Sensors and Materials, Vol. 36, No. 9 (2024) 4017–4028
MYU Tokyo

S & M 3781

Development of Segmentation Technology for Fall Risk Areas in Small-Scale Construction Sites Based on Bird's-eye-view Images

Na Jong-ho,[1] Lee Jae-kang,[2*] Shin Hyu-soung,[3] and Yun Il-dong[4]

[1]Department of Future & Smart Construction Research, Korea Institute of Civil Engineering and Building Technology, Ph.D. Candidate, Department of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies
[2]Department of Civil Engineering, Dong-A University
[3]Department of Future & Smart Construction Research, Korea Institute of Civil Engineering and Building Technology
[4]Department of Computer Engineering, Hankuk University of Foreign Studies

(Received August 23, 2024; accepted September 25, 2024)

Keywords: construction site, bird's eye view, process area classification, risk area classification, instance segmentation

Construction sites have shown the highest incidence of safety accidents across industries in recent times. Small-scale sites, in particular, often operate without on-site safety managers, leading to significant safety oversights. In this study, we developed a method of identifying risk areas during construction procedures by using bird's-eye-view image data throughout the construction cycle. Actual construction site images were collected and specific target objects were selected to create an AI training dataset. The segmentation model's performance was validated, and a system was developed to identify fall risk areas by establishing interconnections between these target objects within the images. The findings of this study can help enhance compliance assessment with construction procedures and improve safety management oversight at small-scale construction sites.

1. Introduction

Recently, the government of South Korea has intensified efforts to reduce the number of fatalities and safety accidents in industrial sites through legislation such as the Occupational Safety and Health Act and the Serious Accidents Punishment Act. Despite these measures, the construction sector reported the highest number of fatalities in 2021, with 417 deaths, comprising 50.5% of all industrial accidents.(1) While large companies deploy on-site safety managers who independently oversee safety tasks, small-scale construction sites with budgets below 5 billion KRW are not obligated to do so, resulting in significant safety oversights. Statistical analysis from the Construction Safety Management Integrated Information Network reveals that such small-scale construction sites exhibit the highest accident rate, predominantly due to falls. This

*Corresponding author: e-mail: jaekanglee@dau.ac.kr
https://doi.org/10.18494/SAM5334

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

underscores the urgent need for technical solutions to enhance safety management in these overlooked areas of the construction industry.

Therefore, this study focused on detecting process areas and fall risk zones using closed-circuit television (CCTV) footage, with a specific emphasis on small-scale construction sites where fall accidents are prevalent. Unlike in previous studies where data were typically analyzed using CCTV installed on-site, in this study, a novel approach that utilizes bird's-eye view images was considered. To gather real-world data, a small-scale construction site was selected and a bird's eye-view image collection system was deployed onsite. Data spanning the entire construction cycle were collected and time-series analysis was conducted to capture images of construction areas at various stages. The key process areas and elements associated with fall risks were identified to form datasets, and a real-time fall risk detection system was implemented using a segmentation model. The reliability of the system was validated through evaluations of object detection rates and segmentation performance using experimental data.

In small-scale construction sites, construction progresses rapidly with multiple concurrent processes, leading to dynamic changes in fall risk zones and heightened worker accident risks. In this study, we propose a distinctive data analysis method, providing an easy way to detect fall risk areas at different floor levels without the need for additional sensors (e.g., distance sensors). Implementing this system is expected to reduce accident rates at construction sites and enhance compliance with safety protocols.

2. Related Research

Traditional computer vision theories have extensively been used in research on object recognition technologies.[2] These methods relied on predefined rules for detection, which proved challenging to define comprehensively, limiting their performance. However, recent advancements in deep learning-based object recognition have substantially enhanced detection capabilities, even in challenging conditions with poor data quality. Consequently, deep learning technologies are now applied to enhance CCTV-based safety monitoring in construction sites, which often face varying light, weather, and dust conditions.

For instance, Xiang *et al.*[3] utilized the Faster Region-based Convolutional Neural Network (R-CNN) method to monitor large-scale construction sites and detect intruding vehicles. Similarly, Guo *et al.*[4] employed CNN's multi-level features to detect densely packed vehicles in small-scale construction sites, introducing an orientation-aware bounding box technique for enhanced vehicle identification in crowded areas. Yang *et al.*[5] applied the Mask R-CNN method to identify crane operators and hazardous zones within construction sites, while Fang *et al.*[6] used Mask R-CNN to classify various construction-related objects. Chen *et al.*[7] utilized three CNNs to track and analyze multiple excavators' activities from surveillance images to assess construction site productivity. Luo *et al.*[8] estimated construction equipment posture using various deep learning architectures.

