A Deep Learning-based Predication of Fall Portents for Lone Construction Worker

Numan Khan[a,b], Sharjeel Anjum[a], Rabia Khalid[a], JunSung Park[a], and Chansik Park [a,*]

[a] Department of Architectural Engineering, Chung-Ang University, Seoul 06974, South Korea.
[b] Department of Architecture, Ajman University, Ajman, United Arab Emirates.
E-mail: numanpe@gmail.com, sharjeelanjum@cau.ac.kr, rabiakhalid@cau.ac.kr, qkrwnstjd0@cau.ac.kr, and cpark@cau.ac.kr

Abstract –

As construction projects resume worldwide and workers return to the job site, the possibility of transmitting the Covid-19 could be added to the extensive list of risks confronting workers in the construction sites; thus, the workers need to work alone in an assigned activity. Many workers are already working alone in the construction sites, such as utility workers, repair technicians, teleworkers, operators, and drivers. Lone workers in construction are subjected to greater safety risks compared with those working alongside others. Considering the accidents faced by lone workers, it's less likely that another person would be there to aid them - and if they don't get treatment quickly enough, serious injuries might prove deadly. Currently, the construction sites depend on physical inspections to the construction sites and manual observation of video streams generated through close circuit television (CCTV). To solve this issue, this research work presents an automated deep learning-based fall detection system of a lone worker to provide information of severe situations and help the workers in their golden time. A diverse dataset of multiple scenarios having workers with the excavator, forklift, ladder, and mobile scaffold is established, and a deep learning algorithm has been trained to validate the concept. The developed system is expected to reduce the efforts being made in manual inspection, enhance the timely access of the due aid from co-workers and supervisors, which is more easily obtainable in non-lone working situations.

Keywords –

Deep Learning; Covid-19; Construction Hazards; Worker Safety; Lone Person Fall

1 Introduction

Despite limited access to timely assistance, lone workers independently cope with potentially risky situations such as extreme weather conditions, tools, and equipment failure. If alone the injured worker receives aid promptly, the injuries could turn out to be fatal, among other accidents in a construction job site. Fall accidents are common and severe accidents that can happen anywhere on a construction site. Even if workers fall from a low height (1m or 2m), it frequently results in significant mishaps, probably death.

Substantial efforts are being made to significantly reduce fall accidents in Korean construction job sites and enhance construction safety management. The extensive efforts include quality safety education, advanced safety training, high-elevation work management rules, and the use of fall prevention protective equipment. As per the Korea Occupational Safety and Health Agency (KOSHA)'s industrial accidents and analysis (2009 to 2017), fall accidents in the construction industry represent a considerable share of 47.7% to 52.1% [1]. This tendency was also observed in various countries, including the United States [2], Singapore [3], Norway [4], and Hong Kong [5].

Consequently, falls from great heights have received extensive research attention and have become an essential topic in the construction industry. Construction workers are prone to weariness, drowsiness, and loss of balance, increasing safety risks and fall accidents due to their severe physical needs and irregular lifestyle (e.g., alcohol, misuse, night shifts, and insufficient rest interval). However, most of the studies focused on safety facilities and PPE inspection; this only helps reduce the severity of damages rather than providing quick aid and timely assistance when a fall accident happens. Therefore, this research work presents the inevitable approach to detecting a person in falling conditions, which will ultimately help employers automatically monitor lone workers and respond timely to any falling accidents.

The current stage of computer vision application in the construction industry is covered in Section 2. Section-3 describes the dataset and model development. Section-4 includes evaluating the developed model using the performance indicators to validate the feasibility and practicality of the proposed system's for actual implementation. A conclusion is included at the end of the work.

