Automation in Construction 131 (2021) 103862

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Computer vision approaches for detecting missing barricades

Eugene Chian, Research Associate [a], Weili Fang, PhD, Research Fellow [a], Yang Miang Goh, PhD, Associate Professor [a,*], Jing Tian, PhD, Senior Lecturer [b]

[a] Safety and Resilience Research Unit (SaRRU), Dept. of Building, School of Design and Environment, National Univ. of Singapore, 4 Architecture Dr., 117566, Singapore
[b] Institute of System Science, National University of Singapore, 25 Heng Mui Keng Terrace, 119620, Singapore

ARTICLE INFO

Keywords:
Falls from height
Safety
Computer vision
Unsafe behavior
Deep learning

ABSTRACT

The installation of barricades effectively prevents falls from height (FFH) on construction sites. Common approaches for detecting missing barricades (e.g., manual inspection of the site or three-dimensional models) are not practical due to two inherent challenges: (1) these approaches are labor-intensive and time-consuming; and (2) FFH hazards are dynamic and changing as construction work progresses. To address these challenges, two computer vision-based detection approaches, including *Masks Comparison Approach* (MCA) and *Missing Object Detection Approach* (MODA), are developed in this study to automatically detect missing barricade. The performance of the proposed approaches and their benefits and implementation challenges were evaluated through a case study. The results demonstrate that MODA can achieve better performance and have several implementation advantages over MCA. The average precision and average recall for MODA were 57.9% and 73.6%, respectively. These two approaches can help site managers take action promptly to reduce the risks of FFH accidents.

1. Introduction

Fall from height (FFH) has been identified as the major contributor to fatalities in the construction industry [50]. One of the most common practices to prevent workers from FFH is installing barricades or edge protection [46]. According to the Workplace Safety and Health Council [68], barricades are required for all building edges and edges of excavations, holes, floor openings, and roofs in construction sites to prevent FFH. However, missing barricade is a serious problem in construction. For example, Zlatar et al. [70] analyzed 114 cases and found that guardrails, handrails, barriers, and edge protection failed while working at heigh, accounting for 33% of safety management measures. Likewise, in Navon and Kolton's [37] work, they conducted interviews in 12 construction sites and found a lack of protective measures for openings in the external wall or at the edge of a slab. Furthermore, parts of the guardrails were missing at times. Similarly, in Singapore, open sides and missing guardrails, barriers, or barricades are perennial problems in construction sites (e.g., [2,47]). Therefore, safety inspections and monitoring are required to be conducted based on these requirements to minimize the FFH risk of construction sites.

Planning for FFH risks (e.g., open edges of construction floors) can be undertaken before construction and may be performed using manual observation of floor plans and/or three-dimensional (3D) models [46,65]. However, hazards can change once construction commences. Furthermore, manual safety compliance checks during construction can be labor-intensive, time-consuming, and inconsistent. As a result, safety compliance is difficult to be assured, and therefore FFH remains a major risk for construction sites.

To address the drawbacks of manual monitoring approaches, a computer vision-based approach is proposed in this study to automatically detect missing barricades for the mitigation of FFH risks in construction sites. Despite the proliferation of computer vision solutions for construction safety [15,25,28], the authors were not able to find any research focusing on the detection of missing barricades. To achieve accurate detection of missing barricade, our study focuses on addressing the following three research questions:

(1) What are the possible computer vision-based approaches for automatic detection of missing barricades on construction sites?
(2) Which approach can achieve higher accuracy for detecting missing barricades?
(3) Which approach is more feasible to be implemented on construction sites?

To answer the above research questions, this study identified two possible computer vision-based approaches or strategies for detecting

* Corresponding author.
E-mail address: bdggym@nus.edu.sg (Y.M. Goh).

https://doi.org/10.1016/j.autcon.2021.103862
Received 5 April 2021; Received in revised form 27 July 2021; Accepted 28 July 2021
Available online 25 August 2021
0926-5805/

missing barricades, including the *masks comparison approach* (MCA) and the *missing object detection approach* (MODA). MCA uses computer vision techniques to detect the barricades from each frame and then infer if a barricade is missing by finding the difference between consecutive frames. MODA applies computer vision techniques to detect missing barricade as an object. It must be noted that MCA and MODA are different from algorithms in the sense that they are higher-level strategies to detect missing barricades. Furthermore, it must be emphasized that detecting the absence of an object is not as straightforward as detecting the presence of an object. To validate the performance of our proposed approaches, we implemented the prototype system on a public housing construction site in Singapore. It is noted that to trigger an alert when a worker enters the vicinity of an open edge with missing barricade, the system must be able to detect the worker and the missing barricade. As our previous studies have achieved an acceptable level of accuracy for worker detection (e.g., [13,17]), this study focuses on the detection of missing barricades.

The rest of this study is organized as follows. This study commences by providing a review of related research works on FFH and computer vision for object detection in Section 2. Then, Section 3 describes our developed computer vision approaches. Next, a case study is used to validate the feasibility and effectiveness of the proposed approaches in Section 4. The contributions, limitations, and conclusions are discussed in subsequent sections.

2. Literature review

2.1. Studies on prevention of FFH

According to the hierarchy of control [23], there are five general types of safety controls (in descending level of effectiveness): elimination, substitution, engineering control, administrative control, and personal protective equipment (PPE). Within the context of FFH, there is a strong emphasis on the need for engineering controls (e.g., barricades and guardrails) to be implemented on construction sites. As highlighted earlier, barricades are required for all building edges and edges of excavations, holes, floor openings, and roofs in construction sites to prevent FFH [68]. However, open edges remain a common problem in construction sites around the world [37,47,70]. This could be due to the unique, dynamic, and complex nature of the working environment in construction sites [54].

Despite the importance of barricades in preventing FFH and the difficulties in ensuring their installation on-site, past research on FFH tends to focus on fall arrest systems (a type of PPE) (e.g., [21,24,26,32,63]), and administrative control (e.g., [22]). There are also studies that developed technologies to facilitate the early identification of FFH hazards [67,71], including the use of BIM model for safety compliance checks during the planning phase. For example, Zhang et al. [71] developed a 4D BIM (3D and schedule) based safety rule checking, which can identify and eliminate potential fall hazards during the planning phase. Likewise, Qi et al. [53] expanded the industry foundation classes (IFC) hierarchy to facilitate compliance checks and optimization of building design for preventing FFH accidents during the construction phase. Despite their usefulness, these technologies are used during the design phase and are not directly applicable to the construction phase, where FFH hazards arise as activities progress and change dynamically.

With the increasing interest in using computer vision to solve construction safety problems, some studies use computer vision approaches for mitigating FFH risks in construction sites. For example, Fang et al. [17] developed a computer vision approach with Mask R-CNN to identify construction worker traversing on structural supports. Then, an overlap detection module was used to determine relations between workers and structural support. Similarly, Fang et al. [12] applied a computer vision approach to identify workers not wearing a safety harness when working at height. However, to the best of our knowledge, there is no research that uses a computer vision approach to automatically detect missing barricades for the prevention of FFH accidents. An example of studies focusing on guardrails and barricades is the study by Navon and Kolton [51], who attached sensors to guardrails for inspection of the installation, and warnings alert will be trigged when guardrails were missed or different from the planned ones. Similarly, Zuluaga and Albert [76] use virtual prototyping methods to check bridge guardrails' usage. Cheung and Chan [7] invented a Rapid Demountable Platform (RDP) device to prevent external workers from falling from height. These studies on barricades had not utilized computer vision.

Thus, our review shows that barricade is an important safety control for the prevention of FFH, but the construction industry continues to face problems in preventing the occurrence of the missing barricade. Furthermore, despite the advancement of computer vision, it had not been used to detect missing barricades.

2.2. Object detection and segmentation in construction

A plethora of studies had used computer vision to detect objects in construction sites [12–15]. With significant advancements in deep learning and computer vision (e.g., Faster R-CNN, Yolov3, and SSD), they have been adopted to automatically identify various construction "objects," such as workers, heavy equipment, and plants. For example, Fang et al. [12,13] employed an improved Faster RCNN model to detect workers and excavators in construction sites, which is able to achieve an accuracy of 91% and 95% for detection of workers and excavators, respectively. Son et al. [60] applied a deep residual network-152 for worker detection with varying poses and changing backgrounds in construction sites. The accuracy, precision, and recall of worker detection were 94.3%, 96.03%, and 98.13%, respectively. Hou et al. [30] proposed an improved Mask R-CNN to simultaneously detect and segment object signatures in ground penetrating radar scans. The average accuracy (AP) of detection and segmentation was 58.64% and 47.64%, respectively. Tang et al. [62] employed Faster R-CNN to detect workers, eye protection, face protection, foot protection, hand protection, where the AP was 89.4%, 28%, 25.1%, 67.8%, 66.1%, respectively. Table 1 presents prior research works on deep learning and computer vision-based object detection in the construction industry.