SODA[9] and AIM[10] have developed and released AI datasets specifically for monitoring construction sites, integrating deep learning technologies into surveillance tasks. In contrast to prior research, in this study, we collected CCTV footage in bird's-eye-view format and constructed a dataset for experimentation with the compliance assessment of construction procedures and the segmentation of fall risk areas.

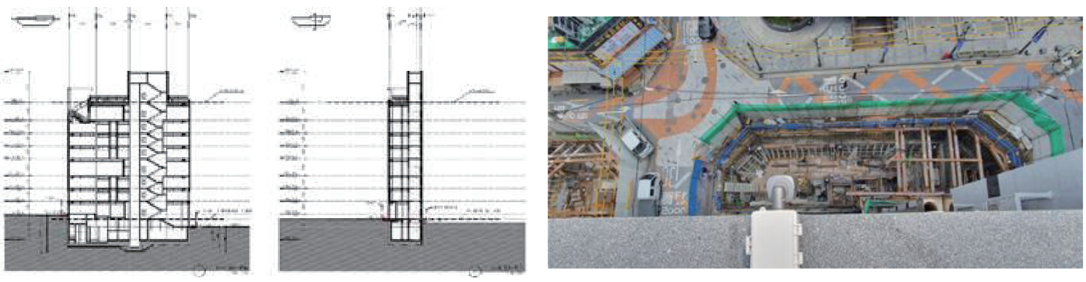
3. Materials and Methods

3.1 Small-scale construction site

To construct a standardized dataset, a neighborhood living facility construction site in Daehyeon-dong, Seodaemun-gu, Seoul was chosen. Figure 1 provides details of the site layout. Figure 1(a) illustrates the overall site plan, encompassing cross and longitudinal views, with the site comprising one basement level and eight above-ground floors. Given the study's focus on safety management through construction procedure monitoring, it was essential to establish an image collection system capable of overseeing the entire site. Unlike conventional approaches that deploy multiple CCTVs within the site, in this study, we positioned CCTVs at elevated vantage points near the site perimeter to capture comprehensive bird's-eye view of the entire area, as depicted in Fig. 1 (b).

Table 1 shows the specifications for CCTV installation. To capture a comprehensive view of the evolving site in real time, a fixed camera model DS-2CE16HOT-IT5F from HIKVISION was deployed. This camera features a 5 MP resolution of 2560 × 1944 pixels and is equipped with waterproof and dustproof capabilities, suitable for high outdoor locations.

The monitoring system, depicted in Fig. 2, comprises an image collection CCTV, a network video recorder (NVR), and an internet network, enabling data access from mobile devices and



(a) (b)

Fig. 1. (Color online) Construction site information: (a) cross and longitudinal sections of the site and (b) example of CCTV installation at the construction site.

Table 1
Dataset status according to number of objects and dataset type.

Type	Manufacturer	Resolution TVL	HD	Digital (DSP)	Focal length	Picture elements $H \times V$	Additional info
Spec	Hikvision	5 MP	Yes	Yes	2.8; 3.6; 6	2560 × 1944	High-quality imaging with 5 MP, 2560 × 1944 resolution, and water and dust resistance (IP67)

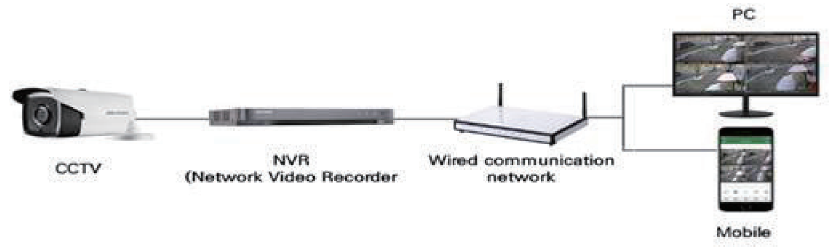


Fig. 2. (Color online) Image collection monitoring system configuration.

PCs via the internet. The system was configured to capture site footage through CCTV during operational hours from 06:00 to 22:00, saving recordings at 30 min intervals.

3.2 Experimental data

Using the monitoring system's image backup feature, approximately three months of original footage were obtained. To encompass all site activities in the acquired footage, the initial step involved sampling from all images. The sampled images were then organized by floor.