2 Literature Review

Despite the efforts of researchers, safety specialists, and strict enforcement of safety rules, construction accidents, and fatalities have not appreciably decreased. Falls are a significant public health concern worldwide and one of the leading causes of severe and fatal injuries among construction workers. Researchers and experts in construction safety and health management have devoted tremendous efforts to prevent falls [6]. Proactive and passive strategies such as safety training education and preventive measures on safety accident analysis are developed to prevent and minimize the severe injuries generated from FFH [7]. For instance, using accident records and data from routine safety inspections to identify factors contributing to deadly occupational falls [7]. Fixed safety equipment (e.g., guardrails and opening covers), fall arrest systems (e.g., full-body harness), and travel restraint systems (e.g., belts); are FFH preventative measures generated from an examination of accident data. However, the social distancing regulations imposed by governments worldwide lead many workers to deal with assigned activity solely, and the construction workers already working alone are comparatively facing a greater risk of injuries. Considering the accidents faced by lone workers, it’s less likely that another person would be there to aid them - and if they don’t get treatment quickly enough, serious injuries could become a fatality.

Researchers have leveraged recent technological advancements to automate safety management procedures. Until yet, very little attention has been devoted to protecting the lone worker accidents generated through FFH. Sensor-based technology has garnered a lot of attention in recent decades for monitoring PPEs worn by construction workers [8,9], such as activity recognition [10,11] and safety of the construction workers [9,12–16]. In particular, most FFH prevention studies used sensors to detect risky behaviour to avoid FFH accidents by analyzing workers' motions, body positions, and walking patterns [17–19]. Furthermore, many researchers have developed methods to detect workers' actions and body postures using wearable sensors' signals [10,20]. Yang et al. [10] Yang et al. [10] investigated workers' behavior patterns from three angles: while conducting the lab experiment, they attached the sensor to the workers' waist and recorded acceleration and angular velocity.

The system's accuracy in predicting unsafe behavior, was 98.6% and 60.9 percent, respectively. Owing to construction site complexity, the prediction model had a hard time distinguishing behaviors related to moves that weren't in the training data. Furthermore, most sensor-based monitoring devices possess issues concerning noise, precision, loss, and errors in the data they acquire.

Apart, In the recent decade, several researchers have been attracted to use computer vision in their fields, such as the construction industry for worker safety monitoring, progress monitoring, and worker action recognition to automate the manual procedures at the construction site. Recently researchers have been focusing on computer-vision-based safety monitoring of the worker [6,21–25]. A growing number of recent studies in the construction industry have focused on using CNN-based methods, such as Faster R-CNN, R-FCN, SSD, Retinanet, YOLOv3 for detecting workers, faces to recognize non-certified persons, non-hardhat-use, equipment for activity recognition, guardrails, PPE for steeplejacks [6,24–26]. Fang et al. [27] developed a deep learning-based approach to detect non-hardhat users in a construction site. Fang et al. [6] developed an automated approach to detect safety harnesses to prevent heights falling using double-layer CNN. Likewise, Ding et al. [28] presented convolution neural networks (CNN) and long short-term memory (LSTM) that automatically identify unsafe behavior by detecting humans climbing on a ladder. Khan et al. [23] proposed Mask R-CNN-based algorithm to monitor the worker's safety while working on the mobile scaffolding. Weili Fang et al. [29] used Mask R-CNN-based algorithm to recognize the unsafe behavior of construction workers traversing structural support during the construction. Nath et al. [30] developed three Deep learning (DL) models on YOLO architecture to detect PPEs.

Despite these efforts being made to integrate the computer vision for an intelligent construction site, considerable efforts are required to extend the computer vision application in the construction industry. Currently, the construction sites around the globe is recommencing again, and the workers are joining back the construction job site. The possibility of transmitting the Covid-19 could be another risk confronting by the workers in the construction sites; thus, the workers need to work solely or on a distance from another worker to complete the given tasks. Moreover, many workers are already working solely in the construction sites, such as utility workers, repair technicians, teleworkers, operators, and drivers. As we all know, lone workers in construction are subjected to greater safety risks than those working alongside others. Contemplating the lone worker's mishaps or severe accidents, the probability of helping another person would be less in that situation - and if they don’t get first-aid quick enough in their golden times, these serious injuries might result in death. Currently, the construction sites depend on physical inspections to the construction sites and manual observation of video streams generated through close circuit television

(CCTV). Correspondingly, this research study has tried to put efforts on the appropriate and exceptional use of the provided rich information generated through digital data. Thus, this research proposed a deep learning-based worker's fall detection system, enhancing timely access to due aid from co-workers and supervisors.