Although there has been significant progress in deep learning-based object detection, visual detection of construction objects [34] is still challenging due to various reasons such as camera movement and shaking, background clutter, intra-class variation, and occlusion [35]. These construction-specific challenges prevent the simplistic application of existing computer vision algorithms. Furthermore, unlike typical object detection, which detects specific types of objects with a more stable set of features, detecting a "missing object" involves more uncertainties in identifying the features of the missing object. Therefore, a

Table 1
A summary of prior works on deep learning and computer vision-based object detection in the construction industry.

Target objects	Methods	Author (Year)
Track components	Yolov4	Guo et al. [28]
13 types of moving object (e.g., workers, tower crane)	Yolov3, SSD300, RetinaNet, FCOS, NAS-FPN, Faster R-CNN, TridentFast	An et al. [1]
Worker	Yolo	Son and Kim [61]
Personal protective equipment (PPE) of workers (e.g., hard hat and vest)	CNN-based classifiers (i.e., VGG-16, ResNet-50, and Xception)	Nath et al. [36]
Worker and excavator	Yolov2	Luo et al. [44]
Worker and excavator	Faster R-CNN	Fang et al. [12,13]

reliable method is needed to detect missing objects under various postures and scales in images.

2.3. Object detection and segementation methods

Many different methods for computer vision-based object detection were studied in the past two decades, and the number of publications has been growing exponentially in the last few years since the emergence of deep learning object detection methods [74]. In the early years, the Histogram of Oriented Gradients (HOG) method [49] was popular, and HOG detectors were used in many computer vision applications [18,45,66]. Computer vision-based object detection studies slowed down for a period until the increasing use of Graphics Processing Unit (GPU,) which enabled the adoption of convolutional neural networks (CNN) [40] after 2011. Since then, it started the deep learning era, and CNN evolved into a wide range of deep learning-based object detection methods, which can be classified into two main groups, namely "anchor-based" and "anchor-free" object detection model [9,73,75]. "Anchor-based" models can be further categorized into "two-stage detection" and "one-stage detection". These different categories will be discussed below.

In the "two-stage detection" approach, the model first proposes a set of regions of interests (ROI) follow by the second stage, where the model processes the proposed ROI and provides the final prediction. Whereas for a "one-stage detection," the model directly detects over a dense sampling of possible locations or grids without the region proposal stage to provide the final prediction. In general, the "two-stage detection" model would be more accurate and better in handling scale variation, while the "one-stage detection" model would provide a faster inference speed [74, 33,38]. Some common "two-stage detection" models were those from the region-based convolutional neural networks (RCNN) family, with the Faster RCNN [56] model being more popular due to its capability to provide a better balance between both accuracy and inference speed. Faster RCNN was applied in several studies in different industries to address real-time challenges with promising results [20,42,58]. For "one-stage detection", SSD [43] and YOLO [55] were two popular models that were widely used due to their fast inference speed.

While "anchor-based" models have been widely adopted, their reliance on pre-defined anchor boxes determined by a set of hyper-parameter, e.g., scales and aspect ratios, has limited their ability to address scale variation problems in object detection. In recent years, many studies using "anchor free" models have started to emerge [41,64,73,75] in the hope to address this problem. For example, Zhou et al. [75] proposed an anchorless-based approach, CenterNet, which has proven to be simpler, faster, and more accurate than corresponding bounding box-based approaches. To achieve higher accuracy, CenterNet is adopted in this study to detect missing barricades as it has outperformed current state-of-the-art object detection on the COCO database [73,75].

Background subtraction is widely used in computer vision applications involving video taken by fixed cameras [11] (e.g., traffic monitoring and industrial machine vision). Background subtraction can segment static and moving foreground objects in a video stream [5]. CNN was employed successfully for background initialization, foreground detection, and developing deep learned features. For example, Babaee et al. [4] proposed a novel deep learning method for background subtraction from video sequences. Likewise, Kim et al. [39] proposed a hybrid framework by integrating background subtraction and deep learning for person detection. In this work, background subtraction was applied to find the region of interest.

Many segmentation methods such as Mask R-CNN and U-Net can be used to segment static and moving foreground objects from a video stream [3,10]. Compared with other segmentation methods, U-Net has the following advantages: (1) the model allows for the use of global location and context at the same time; (2) it works with very few training samples and more accurate results for segmentation tasks [57]. Thus, U-Net is adopted as the backbone for our MCA method to segment barricades from video streams.

3. System overview

To prevent workers from falling off open edges at the construction level of a building, it is important to be able to detect missing barricades so that workers working near the open edges can be monitored. As shown in Fig. 1, the real-time video streaming from a surveillance camera mounted on tower cranes will be fed into the vision-based detection module. Then, the vision-based detection module will detect missing barricades if the barricades are removed from the building edges. Once the missing barricade is detected and a worker is detected in the hazardous area near the open edge, the system can alert site personnel of the unsafe activity and capture the incident as a statistic.

If the missing barricade is detected and a worker is detected in the hazardous area, then a "worker near open edges" event will be detected as an unsafe behavior. In this instance, a warning alert will be generated and sent to site personnel (e.g., site managers, supervisors, and workers) using Telegram[1] so that site personnel can take action immediately to prevent an FFH accident. Furthermore, statistics about workers near open edges can be collected to help managers assess the effectiveness of their safety management interventions.

4. Vision-based missing barricade detection

In this study, two approaches or strategies for detecting missing barricades are developed and implemented. These approaches include (1) *masks comparison approach* (MCA); and (2) *missing object detection approach* (MODA). The approaches were developed to solve the missing barricade problem faced by the developer and contractor that the authors were collaborating with in this study. The collaborators described their work processes and specified their requirements for the computer vision system. Based on their requirements, MCA was first developed and implemented. MODA was subsequently developed and implemented to solve the problems identified when implementing MCA. The details of these two approaches are described as follows.

4.1. Masks comparison approach (MCA)

Based on the inputs from the industry collaborators, the installation of barricades is part of the construction process, and it is a requirement for all barricades to be installed before workers are allowed to work on the construction level without a fall protection system. This requirement is aligned with safety regulations which require all open edges to be barricaded [47,68]. In fact, barricades are installed to the formwork that is being lifted to the next level to ensure that the barricades are present once construction of the next level starts. Inspection of the barricades at the construction level is one of the required tasks for work to proceed on the construction level, but the inspection is only at a specific instance in time. Site personnel face problems in ensuring that the barricades are not removed in an unplanned and unsafe manner after construction activities start on the construction level. There had been instances of workers removing barricades prematurely out of convenience. Besides, as part of the construction process, removal of the barricades is necessary, e.g., during the installation of precast walls. During these hazardous instances, the site supervisor needs to be alerted to ascertain that the workers have put on their fall protection system prior to barricade removal. Based on the work process, where the barricades are installed at the beginning of the construction cycle for each level, the MCA approach detects missing barricades by comparing the mask at

[1] Telegram is a free chat and instant messaging service that is available across different platforms, https://telegram.en.softonic.com/

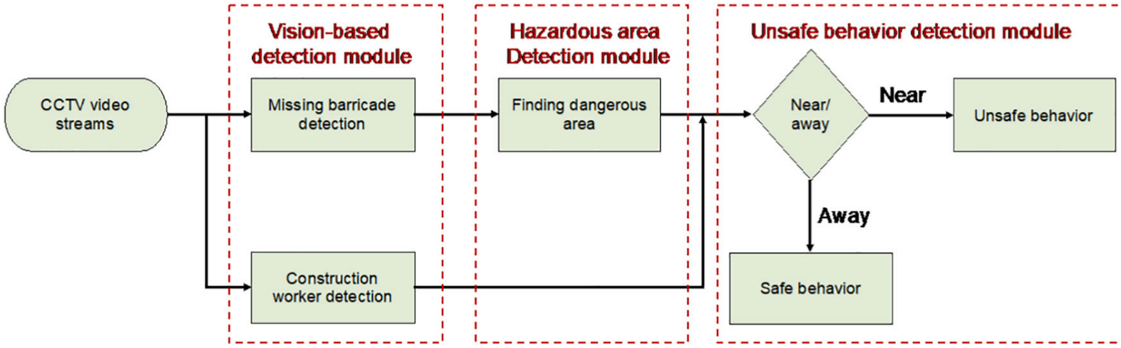

Fig. 1. Workflow of research approach.

consecutive timestamps t and $t+1$. Therefore, a key assumption of MCA is that barricades are installed at the beginning of the construction cycle for each level.