Moreover, the construction activities depicted in the sampled images were categorized, and their procedural cycles were identified. Figure 3 illustrates the visualization of construction activities across three floors. Each floor consistently follows a processing sequence comprising foundation work, formwork, rebar installation, wiring, and concrete pouring.

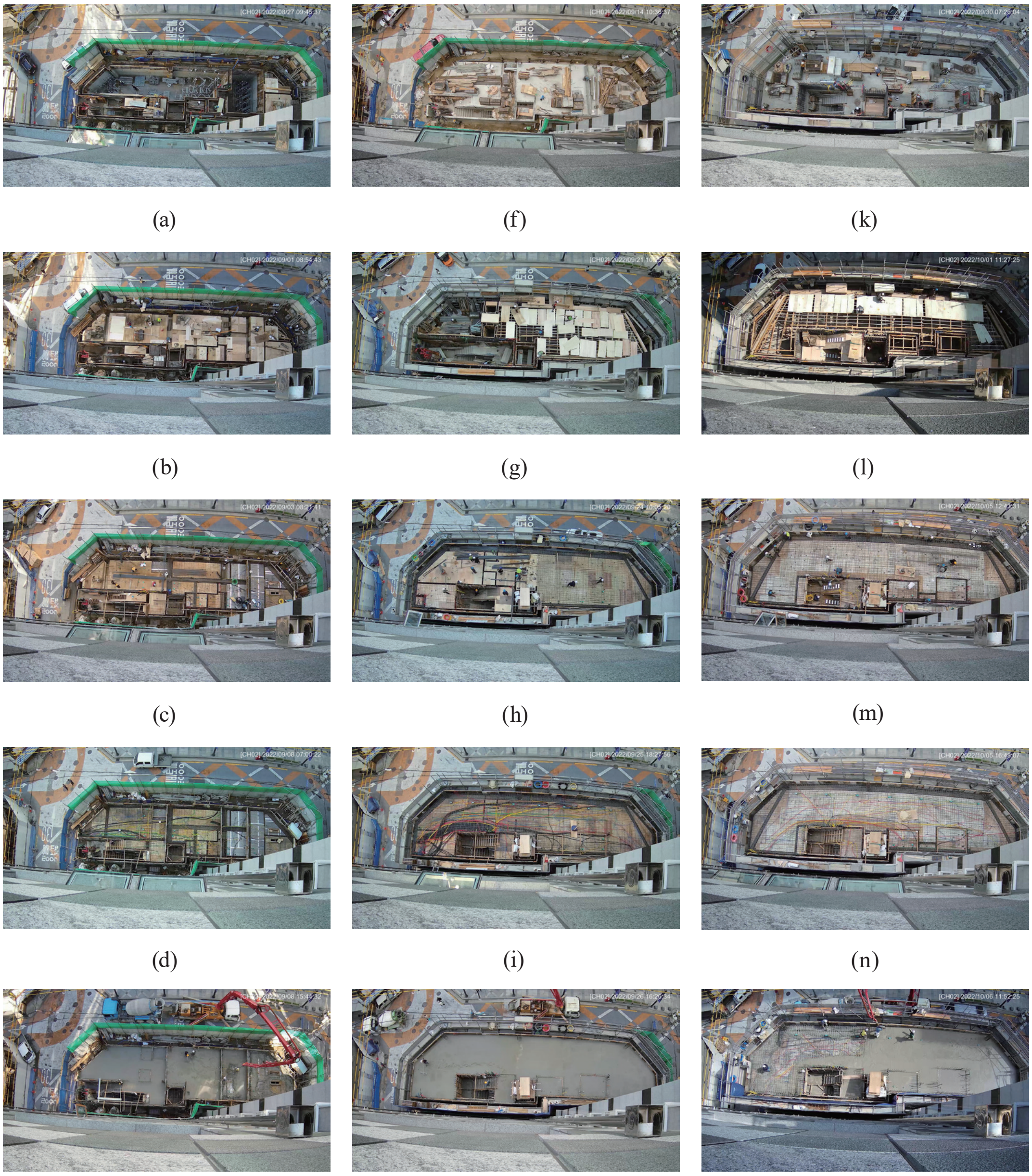
Further analysis was conducted on the basis of the procedural cycles. Table 2 presents the construction durations required for each floor. Generally, the construction period for each floor spans approximately 10 days, although delays occurred in August owing to adverse weather conditions, resulting in extended process durations.