3 System Development

This section includes a comprehensive discussion on dataset Preparation, deep learning model selection, and Model Training. The dataset preparation part focuses on establishing scenarios, digital data collection, cleaning, and finalizing the data for ultimate deep learning. The appropriateness and feasibility of the deep learning model specific to lone person detection are stated under the subsection (deep learning model selection). The model training section stipulates the specification adopted during the training process of the deep learning model.

3.1 Dataset Preparation

A considerable amount of digital image data with diverse patterns is needed to train a vision intelligence-based detection algorithm. As vision intelligence technologies emerged recently in the construction industries, collecting labeled datasets from open-source websites remains problematic. Therefore, images and multiple videos for lone person detection with enough variations considering scenarios with ladder, forklift, Excavator, and Mobile Scaffolding have been recorded at the Korean scaffolding institute in Seoul, Korea. Random frames were extracted using Fast Forward MPEG (FFmpeg) tool in python. FFmpeg is an open-source software package that comprises many libraries dealing with audio, video, and other multimedia files. The inappropriate images such as (irrelevant) and duplicate images from the same scenes were removed from the dataset during the dataset cleaning process.

The decisive image data of 799 images were uploaded to the roboflow platform for pre-processing and labelling. A total of 2037 annotations were labelled across the six classes in the dataset (1) Person Falling, (2) Worker, (3) Ladder, (4) Forklift, (5) Excavator, and (6) Mobile Scaffolding. The dataset was separated in the following ratio: 91:4:5. As a result, 1857 images from the training set, 80 images from the validation set, and 100 images from the test set were gathered for further experiment.

As mentioned in the literature section, the deep learning models required large enough data for training to correctly identify and recognize the interested objects. The data augmentation techniques have been utilized to increase the image data. Apart from dataset maximization, data augmentation techniques increase the accuracy of deep learning models. According to a conducted experiment [31], a deep learning-based model with image augmentation outperforms a deep learning model without image augmentation in terms of training loss and accuracy and validation loss and accuracy for image classification tasks. In the augmentation process, we have performed flip (Horizontal), shear (15°), hue (between -25 to +25), and brightness (between -19 to +19) on the training set to increase the dataset. After the augmentation process, the total number of image data is increased to 2037 images. The process of labeling the dataset is exhibited in the below figure (see Figure 1).

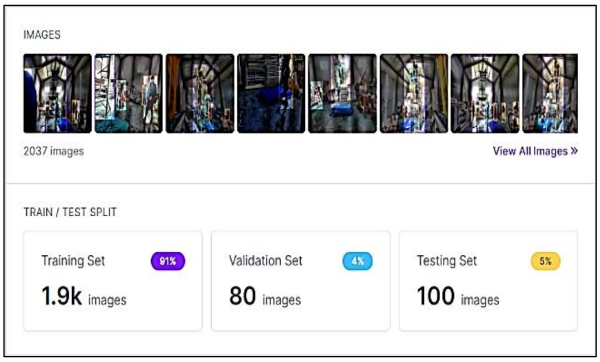

Figure 1. Dataset establishment in Roboflow platform

3.2 Deep learning Model Selection

In recent years, cutting-edge technology has been used in computer vision applications to detect resources. The previous efforts reported two primary categories of object detectors: single-stage object detectors and double-stage object detectors. Depending on the problem to be solved, one could choose among them (single-stage or double-stage object detector). The main difference between a single-stage and a two-stage object detector is that a single-stage object detector's output can be acquired after the first CNN (Convolution Neural Network) operation. The high-score area proposals obtained from the first-stage CNN are typically passed to the second-stage CNN for the final prediction in the case of the double-stage object detector.

The inference times of single-stage and double-stage detectors could be defined as:

$$T_{one} = T_1^{st} \text{ and } T_{two} = T_1^{st} + mT_2^{nd}$$

The above equation defines m as the number of area recommendations with a confidence score greater than a threshold. In the case of real-time object detection, the single-stage object detectors are preferred to use because the inference time of the single-stage detectors is constant but for the double-stage object detectors are variable [32]. Consequently, in our case, a real-time object detection of the lone falling person on the construction site is significant to report and assist them as soon as convenient.