The basic concept underlying MCA is to use barricade present in past frames as a reference to identify missing barricades in the current frame. More specifically, during the training stage, a barricade segmentation model was constructed using a training image dataset, which consists of images with annotated barricade masks. The masks are used to indicate each pixel location as part of a barricade or not. In the inference stage, it first applies the trained barricade segmentation model to detect the presence of barricades in each frame and generate the barricade mask. Then, it compares the barricade mask between consecutive frames to identify whether there are missing barricades.

To ensure that the reference barricade mask continues to be relevant, it needs to be updated as construction work progresses. The updating of the reference mask is implemented using an 'exponential smoothing function'. In doing so, exponential smoothing is used to update the previous mask. The workflow of MCA is described as follows and illustrated in Fig. 2.

Step 1: For the first frame at the time index t, segment barricade mask as m_t.

Step 2: For the next frames starting from the time index $t+1$, segment barricade mask as m_{t+1}.

Step 3: Subtract the barricade mask at the current frame m_{t+1} from that in the previous frame m_t to get missing barricade mask, $m_l = m_t - m_{t+1}$.

Step 4: Update the previous barricade mask using an exponential smoothing function as $m_t = 0.9 * m_t + 0.1 * m_{t+1}$.

Step 5: Find the coordinate of the missing barricade mask (x_{max}, y_{max}, x_{min}, y_{max}), m_l as the bounding box of the missing barricade.

Step 6: Repeat Steps 2–5 for the rest frames in the video streaming.

Fig. 2 presents the workflow of our MCA method. We segment the barricade mask in the time indices t (Fig. 2(a)) and $t+1$ (Fig. 2(b)), respectively. Then, we subtract the barricade mask at the time index $t+1$. Thus, we can get the missing barricade mask, as noted in Fig. 3(c). An example is presented to illustrate the workflow, as illustrated in Fig. 3.

As noted above, U-Net, one of the most popular object segmentation approaches, can achieve more accurate results with fewer training images for the task of object segmentation. Thus, U-Net is used in this study to segment the barricades in the video streams. U-Net, an end-to-end fully convolutional network, was proposed by Ronneberger et al. [57], which contains two key parts. The first part is a downsampling strategy with the max-pooling operator upon receiving the input image and is referred to as the encoder or the contraction path. The second part is an asymmetric expansion of the feature map resulted from the first part with the upsampling operator and is referred to as the decoder or the expansion path. The model was trained with an image dataset with a resolution of height 320 pixels and width 576 pixels, 100 epochs with 200 steps per epoch using the Adam optimizer at a learning rate of 0.001. More details on U-Net can be found in Ronneberger et al. [57].

4.2. Missing object detection approach (MODA)

MODA was developed after identifying that the main cause of false positive detections of MCA is the excessive shaking of the CCTV cameras on the tower crane, which will be discussed in more detail in the case study. Furthermore, from a practical standpoint, the labeling of segmentation data is expensive and resource-intensive, making it difficult to increase the size of the annotated dataset to improve the accuracy of the MCA model. Thus, MODA was developed based on the requirement to overcome the problem of detection errors arising from camera shaking and to facilitate the accumulation of training dataset. At the same time, MODA removes the need to assume that the barricades are installed at the beginning of the construction cycle for each level, which is the key assumption of MCA.

The key idea of the MODA method is to treat the 'missing barricade' as a type of object and exploit deep learning methods to directly detect the object in the image. More specifically, in the training stage, the approach requires a training image dataset consisting of images with annotated missing barricade regions (via four corners of each region). In the inference stage, it detects the "missing object" in each frame, by identifying four corners of the missing barricade region. Fig. 4 presents the workflow of our MCA method.

A key point detection approach, CenterNet, is used to detect missing barricades at the construction level of the building in our case study. CenterNet is an anchor-free and key point-based detection model, which outperformed current state-of-the-art object detection approaches when tested on the MS COCO database [73,75]. It detects an object first as a center point and then regresses the object bounding box's height and width with respect to the center point. The center point is used to predict other object properties, such as object height and width. The model feeds the input image to a fully convolutional network that generates a heatmap, which is further used to obtain three outputs. First, peaks in this heatmap are used as predicted object centers. Second, image features obtained at each peak are used to predict the objects bounding box height and width. Third, offset values are predicted to refine the

Fig. 2. Workflow of MCA.

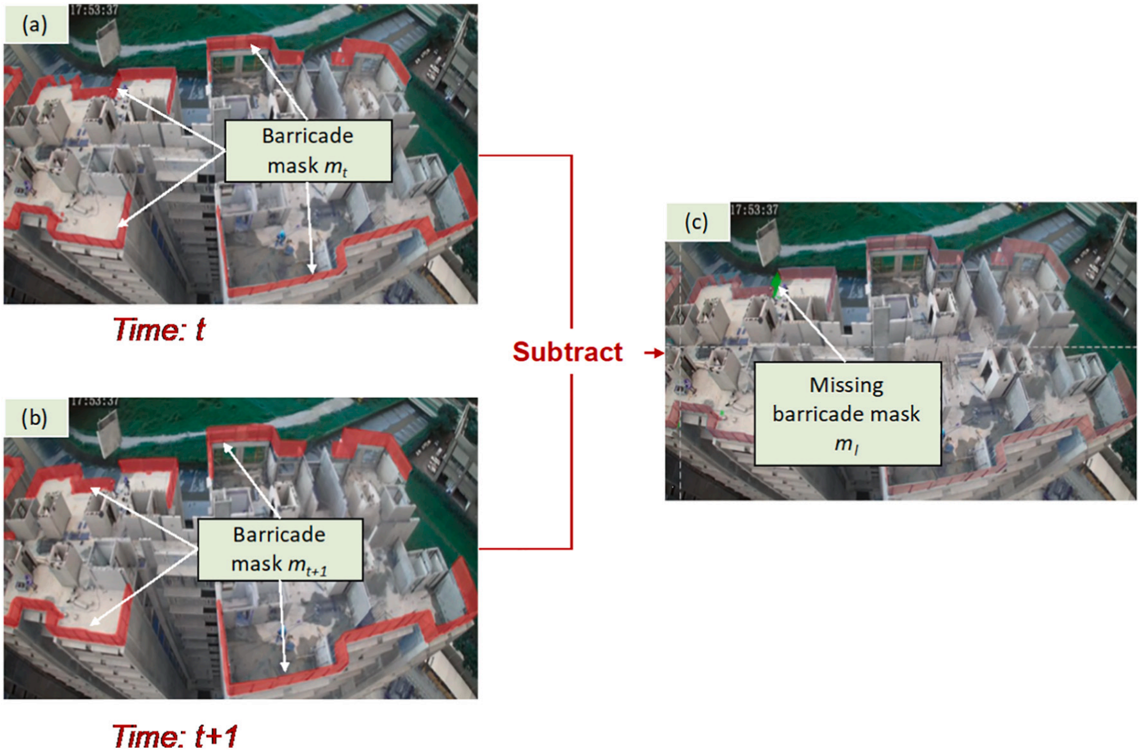

Fig. 3. An example of the workflow of MCA.

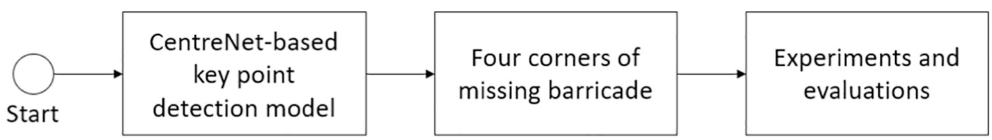

Fig. 4. Workflow of MODA.

predicted centers locations, which are incurred due to the down-sampling process in the model. In this study, Deep Layer Aggregation-34 (DLA-34) is used as the backbone network for the fully convolutional network to generate the heatmap. To extract the peak of each heatmap, all responses with a value higher or equal to its 8-connected neighbors are detected, and the top 100 peaks are retained. More details about the approach for peak extraction can be found in Zhou et al. [73,75].