To identify fall risk areas in the images over time, in this paper, we selected specific target objects, as detailed in Table 3. The primary targets for detection were workers and active construction zones. These zones included areas dedicated to foundation construction, formwork construction, rebar work, and concrete pouring. Additionally, rebar work and wiring were grouped together, while formwork construction was subdivided into formwork foundation and formwork ceiling construction.

Table 4 presents the training data status, primarily comprising monthly and categorized image data based on construction types. Monthly image data were sampled at 30 min intervals throughout the acquisition period to capture comprehensive site conditions (e.g., lighting and weather). The monthly samples were further refined to enhance the dataset, specifically targeting critical construction areas to ensure robust detection performance. The experimental dataset comprised a total of 2340 images, split into training and test data in an 8:2 ratio.

3.3 Experimental method

Small-scale construction sites undergo constant changes, resulting in real-time fluctuations in risk areas. Particularly in multi-story building construction, the construction period for each floor progresses rapidly, typically within approximately two weeks, barring severe weather



(a) (f) (k)

(b) (g) (l)

(c) (h) (m)

(d) (i) (n)

(e) (j) (o)

Fig. 3. (Color online) Video samples by floor and construction type: (a) 1st floor foundation work, (b) 1st floor formwork, (c) 1st floor rebar work, (d) 1st floor wiring, (e) 1st floor concrete pouring, (f) 2nd floor foundation work, (g) 2nd floor formwork, (h) 2nd floor rebar work, (i) 2nd floor wiring, (j) 2nd floor concrete pouring, (k) 3rd floor foundation work, (l) 3rd floor formwork, (m) 3rd floor rebar work, (n) 3rd floor wiring, and (o) 3rd floor concrete pouring.

Table 2
Construction period by floor.

Floor	Process name	Period	Construction duration
1st floor	Foundation work	20220818 07:00 – 20220825 11:37	21 days
	Formwork	20220827 08:02 – 20220901 15:19	
	Rebar work	20220902 07:29 – 20220903 11:25	
	Wiring	20220903 11:31 – 20220903 15:22	
	Concrete pouring	20220908 09:07 – 20220908 16:28	
2nd floor	Foundation work	20220913 06:59 – 20220920 10:10	13 days
	Formwork	20220920 13:03 – 20220923 15:46	
	Rebar work	20220924 06:37 – 20220924 12:25	
	Wiring	20220924 12:25 – 20220924 17:18	
	Concrete pouring	20220926 09:13 – 20220926 16:28	
3rd floor	Foundation work	20220927 08:22 – 20220930 11:39	9 days
	Formwork	20221001 07:15 – 20221001 15:33	
	Rebar work	20221005 06:35 – 20221005 13:20	
	Wiring	20221005 13:20 – 20221005 16:14	
	Concrete pouring	20221006 08:36 – 20221006 12:29	

Table 3
Detection target object selection.

Detection object	Subdivided detection object	Class name
Worker	Worker	worker
	Worker without safety equipment	worker_no_helmet
Foundation work	Foundation construction	foundation_construction
Formwork	Formwork foundation construction	form_foundation_construction
	Formwork ceiling construction	form_ceiling_construction
Rebar work wiring	Rebar work	reinforcing_bar_construction
Concrete pouring	Concrete pouring	concrete_construction

Table 4
Detection target object selection.

Category	Count	Total
August	602	1657
September	827	
October	228	
Formwork	342	683
Rebar work	171	
Concrete pouring	170	

conditions. To effectively segment risk areas according to the dynamic zones, it is essential to analyze CCTV footage to distinguish construction areas and evaluate risk zones through post-processing.

Figure 4 illustrates the entire experimental process, divided into the training and inference phases. The training phase involves developing a model based on instance segmentation using the collected data, assessing object recognition and area segmentation performance to ensure model reliability. The inference phase sequentially conducts object recognition and segmentation on input images, followed by post-processing to identify risk areas.

The risk area designation changes the current floor area to a fall risk area on the basis of the progression to the next floor. Figure 5(a) depicts the image of the *N*th floor, with Fig. 5(c)