Therefore, this research work adopted a recently emerged single-stage object detector type named as Scaled-YOLOv4 algorithm. The model scaling technique is an important approach for effectively detecting objects with significant accuracy and real-time recognition on numerous kinds of devices, such as embed devices, while minimizing computing resources. The most common model scaling techniques are the depth (convolutional layers' numbers in a CNN) and width (convolutional filters' numbers in a convolutional layer) of the architecture's backbone. A deep learning-based algorithm (scaled YOLOV4) for object detection was recently published as an addition to the pool of detection techniques integrating scaling techniques. This scaling technique showed significance in terms of performance on small and large networks without compromising the model's performance [32].

The authors of the Scaled YOLOv4 manuscript have designed a different version of the Scaled YOLOv4 for low-end GPU, and high-end GPU, and the general GPU. The YOLOv4-CSP is re-designed of the original YOLOv4 for the general GPUs with high performance and accuracy. The authors designed YOLOv4-tiny with a simpler structure and reduced parameters to make it feasible for the development on mobile and other edge devices (Andriod, etc.). YOLOv4-large is designed for the high-end GPUs such as cloud servers, the main goal of proposing this is to achieve high accuracy while minimizing training time and achieving efficient performance. A fully CSP-ized model has been created named as YOLOv4-P5 and scaled it up to YOLOv4-P6 and YOLOv4-P7 [32]. The structure of YOLOv4-P5, YOLOv4-P6, and YOLOv4-P7 is depicted in Figure 2. The authors of the Scaled YOLOv4 have performed the compound scaling of the architecture backbone on size input, stage, and width scaling. The inference time is used as a constraint for the additional width scaling of the architecture backbone. The authors conducted experiments on the MSCOCO-2017 dataset to validate the proposed scaled-YOLOv4-large, and the results show that the YOLOv4-P6 can achieve real-time performance at 30 frames per second (when the width scaling factor is set to 1) and the YOLOv4-P7 can achieve the real-time performance of 16 frames per second (When scaling factor of the width is 1.25) [32]. As a Result, scaled YOLOv4-large is selected to detect lone person falls on a construction.

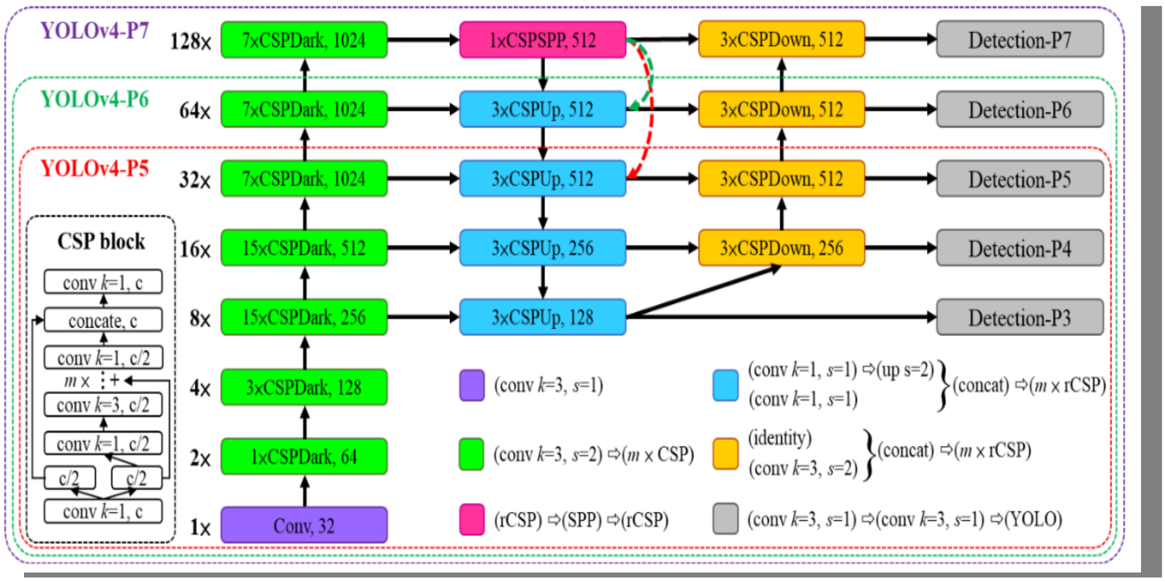

Figure 2 System architecture of YOLOv4-large, with YOLOv4-P5, YOLOv4-P6, and YOLOv4-P7 [32]