To detect missing barricades using CenterNet, we considered each barricade as a set of four corner keypoints, each of which is parameterized by an offset to the center point. To refine the keypoints, we further estimate the barricade's heatmap using standard bottom-up pose estimation [6]. The CenterNet model is trained by exploiting three types of losses, loss of focal, loss of center offset, and loss of bounding box size, which are defined as follows.

Loss of focal is calculated using on Eq. (1).

$$L_k = \frac{-1}{N}\sum_{xyc}\begin{cases}\left(1-\widehat{Y}_{xyc}\right)^{\alpha} log\left(\widehat{Y}_{xyc}\right) & if\ Y_{xyc}=1\\ \left(1-Y_{xyc}\right)^{\beta}\left(\widehat{Y}_{xyc}\right)^{\beta} & otherwise\\ log\left(1-\widehat{Y}_{xyc}\right) & otherwise\end{cases} \tag{1}$$

where, α *and* β are hyper-parameters of the focal loss, which are based on Zhou et al. [73,75].

The loss of center offset L_{off} is noted in Eq. (2).

$$L_{off} = \frac{1}{N}\sum_{p}\left|\widehat{O}_{\tilde{p}} - \left(\frac{p}{R} - \tilde{p}\right)\right| \tag{2}$$

The loss of bounding box size is noted in Eq. (3).

$$L_{size} = \frac{1}{N}\sum_{k=1}^{N}\left|\widehat{S}_{pk} - s_k\right| \tag{3}$$

5. Case study

5.1. Background

To evaluate the feasibility and effectiveness of the approach, a public housing construction site in Singapore was used as a case study. In this case study, we collected data from two specific CCTV cameras covering one 20-story residential block under construction. The construction used precast and cast-in-situ technology. The approximate distance between the camera and the construction floor varies between 6.6 m to 17.8 m as the tower crane is jacked up when construction progresses, as illustrated in Fig. 5. Therefore, the size of barricades in the video footage will change when the distance between the camera and the construction floor varies, i.e., a multi-scale problem exists. Two sample snapshots of the CCTV images are presented in Fig. 6. Data was collected over a period of five months, including March, July, August, September, and October 2020. The break in the collection of data in April, May and June 2020 is due to the stop-work order arising from the COVID-19 pandemic. March, July, August, and September 2020 were set as the training

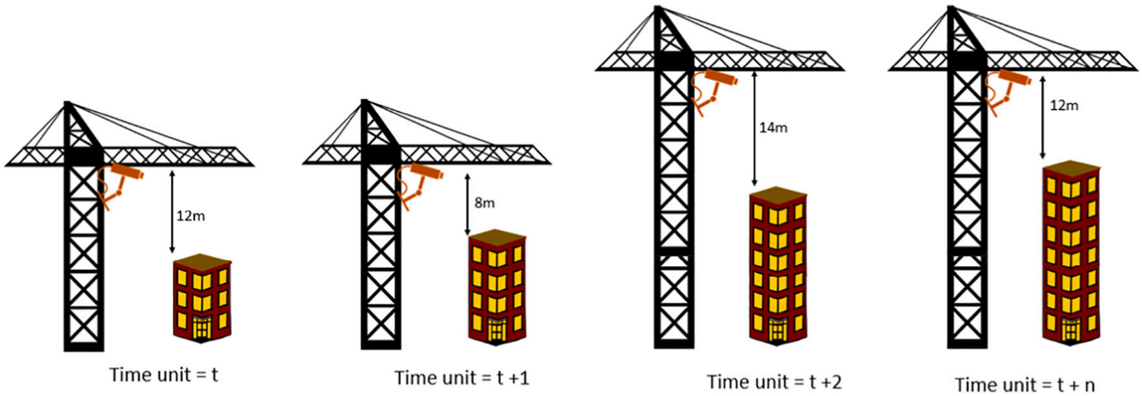

Fig. 5. Site tower crane jack-up cycle.

Fig. 6. Snapshots of CCTV images.

period, and data from this period were used to train models. Videos from October 2020 were used as the test data, and the data from this period was unseen by trained models. Both MCA and MODA methods were implemented on a server equipped with Intel i7 9th Generation CPU Computer with Nvidia GeForce RTX 2070 graphics card.

5.2. Evaluation performance metrics

To evaluate the performance of our proposed approach in this case study, three common criteria are used, which are Average Precision (AP), Average Recall (AR), and detection speed frames per Second (fps) [30,31,61]. AP measures the area under the Precision-Recall curve, and AR measures the area under Recall-Intersection over Union (IOU) threshold curve where the y-axis is Recall and x-axis is IOU threshold. In this study, the IOU is set to 0.5 as used COCO dataset. As illustrated in Fig. 7, if the IOU between the ground-truth bounding box and the predicted bounding box is greater than 0.5, then the prediction bounding box is identified as a true positive. Otherwise, it will be identified as a false positive. AP with an IOU threshold of 0.5 is abbreviated as $AP^{0.5}$ and AR with an IOU threshold of 0.5 is abbreviated as $AR^{0.5}$ in our following results.

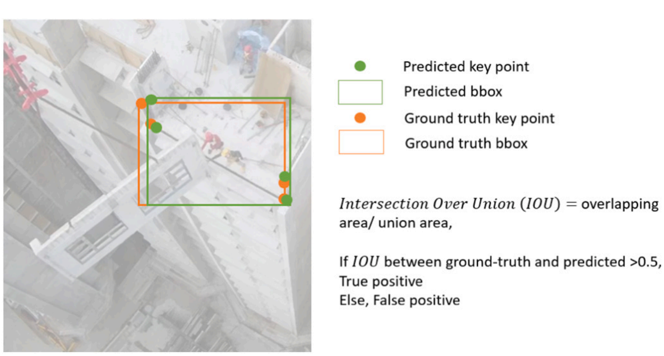

Fig. 7. Rules for evaluation of prediction results.

5.3. Solution #1: MCA for missing barricade detection

5.3.1. Data annotation

To validate the effectiveness and feasibility of the MCA, an image dataset is needed for model training and testing. To reduce the potential bias in selecting the images, the images in the image dataset were obtained from different viewpoints, at varying scales, poses, and illumination. Fig. 8 presents examples from our image dataset, including missing barricades and barricades. Then, the created image dataset was randomly divided into a training dataset and a testing dataset. In this study, our created database has 853 images, out of which 764 images were used for training, and 89 images were used for testing. Prior to testing, the created training dataset was labeled to annotate the barricade in the image using an annotation tool (e.g., 'LabelImg'). An example of image annotation is presented in Fig. 9.

5.3.2. Detection result

Table 2 presents the detection results of MCA. Dice coefficient (DC) is a metric to evaluate the performance of the object segmentation model, and it is measured by the area of overlap divided by the total number of pixels in both the ground truth and the predicted label [27,59]. DC can be computed using the following equation,

$$D_C = \frac{2N_0}{N_R + N_I} \tag{4}$$

whereN_0 is the number of overlap pixels; N_R and N_I are the numbers of pixels of the ground-truth and the predicted label, respectively.

From Table 2, we can conclude that the barricade can be accurately detected, but the missing barricades from video streaming are not able to be detected by MCA. In addition, the detection speed of MCA is 4.25 fps, as it needs more computing resources to extract features, which may not meet the requirement of real-time detection in construction sites via

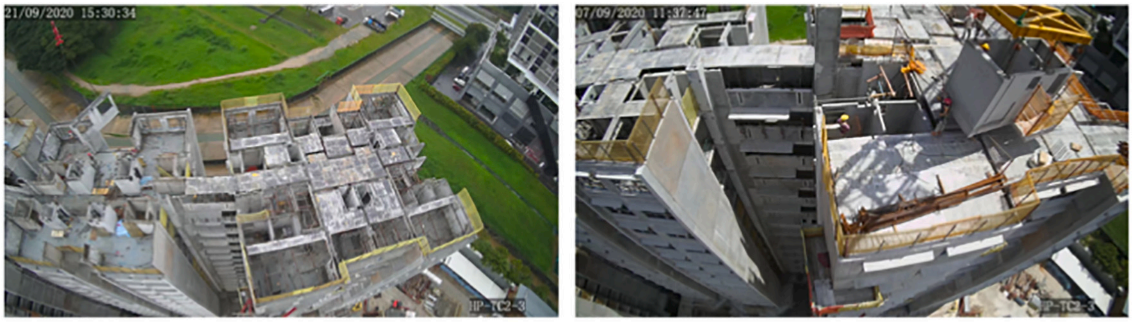

Fig. 8. Examples of images in the image dataset.