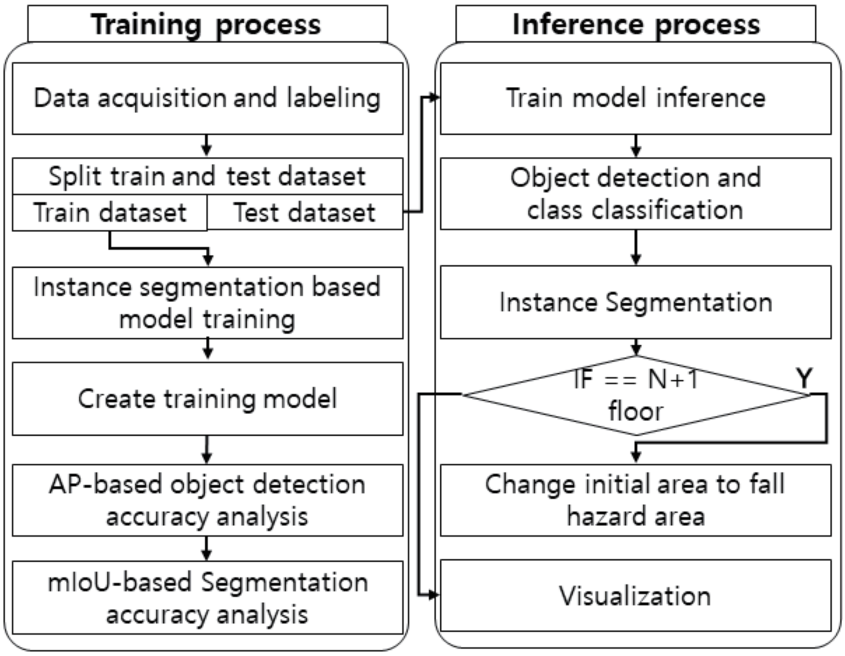
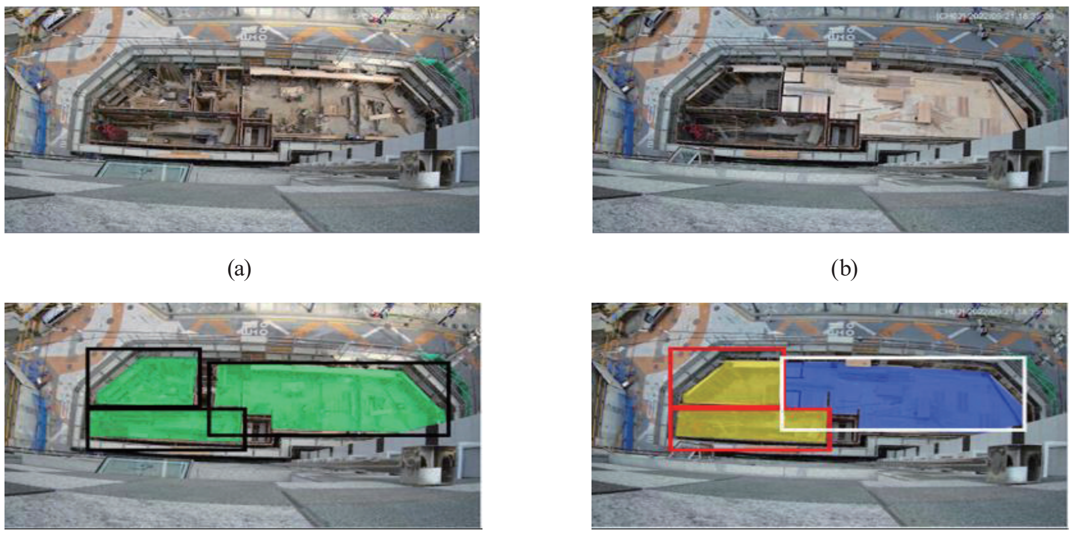


Fig. 4. Experimental process.



(a) (b)

(c) (d)

Fig. 5. (Color online) Criteria for changing risk areas: (a) 1st floor image, (b) 2nd floor image, (c) 1st floor foundation construction area, and (d) 2nd floor formwork ceiling and risk areas.

showing the designated green area for foundation construction. Figure 5(d) illustrates the blue area for the formwork ceiling and the yellow area, also designated for foundation construction. As construction progresses on the *N*th floor, the appearance of the blue area indicates the onset of the $N + 1$ floor area. Thus, the presence of the formwork ceiling area signals the occurrence of the $N + 1$ floor area. Consequently, the post-processing procedure for determining risk areas is designed to designate the foundation construction area as a fall risk area based on the presence of the formwork ceiling area.