3.3 Model Training

The experiment was conducted using an open-source Scaled YOLOv4 GitHub repository. The required repository of YOLOv4 was clone to the colab environment, all the dependencies, such as the torch mish activation function for Cuda, were imported. The training of the model is performed using Intel Core i7 9th generation with NVIDIA GeForce RTX 2080Ti. The hyper-parameters of the Scaled YOLOv4-large can be seen in the given Table1. To perform detection, we have modified hyper-parameters in the configuration file so that the step size is set to 9600,10800, the learning rate to 0.0026. The momentum and the weight decay were set 0.949 and 0.0005, respectively. Maximum batches were set to 12000, and the epochs were set to 1000. An iteration indicates a specific change to a model's weights, while an epoch, on the other hand, determines one iteration across the whole dataset.

Table 1. Parameters of Scaled YOLOV4

Parameters	Values
Input Size	416 x 416
Batch Size	16
Learning Rate	0.00261
Momentum	0.949
Decay	0.0005
Iterations	1000
Classes	6

4 Results and Evaluation

The feasibility of the trained lone person fall detection model is tested with the different performance indicators. The model test and evaluation are performed on test and validation dataset. The Figure 3 reveals the correctly detected results of the lone person falls. The developed model successfully detected all the objects in a given testing and validation dataset. The efficacy of the trained model is quantified using mean average precision (mAP), a single numerical number that indicates the effectiveness of the entire system in object identification (and information retrieval). Other evaluation matrices such as Precision, Recall, are also used to check the performance of the developed system.

Figure 3. Detected Results of the Trained Model

4.1 Precision

The number of true positives (Tp) divided by the number of false positives (Fp) plus the number of true positives (Tp) makes precision *(P)* value of the model. False positives (Fp) are instances where the model mistakenly identifies something as positive when it is truly negative. Precision actually measures how many of the predicted positives were truly positive. It is mathematically denoted as follows:

$$P = \frac{Tp}{Fp + Tp} \quad (1)$$

The below figure (see Figure 4.) shows the precision (48.5%) of our proposed model. As we detect the person falling, the precision and recall could be adjusted by selecting an optimum point on the precision-recall curve.

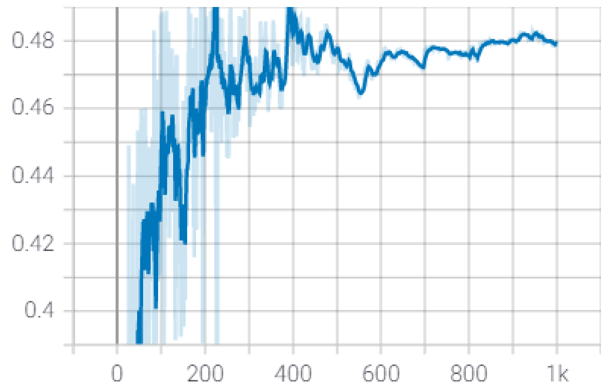

Figure 4. Graphical representation of precision matrix

4.2 Recall

A recall expresses the capacity to discover all relevant instances in a test dataset. Recall (R) measures how many true positives were successfully found. Simply, the number of true positives (Tp) is divided by the number of false negatives (F_N) plus true positives (Tp). The mathematical form of recall can be written as:

$$R = \frac{Tp}{F_N + Tp} \quad (2)$$

Recall (83%) of our proposed scaled-YOLOV4 based model is depicted in the Figure 5.