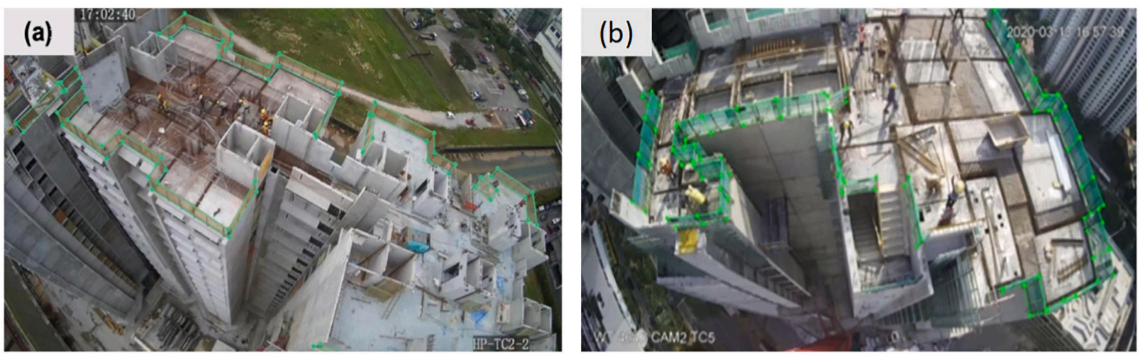

Fig. 9. Example of image annotation, where the green color indicates the annotated barricade in the image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Detection results using the MCA.

Performance metrics	$AP^{0.5}$	$AR^{0.5}$	Dice coefficient	Detection speed (fps)
Barricade Segmentation	–	–	0.81	–
Missing barricade detection	0.20	0.60	–	4.25

video streaming. Fig. 10 presents examples of true positive detection and false positive detection results.

As highlighted earlier, after verifying the detection results, we found that the key reason affecting the accuracy of MCA is “Camera shaking”. An example of the false positive detection due to the shaking of the camera is presented in Fig. 11. Tower cranes have the tendency to vibrate when it is lifting a load. When it vibrates, the camera mounted on top of the crane shakes. In our case study, the camera shakes during almost every lift, but the severity of the problem varies. When the camera is shaking, MCA will cause a false positive to be detected as the originally detected barricades will move out of positions when the camera shakes. When the detected barricades move out of positions, the previously registered barricade will be deemed as not present and will cause a missing barricade to be detected, i.e., false positive will occur.

Another practical challenge of the MCA method is that it is resource intensive and time-consuming to label segmentation data since the MCA method requires dense pixel-level annotations [72]. Therefore, it is not feasible to label a large amount of training data for the MCA model, which influences its detection accuracy due to insufficient annotated data for the model training.

5.4. Solution #2: MODA for missing barricade detection

To address the key problem of “camera shaking” and overcome the difficulties in accumulating labeled datasets for training, we developed MODA and implemented it in our case study.

5.4.1. Data annotation

A dataset with 1689 images was used to train and test the MODA model. More specifically, 1560 images were used for training, and 129 images were used for testing. In this study, four key points are used to demarcate a missing barricade at its four corners. As four corners points of every missing barricade may not present themselves clearly, there needs to be a standardized way of labeling thee missing barricade four corner points. The following two rules are considered during data annotation:

- Cut back the corner point if the corner point of the missing barricade is occluded by an opaque object, as shown in Fig. 12(a), Fig. 12(b) and Fig. 12(c).
- Retain the corner point of the missing barricade if the corner point is occluded by another translucent barricade.

Fig. 12 shows how the missing barricade was annotated with two general rules.

5.4.2. Detection result

Table 3 presents the detection results for missing barricades at the construction level using these two approaches. Based on the results, we can conclude that MODA can achieve a higher accuracy in the detection of missing barricade in images. In addition, the MODA has a higher detection speed compared with MCA.

Figs. 13 and 14 present examples of correct detection and error detection of MODA, respectively. Here, the blue bounding box highlighted in Fig. 13 is a danger area (near the open edge) when the barricade was missed. We found that the main reason for error detection is that the color of the barricade is similar to the background. However, the issue of “camera shaking” had been minimized.

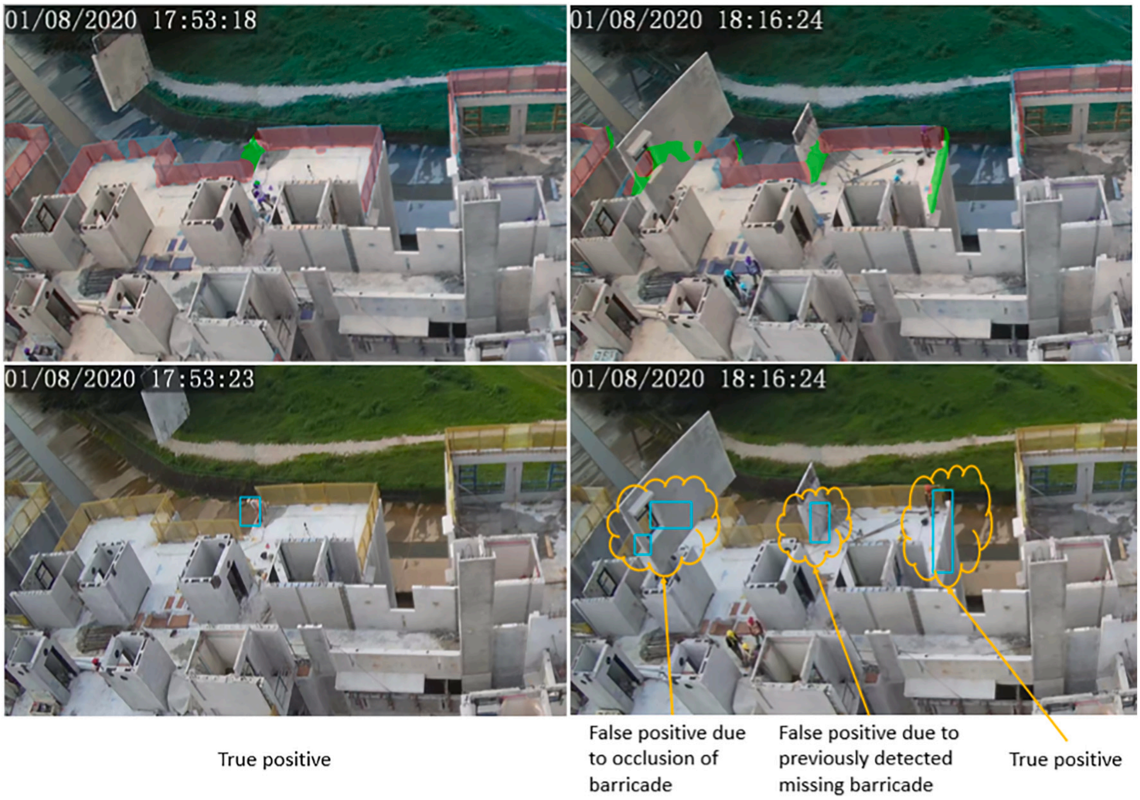

Fig. 10. Example of detection results: True positive detection (*left*); True and False positive detections (*right*).

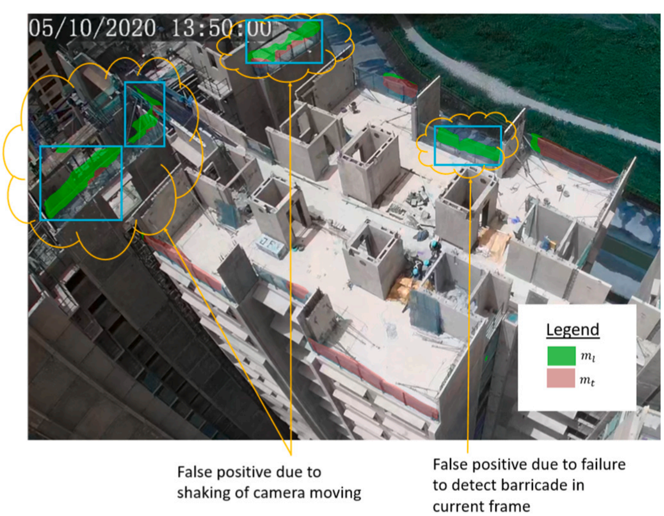

Fig. 11. Example of false detection due to camera shaking.