4. Results and Discussion

The primary model used for training was Mask R-CNN,[11] a prominent model in the instance segmentation field. Figure 6 illustrates the model structure of Mask R-CNN. Mask R-CNN operates in a two-stage manner, incorporating a Fully Convolutional Network (FCN) into the foundational structure of Faster R-CNN.[12,13]

Mask R-CNN applies FCN to the region of interest (RoI) extracted during the Region Proposal Network (RPN) process. RPN functions as a localization technique that estimates object positions in images on the basis of anchor box parameters. In this study, we conducted experiments analyzing visual inference and numerical results from the training dataset using Mask R-CNN. Numerical evaluations employed mean average precision (mAP) and mean intersection over union (mIoU), standard metrics in object recognition and segmentation models.[14,15] mAP assesses the accuracy of object detection via bounding boxes, while mIoU measures how closely the predicted area within a bounding box aligns with the ground truth.[15] The IoU threshold used for evaluation was set at 0.5.

4.1 Numerical performance evaluation results

Table 5 presents the numerical evaluation results of the model using various metrics. Both structured (related to workers) and unstructured (construction areas) objects exhibited an overall object detection accuracy exceeding 80%. The segmentation accuracy (mIoU) also achieved performance close to 80%, albeit slightly lower than the object detection accuracy. Specifically for fall risk areas, the experiment achieved an accuracy of approximately 70% based on the criteria for determining the risk areas described earlier.

Figure 7 depicts the visual evaluation results of the model for various objects in the images. The experiments were conducted under consistent conditions, confirming the accurate classification of object types at their respective positions. However, upon comparing the detection boundaries in each image to the ground truth, minor mis-segmentations were

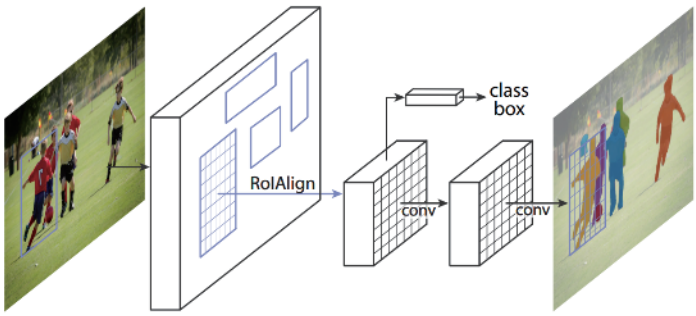


Fig. 6. (Color online) Experimental process (He *et al.*[11]).

Table 5
Detection results by target object.

Target object	Inference results	
	Mean AP	Mean IoU
Worker	0.88	0.82
Worker without helmet	0.81	0.75
Foundation construction	0.83	0.77
Formwork foundation construction	0.82	0.78
Formwork ceiling construction	0.81	0.75
Rebar work	0.84	0.78
Concrete pouring	0.85	0.79
Fall risk area	0.71	0.68

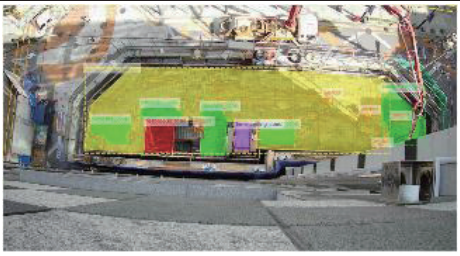
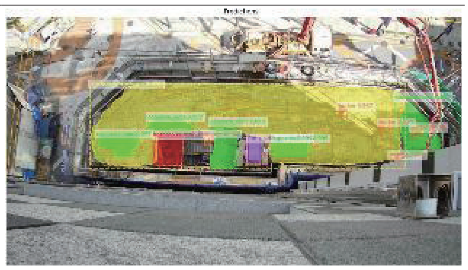
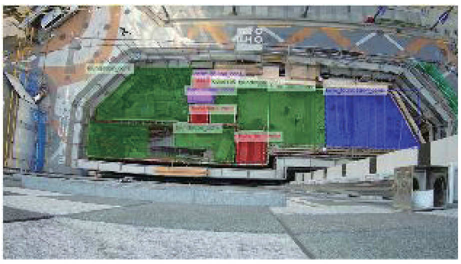
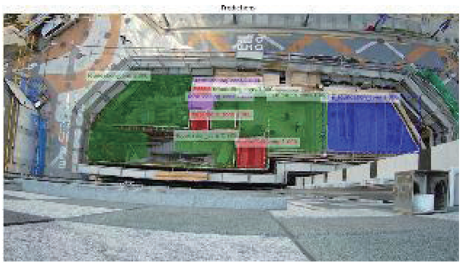
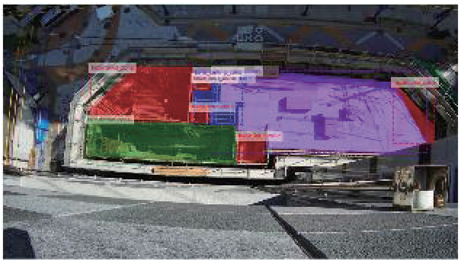
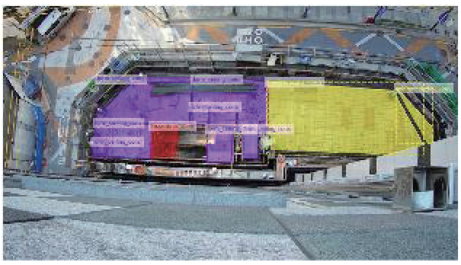
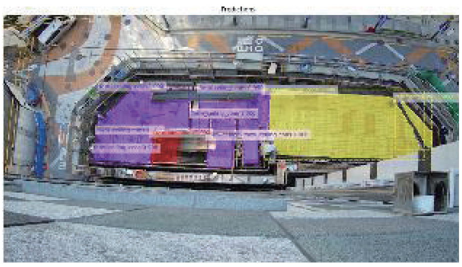
Process name	Example	
	Ground truth	Inference result
Worker		
Formwork foundation construction		
Formwork ceiling construction		
Rebar work construction		