4.3 F1-Score

The F1-score is a metric for how accurate a model performs on a given test data. It is commonly applied to examine binary classification algorithms that categorize examples as either "negative." or "positive". The F-score, which is defined as the harmonic mean of the model's recall and accuracy, is an average of combining the recall and precision of the model.

$$F1 - Score = 2.\frac{(Precision)\ .\ (Recall)}{(Precision)\ +\ (Recall)}$$

(3

The F1-score observed during the experiment was 60.82%.

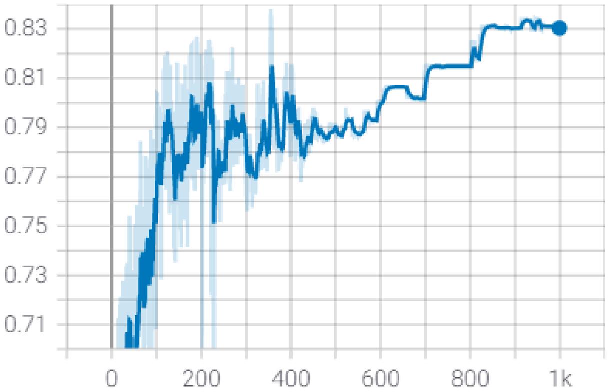

Figure 5. Graphical representation of recall

4.4 Mean Average Precision

The average of Average Precision (AP) is called mAP (mean average precision); mAP allows you to demonstrate the overall system's usefulness as a single numerical value. The following Figure 6 shows the mAP calculated when the Intersection of Union IoU was set to 0.5 and the mean average precision (mAP) obtained in this case was 72.5%.

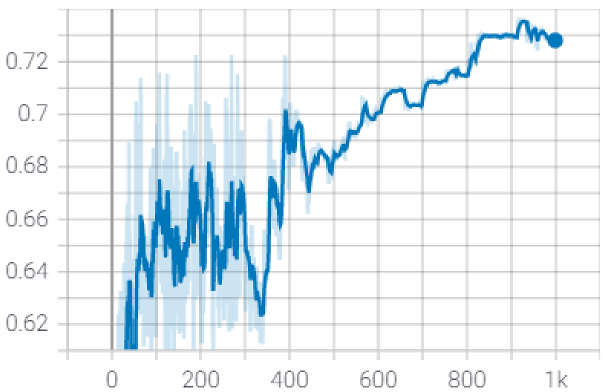

Figure 6. Graphical representation of mAP at 0.5

The following picture shows the mAP calculated on the different IoU thresholds ranging between 0.5 to 0.95 and achieved 37%mAP (see Figure 7.).

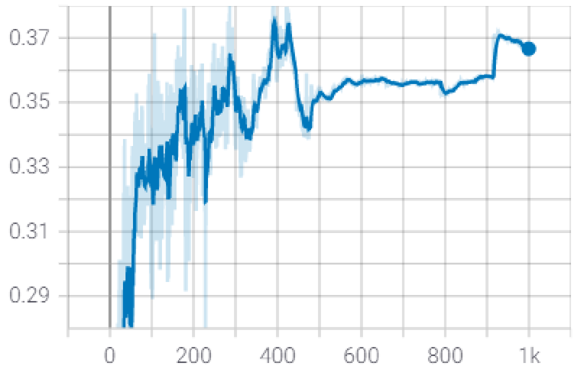

Figure 7. Graphical representation of mAP at 0.5 to 0.95

5 Conclusion

This research work introduced, developed, and evaluated a deep learning-based system for lone person detection on scaled-YOLOv4 architecture. For instance, if an alone worker faces any fall accidents on the construction site, the system will automatically recognize the falling person. To develop the proposed system, diverse digital image data has been gathered with different scenarios. The designed deep learning-based lone person fall detection system was successfully implemented on the test dataset. The detection model's quantitative data indicated a mean average precision (mAP) of 72.50 percent, precision of 47.5 percent, recall of 83 percent, and F1-score of 60.82 percent. According to the findings, the presented system had good accuracy in detecting person falling conditions. As a result, it is expected that this technique will have a major impact on the automated detection of a lone person fall among construction workers. It is also anticipated that the construction industry could take advantage of the system, particularly in the current scenario of COVID-19.

