Li, H., Wang, H., Wu, Z., Dai, F., & Cao, D. (2019). Vision-based detection and visualization of dynamic workspaces. Automation in Construction, 104, 1-13. https://doi.org/10.1016/j.autcon.2019.04.001
14. Memarzadeh, M., Golparvar-Fard, M., & Niebles, J. C. (2013). Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors. Automation in Construction, 32, 24-37. https://doi.org/10.1016/j.autcon.2012.12.002
15. Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020). Deep learning for site safety: Real-time detection of personal protective equipment. Automation in Construction, 112, 103085. https://doi.org/10.1016/j.autcon.2020.103085
16. Shafique, M., & Rafiq, M. (2019). An overview of construction occupational accidents in Hong Kong: A recent trend and future perspectives. Applied Sciences, 9(10), 2069. https://doi.org/10.3390/app9102069
17. Son, H., Seong, H., Choi, H., & Kim, C. (2019). Real-time vision-based warning system for prevention of collisions between workers and heavy equipment. Journal of Computing in Civil Engineering, 33(5), 04019029. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000845
18. U.S. Bureau of Labor Statistics (BLS) (2018). National Census of Fatal Occupational Injuries in 2017. [Online]. Available: www.bls.gov/iif/oshcfoil.htm. [Accessed: 18-Dec-2022]
19. Wojke, N., Bewley, A., & Paulus, D. (2017, September). Simple online and realtime tracking with a deep association metric. In 2017 IEEE international conference on image processing (ICIP) (pp. 3645-3649). IEEE. https://doi.org/10.1109/icip.2017.8296962
20. Wong, P. K. Y., Luo, H., Wang, M., Leung, P. H., & Cheng, J. C. (2021). Recognition of pedestrian trajectories and attributes with computer vision and deep learning techniques. Advanced Engineering Informatics, 49, 101356. https://doi.org/10.1016/j.aei.2021.101356

21. Zhang, M., Cao, Z., Yang, Z., & Zhao, X. (2020). Utilizing computer vision and fuzzy inference to evaluate level of collision safety for workers and equipment in a dynamic environment. Journal of Construction Engineering and Management, 146(6), 04020051. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001802