Fig. 7. (Color online) Comparison of ground truth and inference results.

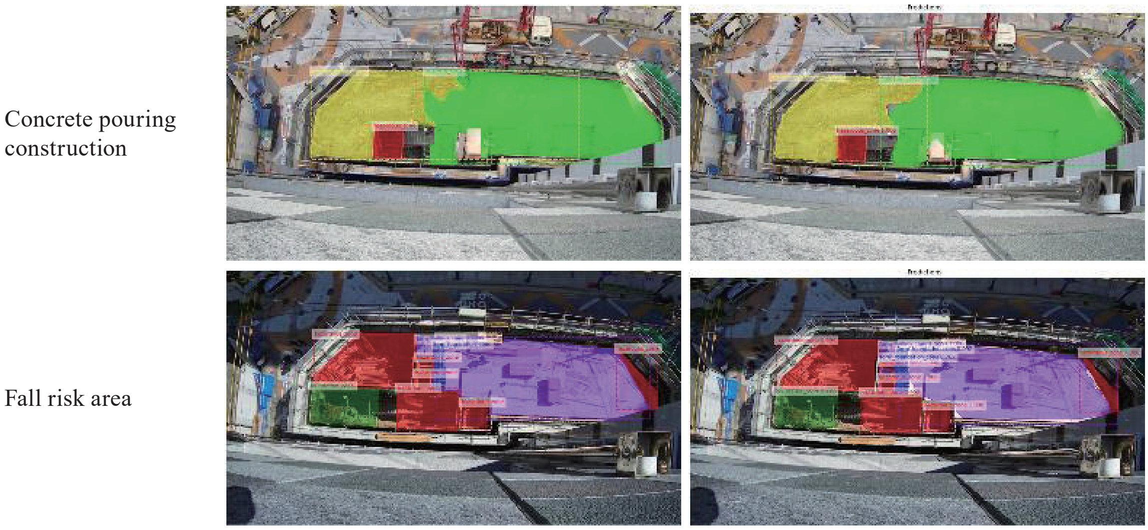


Fig. 7. (Color online) (continued) Comparison of ground truth and inference results.

observed. It is anticipated that incorporating additional training data and fine-tuning model parameters will enhance accuracy levels suitable for practical field applications.

5. Conclusion

In this study, we investigated the real-time detection of fall risk areas from bird's-eye-view images for safety monitoring at small-scale construction sites. CCTV footage was collected from actual small-scale construction sites, and key detection targets were identified by segmenting the stages of small-scale layered construction processes. The entire dataset was sampled and processed during the construction period to create the training dataset. Experimental results demonstrated an object detection accuracy of approximately 80%, with segmentation performance exceeding 70%. Further improvements are anticipated through the inclusion of additional data and the optimization of the model's training parameters. Furthermore, a methodology was proposed for determining fall risk areas through inter-object analysis, achieving a performance level of 70%, underscoring the study's potential. This research defines critical objects at construction sites and utilizes image-based information for object recognition, contributing to the evaluation of compliance with construction procedures and enhancing safety management through monitoring at small-scale construction sites. In this study, we recognize the need to improve detection performance by securing various sites and further processing the currently available datasets. Additionally, we also recognize the need to develop a platform that allows safety managers to easily monitor hazardous area through the integration of Building Information Modeling (BIM).