Acknowledgment

This work was supported by the National R&D project for Smart Construction Technology (No. 21SMIP-A158708-02)” funded by Korea Agency for Infrastructure Technology Advancement under the Ministry of Land, Infrastructure and Transport, and managed by the Korea Expressway Corporation. The dataset used for training is from “Construction Site Safety Equipment Recognition Image Dataset (https://aihub.or.kr)" supported by the National Information Society Agency with the funding of Korea Ministry of Science and ICT.

References

[1] K. Kang, H. Ryu, Predicting types of occupational accidents at construction sites in Korea using random forest model, Saf. Sci. 120 (2019) 226–236. https://doi.org/10.1016/J.SSCI.2019.06.034.

[2] B. Evanoff, A.M. Dale, A. Zeringue, M. Fuchs, J. Gaal, H.J. Lipscomb, V. Kaskutas, Results of a fall prevention educational intervention for residential construction, Saf. Sci. 89 (2016) 301–307. https://doi.org/10.1016/J.SSCI.2016.06.019.

[3] Y.M. Goh, W.M. Goh, Investigating the effectiveness of fall prevention plan and success factors for program-based safety interventions, Saf. Sci. 87 (2016) 186–194. https://doi.org/10.1016/J.SSCI.2016.04.007.

[4] S. Winge, E. Albrechtsen, Accident types and barrier failures in the construction industry, Saf. Sci. 105 (2018) 158–166. https://doi.org/10.1016/J.SSCI.2018.02.006.

[5] Occupational Safety and Health Statistics, (n.d.).

[6] W. Fang, L. Ding, H. Luo, P.E.D. Love, Falls from heights: A computer vision-based approach for safety harness detection, Autom. Constr. 91 (2018) 53–61. https://doi.org/10.1016/j.autcon.2018.02.018.

[7] E. Nadhim, C. Hon, B. Xia, I. Stewart, D. Fang, Falls from Height in the Construction Industry: A Critical Review of the Scientific Literature, Int. J. Environ. Res. Public Health. 13 (2016) 638. https://doi.org/10.3390/ijerph13070638.

[8] X. Yang, Y. Yu, S. Shirowzhan, S. Sepasgozer, H. Li, Automated PPE-Tool pair check system for construction safety using smart IoT, J. Build. Eng. 32 (2020) 101721. https://doi.org/10.1016/j.jobe.2020.101721.

[9] M. Khan, R. Khalid, S. Anjum, N. Khan, C. Park, IMU based Smart Safety Hook for Fall Prevention at Construction Sites, in: IEEE TENSYMP, 2021: pp. 1–6.

[10] K. Yang, sepi aria, changbum R. Ahn, T.L. Stentz, Automated Detection of Near-miss Fall Incidents in Iron Workers Using Inertial Measurement Units, Constr. Res. Congr. 2014. (2014) 140–149.

[11] S.S. Bangaru, C. Wang, S.A. Busam, F. Aghazadeh, ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors, Autom. Constr. 126 (2021). https://doi.org/10.1016/j.autcon.2021.103653.

[12] M.N. Nyan, F.E.H. Tay, E. Murugasu, A wearable system for pre-impact fall detection, J. Biomech. 41 (2008) 3475–3481. https://doi.org/10.1016/j.jbiomech.2008.08.009.

[13] N.Y. Kim Y., J. Haneul, Koo B.,Kim J., Kim T., Detection of Pre-Impact Falls from Heights Using an Inertial Measurement Unit Sensor, Sensors. (2020).

[14] C. Nnaji, A. Jafarnejad, J. Gambatese, Effects of Wearable Light Systems on Safety of Highway Construction Workers, Pract. Period. Struct. Des. Constr. 25 (2020) 04020003. https://doi.org/10.1061/(asce)sc.1943-5576.0000469.

[15] I. Okpala, C. Nnaji, A.A. Karakhan, Utilizing Emerging Technologies for Construction Safety Risk Mitigation, Pract. Period. Struct. Des. Constr. 25 (2020) 04020002. https://doi.org/10.1061/(asce)sc.1943-5576.0000468.