6. Discussion, limitations and future works

In relation to the first research question on possible computer vision-based approaches for detection of missing barricades, this study identified and evaluated two computer-vision approaches (MCA and MODA) to detect missing barricades. The approaches aim to provide site management with an automatic mechanism to proactively identify missing barricades and take immediate action to mitigate the likelihood of FFH. The automatic nature of the approaches improves the efficiency and effectiveness of the safety inspection and monitoring processes to prevent missing barricades and hence reduce the risk of FFH incidents. In addition, the system can send warning alerts to the site supervisor and relevant workers. In this way, the approaches can be used by site management to highlight potential hazards to relevant site personnel. The tight monitoring by the system will encourage better safety compliance.

The second research question is about the accuracy of the identified approaches. The evaluation in this study shows that the $AP^{0.5}$ and $AR^{0.5}$ of MCA were 0.20 and 0.60, respectively. The $AP^{0.5}$ and $AR^{0.5}$ of MCA were 0.579 and 0.736, respectively. Thus, MODA was more accurate in the context of this study.

The third research question is about the feasibility of the two approaches. The study uncovered several practical challenges in implementing the approaches, especially MCA. MODA can address the limitations of camera sharking on construction sites and the experimental results show that MODA has better performance than MCA. However, MCA helps detect fine-grained missing objects when the camera is stationary because it is a pixel level-based object detection approach, especially when the missing object is not easily labeled using the bounding box. For example, the components of scaffolding are missing during operation, MCA can detect this issue more easily by computing the difference between the two consecutive frames. MODA is useful in detection of missing object when the camera is moving as this is a key point-based object detection approach. For example, if we use unmanned aerial vehicle (UAV) for safety inspection, we can use MODA to process the UAV videos to detect missing objects when the missing object is easily labeled using a bounding box. MCA, though involves a more intensive labelling process, is more generalizable to new sites as the model remembers how a barricade looks like and could be used to find barricade on a new site if the barricade used on the new site is similar. However, since MODA remembers the environment when a barricade is missing, and construction sites environment is dynamic, MODA is probably less generalizable to new sites. Thus, both MODA and MCA have their pros and cons, but based on the results of this study, MODA appears to be more feasible for construction sites that are similar to the site in this study.

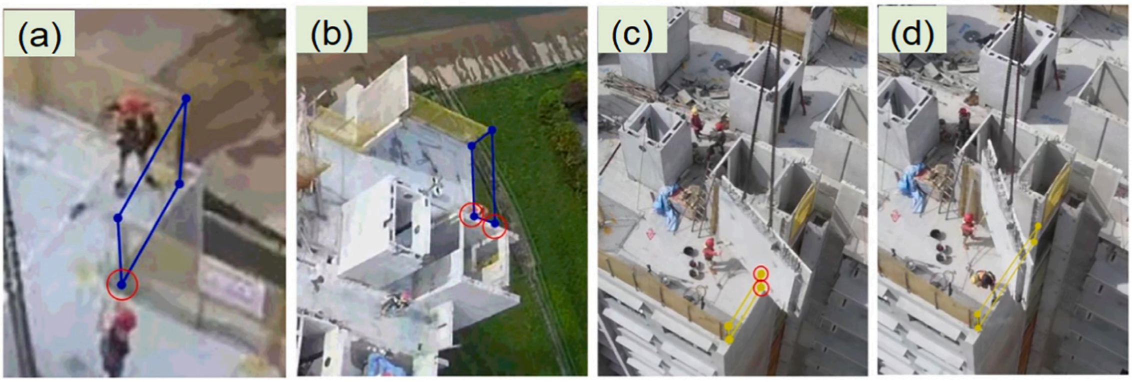

Fig. 12. Example of annotations of the missing barricade: (a) missing barricade with one corner occluded by another transparent barricade; (b) missing barricade with two corner points occluded by an opaque object; (c) missing barricade with two points occluded by hanging load; (d) missing barricade with no points occluded.

Table 3
Missing barricade detection comparison between the MCA method and the MODA method.

Method performance metrics	Masks comparison approach (MCA)	Missing object detection approach (MODA)
Average Precision (AP)	0.200	**0.579**
Average Recall (AR)	0.600	**0.736**
Detection speed (fps)	4.25	**12.50**

The contributions of this study are as follows. Firstly, despite the perennial problem of missing barricades, there has been an absence of study that uses computer vision approaches to detect missing barricades automatically for safety compliance checks. To address this challenge, this study has demonstrated that the use of computer vision, especially MODA, can accurately identify missing barricades so that workers near open edge can be accurately detected. Our proposed approach provides an automated means to monitor and identify worker's unsafe behavior for prevention of FFH accidents, which will also help produce more consistent and comprehensive behavior data to promote behavior-based safety (BBS) in construction. Secondly, as evaluated in our case study, the MODA method can better detect object with varying size in images. In addition, the MODA method can address the limitations of camera sharking on construction sites. The experimental results shown in Table 3 have demonstrated that MODA has better performance than MCA. Therefore, the MODA is more feasible to be implemented on construction sites and achieve a higher accuracy for detecting missing barricades. However, the two approaches may perform differently in other construction scenarios and further evaluation is necessary.

While this study contributes to identifying missing barricades on construction sites, we need to highlight the following limitations, which will be addressed in our future works. Firstly, our model is tested in limited construction scenarios, the generalization of our model still needs to be improved further. In our study, we collect training and testing dataset from two specific CCTV cameras covering one 20-story residential block under construction. In this instance, two limited construction scenarios were used as testing conditions. In our future works, a larger image dataset from different construction projects will be created and used for model training and testing. Secondly, our model is

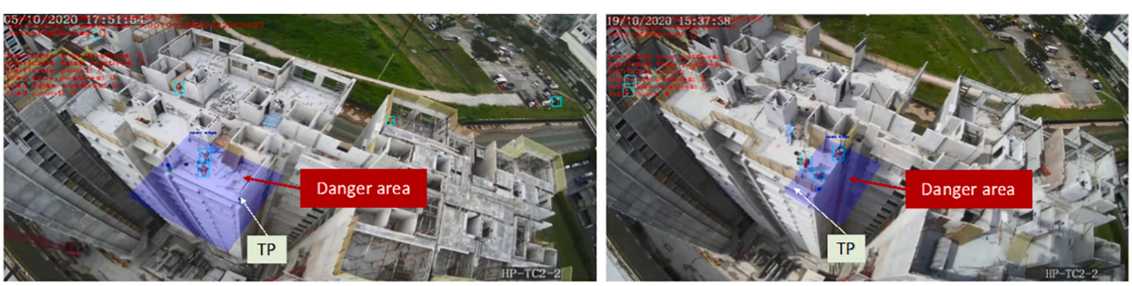

Fig. 13. Examples of correct detection of missing barricade (TP is True Positive).

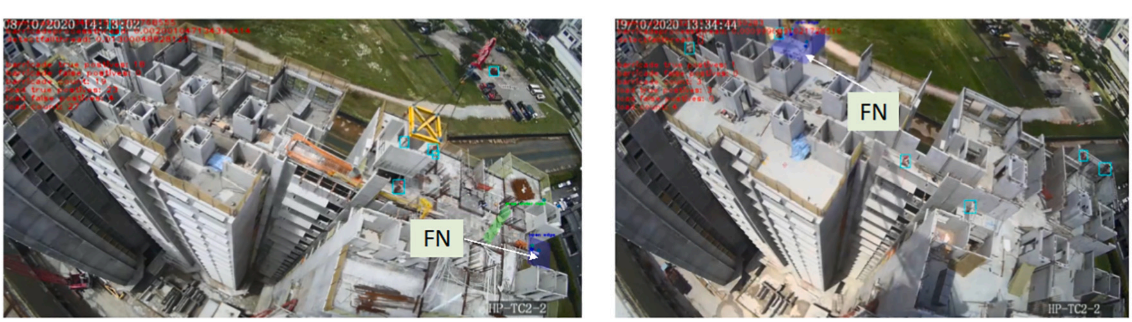

Fig. 14. Example of error detection of missing barricade (FN is False Negative).

trained with a relatively small database and the accuracy of the models can still be improved. However, when compared with other recent relevant studies, e.g., Li et al. [31] and Hou et al. (2020), the $AP^{0.5}$ and $AR^{0.5}$ of 57.9% and 73.6%, respectively, are respectable. In Li et al. [31], an improved YOLOv3 is proposed for rebar counting with $AP^{0.5}$ of 61.8%. Likewise, in Hou et al. (2020), an improved Mask R-CNN is proposed for GPR signature detection and segmentation. The detection AP and segmentation AP were 58.64% and 47.64%, respectively. Nevertheless, we believe that the accuracy of our two proposed approaches can be improved with the help of a larger image dataset.