Acknowledgments

This study was supported by the Dong-A University research fund.

References

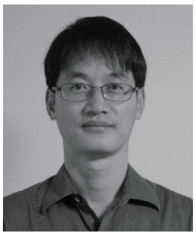
1 Ministry of Employment and Labor (MOEL): Industrial Accident Status (in Korean) 2020.
2 L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen: Int. J. Comput. Vis. **128** (2020) 261. https://doi.org/10.1007/s11263-019-01247-4
3 X. Xiang, N. Lv, X. Guo, S. Wang, and A. El Saddik: Sensors (Basel) **18** (2018) 2258. https://doi.org/10.3390/s18072258
4 Y. Guo, Y. Xu, and S. Li: Autom. Constr. **112** (2020) 103124. https://doi.org/10.1016/j.autcon.2020.103124
5 Z. Yang, Y. Yuan, M. Zhang, X. Zhao, Y. Zhang, and B. Tian: Sensors (Basel) **19** (2019) 2789. https://doi.org/10.3390/s19122789
6 Q. Fang, H. Li, X. Luo, C. Li, and W. An: Comput. Aid. Civ. Infrastruct. Eng. **35** (2020) 979. https://doi.org/10.1111/mice.12541
7 C. Chen, Z. Zhu, and A. Hammad: Autom. Constr. **110** (2020) 103045. https://doi.org/10.1016/j.autcon.2019.103045
8 H. Luo, M. Wang, P. K.-Y. Wong, and J. C. P. Cheng: Autom. Constr. **110** (2020) 103016. https://doi.org/10.1016/j.autcon.2019.103016
9 R. Duan, H. Deng, M. Tian, Y. Deng, and J. Lin: Autom. Constr. **142** (2022) 104499. https://doi.org/10.1016/j.autcon.2022.104499
10 H. Kim, H. Kim, Y. W. Hong, and H. Byun: J. Comput. Civ. Eng. **32** (2018) 04017082. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000731
11 K. He, G. Gkioxari, P. Dollár, and R. Girshick: Proc. IEEE Int. Conf. Computer Vision, Venice: Italy (2017) 2980–2988. https://doi.org/10.1109/ICCV.2017.322
12 S. Ren, K. He, R. Girshick, and J. Sun: IEEE Trans. Pattern Anal. Mach. Intell. **39** (2017) 1137. https://doi.org/10.1109/TPAMI.2016.2577031
13 J. Long, E. Shelhamer, and T. Darrell: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Boston: MA (2015) 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965
14 K. E. A. Van De Sande, T. Gevers, and C. G. M. Snoek: IEEE Trans. Pattern Anal. Mach. Intell. **32** (2010) 1582. https://doi.org/10.1109/TPAMI.2009.154
15 D. M. Powers: arXiv. https://arxiv.org/abs/2010.16061

About the Authors

Jong Ho Na received his B.S. and M.S. degrees from Hankuk University of Foreign Studies, Korea, in 2016 and 2018, Currently, he is Ph.D candidate from Hankuk University of Foreign Studies, Korea. Since 2020, he has been a researcher at Korea Institute of Civil Engineering and Building Technology (KICT). His research interests are in computer vision, artificial intelligence, and intelligence CCTV. (najongho@kict.re.kr)

Jae Kang Lee received his Ph.D degree from University of Nottingham, UK, in 2015. From 2019 to 2022, he was a senior researcher at Korea Institute of Civil Engineering and Building Technology (KICT). He is currently an associate professor at Dong-A University. His research interests range from safety policy in civil engineering to software programming in relation to GNSS positioning, UAV photogrammetry, and the application of the big data platform to civil engineering. (jaekanglee@dau.ac.kr)

Hyu Soung Shin received his Ph.D degree from University of Wales Swansea, UK, in 2002. Since 2003, he has been a research fellow at Korea Institute of Civil Engineering and Building Technology (KICT). He is also a professor at the University of Science & Technology. He is now in intensive technical activity as the chairman of the technical committee in the IT division of The Korea Tunneling Association (KTA). His research interests are in artificial intelligence, and underground space technologies. (hyushin@kict.re.kr)



Il Dong Yun received his Ph.D degree from Seoul National University, Korea, in 1996. Since 1997, he has been a professor at Hankuk University of Foreign Studies. He is now in intensive international and domestic technical activity as the vice president of the technical committee in Korea Conference on Computer Vision (KCCV). His research interests are in artificial intelligence, and image processing. (yun@hufs.ac.kr)