[16] S. Jeon, S. Kim, S. Kang, K. Kim, Smart Safety Hook Monitoring System for Construction Site, 2020 IEEE Int. Conf. Consum. Electron. - Asia, ICCE-Asia 2020. (2020) 19–22. https://doi.org/10.1109/ICCE-Asia49877.2020.9277155.

[17] R. Navon, O. Kolton, Algorithms for Automated Monitoring and Control of Fall Hazards, J. Comput. Civ. Eng. 21 (2007) 21–28. https://doi.org/10.1061/(asce)0887-3801(2007)21:1(21).

[18] H. Jebelli, C.R. Ahn, T.L. Stentz, Fall risk analysis of construction workers using inertial measurement units: Validating the usefulness of the postural stability metrics in construction, Saf. Sci. 84 (2016) 161–170. https://doi.org/10.1016/j.ssci.2015.12.012.

[19] K. Yang, C.R. Ahn, M.C. Vuran, S.S. Aria, Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit, Autom. Constr. 68 (2016) 194–202. https://doi.org/10.1016/j.autcon.2016.04.007.

[20] I. Awolusi, E. Marks, M. Hallowell, Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices, Autom. Constr. 85 (2018) 96–106. https://doi.org/10.1016/j.autcon.2017.10.010.

[21] C. Park, D. Lee, N. Khan, An Analysis on Safety Risk Judgment Patterns Towards Computer Vision Based Construction Safety Management, (2020) 52. https://doi.org/10.3311/CCC2020-052.

[22] D. Lee, N. Khan, C. Park, Stereo vision based hazardous area detection for construction worker’s safety, Proc. 37th Int. Symp. Autom. Robot. Constr. ISARC 2020 From Demonstr. to Pract. Use - To New Stage Constr. Robot. (2020) 935–940.

https://doi.org/10.22260/isarc2020/0129.
[23] N. Khan, M.R. Saleem, D. Lee, M.W. Park, C. Park, Utilizing safety rule correlation for mobile scaffolds monitoring leveraging deep convolution neural networks, Comput. Ind. 129 (2021) 103448. https://doi.org/10.1016/j.compind.2021.103448.
[24] M.-W. Park, N. Elsafty, Z. Zhu, Hardhat-Wearing Detection for Enhancing On-Site Safety of Construction Workers, J. Constr. Eng. Manag. 141 (2015) 04015024. https://doi.org/10.1061/(asce)co.1943-7862.0000974.
[25] H. Luo, X. Luo, Q. Fang, H. Li, C. Li, L. Ding, Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment, Autom. Constr. 93 (2018) 148–164. https://doi.org/10.1016/j.autcon.2018.05.022.
[26] Q. Fang, H. Li, X. Luo, L. Ding, T.M. Rose, W. An, Y. Yu, A deep learning-based method for detecting non-certified work on construction sites, Adv. Eng. Informatics. 35 (2018) 56–68. https://doi.org/10.1016/j.aei.2018.01.001.
[27] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T.M. Rose, W. An, Detecting non-hardhat-use by a deep learning method from far-field surveillance videos, Autom. Constr. 85 (2018) 1–9. https://doi.org/10.1016/j.autcon.2017.09.018.
[28] L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, X. Ouyang, A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory, Autom. Constr. 86 (2018) 118–124. https://doi.org/10.1016/j.autcon.2017.11.002.
[29] W. Fang, B. Zhong, N. Zhao, P.E.D. Love, H. Luo, J. Xue, S. Xu, A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network, Adv. Eng. Informatics. 39 (2019) 170–177. https://doi.org/10.1016/j.aei.2018.12.005.
[30] N.D. Nath, A.H. Behzadan, S.G. Paal, Deep learning for site safety: Real-time detection of personal protective equipment, Autom. Constr. 112 (2020). https://doi.org/10.1016/j.autcon.2020.103085.
[31] A. Takimoglu, Data augmentation Techniques, AI Mult. (2021). https://research.aimultiple.com/data-augmentation-techniques/ (accessed July 30, 2021).
[32] C.-Y. Wang, H.-Y.M. Liao, Scaled-YOLOv4: Scaling Cross Stage Partial Network, 2021.