7. Conclusion

This study proposes two computer vision approaches including the masks comparison approach (MCA), and the missing object detection approach (MODA) for real-time detection of missing barricades at construction sites. A public housing construction project in Singapore was used to assess the effectiveness and feasibility of the proposed approaches. We initially implemented MCA, which is based on a temporal comparison of barricade segmentation. However, we found that the approach does not perform well when the camera on the tower crane shakes. Furthermore, the segmentation approach used in MCA made it costly to accumulate a large amount of training and testing data. Thus, we developed MODA, which detects missing barricades as an object. The results presented in this study demonstrated that MODA can achieve better performance than MCA. The average precision and average recall of missing barricades using MODA were found to be 57.9% and 73.6%, respectively, which are comparable with other recent works on the use of computer vision in construction sites. We suggest that our MODA-based computer vision system can be used to supplement existing safety management measures to reduce the likelihood of FFH accidents. Moreover, the statistics collected from our proposed approach can be used to facilitate implementation of behavior-based safety management at construction sites.

Declaration of Competing Interest

None.

Acknowledgement

This research is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG-100E-2019-026) and the Singapore Housing and Development Board (HDB) (R-296-001-204-490).

References

[1] X. An, L. Zhou, Z. Liu, C. Wang, P. Li, Z. Li, Dataset and benchmark for detecting moving objects in construction sites, Autom. Constr. 122 (2021) 103482, https://doi.org/10.1016/j.autcon.2020.103482.
[2] Q. Ang, MOM to Carry Out 300 More Checks in High-Risk Workplaces after Deadly Accidents in Feb. https://www.straitstimes.com/singapore/300-more-inspections-for-high-risk-workplaces-after-seven-deaths-from-workplace-accidents, 2021.
[3] O. Angah, A.Y. Chen, Removal of occluding construction workers in job site image data using U-Net based context encoders, Autom. Constr. 119 (2020) 103332, https://doi.org/10.1016/j.autcon.2020.103332.
[4] M. Babaee, D.T. Dinh, G. Rigoll, A deep convolutional neural network for video sequence background subtraction, Pattern Recogn. 76 (2018) 635–649, https://doi.org/10.1016/j.patcog.2017.09.040.
[5] T. Bouwmans, S. Javed, M. Sultana, S.K. Jung, Deep neural network concepts for background subtraction: a systematic review and comparative evaluation, Neural Netw. 117 (2019) 8–66, https://doi.org/10.1016/j.neunet.2019.04.024.
[6] Z. Cao, T. Simon, S.E. Wei, Y. Sheikh, Realtime multi-person 2D pose estimation using part affinity fields, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7291–7299, https://doi.org/10.1109/cvpr.2017.143.
[7] E. Cheung, A.P. Chan, Rapid demountable platform (RDP)—a device for preventing fall from height accidents, Accid. Anal. Prev. 48 (2012) 235–245, https://doi.org/10.1016/j.aap.2011.05.037.
[9] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, Q. Tian, Centernet: Keypoint triplets for object detection, in: In: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 6569–6578, https://doi.org/10.1109/iccv.2019.00667.
[10] E. Garilli, N. Bruno, F. Autelitano, R. Roncella, F. Giuliani, Automatic detection of stone pavement's pattern based on UAV photogrammetry, Autom. Constr. 122 (2021) 103477, https://doi.org/10.1016/j.autcon.2020.103477.
[11] B. Garcia-Garcia, T. Bouwmans, A.J.R. Silva, Background subtraction in real applications: challenges, current models and future directions, Comp. Sci. Rev. 35 (2020) 100204, https://doi.org/10.1016/j.cosrev.2019.100204.
[12] W. Fang, L. Ding, H. Luo, P.E. 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.
[13] W. Fang, L. Ding, B. Zhong, P.E. Love, H. Luo, Automated detection of workers and heavy equipment on construction sites: a convolutional neural network approach, Adv. Eng. Inform. 37 (2018) 139–149, https://doi.org/10.1016/j.aei.2018.05.003.
[14] W. Fang, L. Ding, P.E. Love, H. Luo, H. Li, F. Pena-Mora, B. Zhong, C. Zhou, Computer vision applications in construction safety assurance, Autom. Constr. 110 (2020) 103013, https://doi.org/10.1016/j.autcon.2019.103013.
[15] W. Fang, P.E. Love, H. Luo, L. Ding, Computer vision for behaviour-based safety in construction: a review and future directions, Adv. Eng. Inform. 43 (2020) 100980, https://doi.org/10.1016/j.aei.2019.100980.
[17] W. Fang, B. Zhong, N. Zhao, P.E. 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. Inform. 39 (2019) 170–177, https://doi.org/10.1016/j.aei.2018.12.005.
[18] P. Felzenszwalb, D. McAllester, D. Ramanan, A discriminatively trained, multiscale, deformable part model, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2008, June, pp. 1–8, https://doi.org/10.1109/cvpr.2008.4587597.
[20] A. Fuentes, S. Yoon, S.C. Kim, D.S. Park, A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition, Sensors 17 (9) (2017) 2022, https://doi.org/10.3390/s17092022.
[21] Y.M. Goh, P.E. Love, Adequacy of personal fall arrest energy absorbers in relation to heavy workers, Saf. Sci. 48 (6) (2010) 747–754, https://doi.org/10.1016/j.ssci.2010.02.020.
[22] 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.
[23] Y.M. Goh, Introduction to Workplace Safety and Health Management: A Systems Thinking Approach, Second edn, World Scientific, Singapore, 2020. ISBN: 978-981-122-499-7.
[24] B.H. Guo, Y.M. Goh, Ontology for design of active fall protection systems, Autom. Constr. 82 (2017) 138–153, https://doi.org/10.1016/j.autcon.2017.02.009.
[25] B.H.W. Guo, Y. Zou, Y. Fang, Y.M. Goh, P.X.W. Zou, Computer vision technologies for safety science and management in construction: a critical review and future research directions, Saf. Sci. 135 (2021) 105130, https://doi.org/10.1016/j.ssci.2020.105130.
[26] Y.M. Goh, B.H. Guo, FPSWizard: a web-based CBR-RBR system for supporting the design of active fall protection systems, Autom. Constr. 85 (2018) 40–50, https://doi.org/10.1016/j.autcon.2017.09.020.
[27] B. Guindon, Y. Zhang, Application of the dice coefficient to accuracy assessment of object-based image classification, Can. J. Remote. Sens. 43 (1) (2017) 48–61, https://doi.org/10.1080/07038992.2017.1259557.
[28] F. Guo, Y. Qian, Y. Shi, Real-time railroad track components inspection based on the improved YOLOv4 framework, Autom. Constr. 125 (2021) 103596, https://doi.org/10.1016/j.autcon.2021.103596.
[30] F. Hou, W. Lei, S. Li, J. Xi, M. Xu, J. Luo, Improved mask R-CNN with distance guided intersection over union for GPR signature detection and segmentation, Autom. Constr. 121 (2021) 103414, https://doi.org/10.1016/j.autcon.2020.103414.
[31] Y. Li, Y. Lu, J. Chen, A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector, Autom. Constr. 124 (2021) 103602, https://doi.org/10.1016/j.autcon.2021.103602.
[32] W.C. Lim, S.M. Tashrif, Y.M. Goh, S.J. Adrian Koh, Validation of the energy balance approach for design of vertical lifeline systems, Int. J. Occup. Saf. Ergon. (2019) 1–13, https://doi.org/10.1080/10803548.2019.1616948.
[33] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikäinen, Deep learning for generic object detection: a survey, Int. J. Comput. Vis. 128 (2) (2020) 261–318, https://doi.org/10.1007/s11263-019-01247-4.
[34] X. Luo, H. Li, D. Cao, F. Dai, J. Seo, S. Lee, Recognizing diverse construction activities in site images via relevance networks of construction-related objects detected by convolutional neural networks, ASCE J. Comp. Civil Eng. 32 (3) (2018) 04018012, https://doi.org/10.1061/(asce)cp.1943-5487.0000756.
[35] X. Luo, H. Li, H. Wang, Z. Wu, F. Dai, D. Cao, Vision-based detection and visualization of dynamic workspaces, Autom. Constr. 104 (2019) 1–13, https://doi.org/10.1016/j.autcon.2019.04.001.
[36] 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) 103085, https://doi.org/10.1016/j.autcon.2020.103085.
[37] R. Navon, O. Kolton, Model for automated monitoring of fall hazards in building construction, ASCE J. Const. Eng. Manag. 132 (7) (2006) 733–740, https://doi.org/10.1061/(asce)0733-9364(2006)132:7(733).
[38] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, R. Qu, A survey of deep learning-based object detection, IEEE Access 7 (2019) 128837–128868, https://doi.org/10.1109/access.2019.2939201.

[39] C. Kim, J. Lee, T. Han, Y.M. Kim, A hybrid framework combining background subtraction and deep neural networks for rapid person detection, J. Big Data 5 (1) (2018) 1–24, https://doi.org/10.1186/s40537-018-0131-x.

[40] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: In: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105, https://doi.org/10.1145/3065386.

[41] H. Law, J. Deng, Cornernet: detecting objects as paired keypoints, in: In Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 734–750, https://doi.org/10.1007/s11263-019-01204-1.

[42] Z. Ling, R.C. Qiu, Z. Jin, Y. Zhang, X. He, H. Liu, L. Chu, An Accurate and Real-Time Self-Blast Glass Insulator Location Method Based on Faster R-CNN and U-Net with Aerial Images, 2018, pp. 1–8, arXiv preprint, arXiv:1801.05143.

[43] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, SSD: Single shot multibox detector, in: European Conference on Computer Cision, Springer, Cham, 2016, October, pp. 21–37, https://doi.org/10.1007/978-3-319-46448-0_2.

[44] H. Luo, J. Liu, W. Fang, P.E. Love, Q. Yu, Z. Lu, Real-time smart video surveillance to manage safety: a case study of a transport mega-project, Adv. Eng. Inform. 45 (2020) 101100, https://doi.org/10.1016/j.aei.2020.101100.

[45] T. Malisiewicz, A. Gupta, A.A. Efros, Ensemble of exemplar-SVMs for object detection and beyond, in: Proceedings of the International Conference on Computer Vision, IEEE, 2011, November, pp. 89–96, https://doi.org/10.1109/iccv.2011.6126229.

[46] J. Melzner, S. Zhang, J. Teizer, H.J. Bargstädt, A case study on automated safety compliance checking to assist fall protection design and planning in building information models, Constr. Manag. Econ. 31 (6) (2013) 661–674, https://doi.org/10.1080/01446193.2013.780662.

[47] Ministry of Manpower, Kingsford Construction Fined for Safety Lapses at Hillview Worksite. https://www.mom.gov.sg/newsroom/press-releases/2017/0803-kingsford-construction-fined-for-safety-lapses-at-hillview-worksite, 2017 (Accessed on June 11, 2021).

[49] K. Mizuno, Y. Terachi, K. Takagi, S. Izumi, H. Kawaguchi, M. Yoshimoto, Architectural study of HOG feature extraction processor for real-time object detection, in: Proceedings of the IEEE Workshop on Signal Processing Systems, 2012, October, pp. 197–202, https://doi.org/10.1109/sips.2012.57.

[50] E.A. 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 (7) (2016) 638, https://doi.org/10.3390/ijerph13070638.

[51] R. Navon, O. Kolton, Algorithms for automated monitoring and control of fall hazards, J. Comput. Civ. Eng. 21 (1) (2007) 21–28, https://doi.org/10.1061/(asce)0887-3801(2007)21:1(21).

[53] J. Qi, R.R. Issa, S. Olbina, J. Hinze, Use of building information modeling in design to prevent construction worker falls, ASCE J. Comp. Civil Eng. 28 (5) (2014) A4014008, https://doi.org/10.1061/(asce)cp.1943-5487.0000365.

[54] A.A. Raheem, J.W. Hinze, Disparity between construction safety standards: a global analysis, Saf. Sci. 70 (2014) 276–287, https://doi.org/10.1016/j.ssci.2014.06.012.

[55] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788, https://doi.org/10.1109/cvpr.2016.91.

[56] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: Advances in Neural Information Processing Systems, 2015, pp. 91–99, https://doi.org/10.1109/tpami.2016.2577031.

[57] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, 2015, October, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

[58] K. Saleh, M. Hossny, A. Hossny, S. Nahavandi, Cyclist detection in lidar scans using faster R-CNN and synthetic depth images, in: In: IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017, October, pp. 1–6, https://doi.org/10.1109/itsc.2017.8317599.

[59] R.R. Shamir, Y. Duchin, J. Kim, G. Sapiro, N. Harel, Continuous Dice Coefficient: A Method for Evaluating Probabilistic Segmentations, 2019, pp. 1–4, arXiv preprint, arXiv:1906.11031.

[60] H. Son, H. Choi, H. Seong, C. Kim, Detection of construction workers under varying poses and changing background in image sequences via very deep residual networks, Autom. Constr. 99 (2019) 27–38, https://doi.org/10.1016/j.autcon.2018.11.033.

[61] H. Son, C. Kim, Integrated worker detection and tracking for the safe operation of construction machinery, Autom. Constr. 126 (2021) 103670, https://doi.org/10.1016/j.autcon.2021.103670.

[62] S. Tang, D. Roberts, M. Golparvar-Fard, Human-object interaction recognition for automatic construction site safety inspection, Autom. Constr. 120 (2020) 103356, https://doi.org/10.1016/j.autcon.2020.103356.

[63] S.M. Tashrif, W.C. Lim, Y.M. Goh, X. Hu, S.J.A. Koh, Experimental validation of energy balance approach for design of horizontal lifeline systems, Int. J. Occup. Saf. Ergon. (2020) 1–43, https://doi.org/10.1080/10803548.2020.1763031.

[64] Z. Tian, C. Shen, H. Chen, T. He, FCOS: Fully convolutional one-stage object detection, in: In: Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 9627–9636, https://doi.org/10.1109/iccv.2019.00972.

[65] C. Vigneshkumar, U.R. Salve, A scientometric analysis and review of fall from height research in construction, Const. Econ. Build. 20 (1) (2020) 17, https://doi.org/10.5130/ajceb.v20i1.6802.

[66] X. Wang, T.X. Han, S. Yan, An HOG-LBP human detector with partial occlusion handling, in: Proceedings of the IEEE 12th International Conference on Computer Vision, IEEE, 2009, September, pp. 32–39, https://doi.org/10.1109/iccv.2009.5459207.

[67] J. Wang, N. Pradhananga, J. Teizer, Automatic fall risk identification using point cloud data in construction excavation, in: Computing in Civil and Building Engineering 2014, 2014, pp. 981–988, https://doi.org/10.1061/9780784413616.122.

[68] Workplace Safety and Health Council, WSHC. https://www.wshc.sg/files/wshc/upload/infostop/file/2013/WSH+Code+of+Practice+2013_ebook.pdf, 2013 (Accessed on June 11, 2021).

[70] T. Zlatar, E.M.G. Lago, W.D.A. Soares, J.D.S. Baptista, B. Barkokébas Junior, Falls from height: analysis of 114 cases, Production 29 (2019), https://doi.org/10.1590/0103-6513.20180091 pp. e20180091.

[71] S. Zhang, K. Sulankivi, M. Kiviniemi, I. Romo, C.M. Eastman, J. Teizer, BIM-based fall hazard identification and prevention in construction safety planning, Saf. Sci. 72 (2015) 31–45, https://doi.org/10.1016/j.ssci.2014.08.001.

[72] X. Zhao, R. Vemulapalli, P. Mansfield, B. Gong, B. Green, L. Shapira, Y. Wu, Contrastive Learning for Label-Efficient Semantic Segmentation, 2020, pp. 1–13, arXiv preprint, arXiv:2012.06985.

[73] X. Zhou, D. Wang, P. Krähenbühl, Objects as Points, 2019, pp. 1–12, arXiv preprint, arXiv:1904.07850.

[74] Z. Zou, Z. Shi, Y. Guo, J. Ye, Object detection in 20 years: A survey, 2019, pp. 1–39, arXiv preprint, arXiv:1905.05055.

[75] X. Zhou, J. Zhuo, P. Krahenbuhl, Bottom-up object detection by grouping extreme and center points, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 850–859, https://doi.org/10.1109/cvpr.2019.00094.

[76] C.M. Zuluaga, A. Albert, Preventing falls: choosing compatible fall protection supplementary devices (FPSD) for bridge maintenance work using virtual prototyping, Saf. Sci. 108 (2018) 238–247, https://doi.org/10.1016/j.ssci.2017.08.006.