Scaffold safety monitoring using SLAM-based Bird Eye View Point Cloud Data Images

Nahye Koo[a] , Huiying Li[b] , Duho Chung[c]

[a,c]Advanced Infrastructure Management Group, Department of Civil and Environmental Engineering, Yonsei University.
Email: ahappyto@yonsei.ac.kr
Email: jungduho1@yonsei.ac.kr
[b]Construction Management & Information Laboratory, Department of Civil and Environmental Engineering, Yonsei University.
Email: hyeyeong1111@yonsei.ac.kr

INTRODUCTION

Safety management has always been a very important topic on construction sites. Construction personnel has been facing difficulties and limitations in terms of safe construction due to the difficulty of monitoring safety on site in real-time. Accord ing to NIOSH, for every 100,000 people, 15.2 people die in an accident on constru ction sites because of safety issues. This data tells us that safety management will be a problem that can no longer be ignored. Safety management in construction sit es is mainly divided into several aspects: organizational manageability, security of construction site equipment, technical management, and industrial production rela tions. In order to ensure the safety of the construction site's large apparatus as well as temporary facilities, construction personnel has progressed from the original h uman detection to the current artificial intelligence detection. A great breakthroug h has also been made in the field of safety management, such as image recognition detection with robots, stress brainwave detection of workers' brains facing danger, etc. In construction sites, what is often overlooked is the safety of temporary struc tures, its more complex structure and the complexity of the site can make the cons truction site safety personnel also overlook this part of the problem. One of the te mporary structures represented at the construction site is scaffolding. Scaffolding i s also frequently mentioned in the regulations for construction site management, s uch as those proposed in the regulations (KOSHA GUIDE C-32-2020). So a scaffold in which there is a platform deserves to be without stairs, and where there is a stai r deserves to be without a platform. However, if in a place both platform and stairs are not set at the same time, then this place will cause safety problems. For examp le, workers accidentally fall or because of the light limitations of night work, will cause a series of safety accidents. In order to avoid such accidents, this research w ill combine the construction site artificial intelligence technology, using point clou ds to constitute real-time video of the scaffold, identifying whether there are safet y hazards in the current scaffold according to each position in its video frame, taki ng the bird‘s eye view as the base, and using the image detection algorithm YOLO v5 to classify safety into safe and non-safe, and finally using robotic dog at the co nstruction site according to the prescribed route to detect the safety of the site in re al-time.

This research has three contributions. The first one is using robot dog to acquire d ata automatically. The robot dog successfully created a 3D map using LIO-SAM t hrough teleoperation. Based on the trajectory of this path, it automatically creates a waypoint and repeats the patrol with the same trajectory, and automatically acqu ires data. Second, the model implemented real-time safety judgment. By convertin g point cloud data into images, scaffold safety can be judged in real-time. Point cl oud data is not affected by environmental factors such as light to improve accurac y. In this study, image detection was attempted by changing the point cloud to BE V images. The last is applying the detection model to general construction sites. In other words, the goal is to make it a common-use system. Since point cloud data c ontains spatial information, so it is easier to train a general model than using an im age. And scaffolds are standardized; scaffold units exist in other sites as well. The refore, our model can maintain satisfactory performance in various fields.

LITERATURE REVIEW

Issues related to safety at construction sites have constantly been present. Among t he types of data used to monitor construction sites, the most famous methods are t he method using the point cloud and the method using the image. This section revi ews how research has been developed to improve construction site safety and disc usses the pros and cons of construction site safety studies using point clouds and i mages.

The importance of construction temporary structures safety and scaffold in site

Safety management is the root of the construction site. The researchers mentioned the four ways to supervise the construction site or collet data of the infrastructure among the image-based supervision. (Fathi 2015). The safety management problems are identified in different status countries, even though in the large construction market country China. Regarding China’s site safety control, the record is poor than the international standards (Tam, 2003). Also as reported in NIOSH, the actual fatality for the USA is 15.2 per 100,000 workers. (NIOSH,2003). The construction management included various aspects such as people’s role, organization and management, apparatus and equipment, technology, and industrial relationship. In terms of temporary works, in construction projects often include temporary structures, such as excavation support systems, underpinning, scaffolds, bracing and shoring, formwork, falsework, work platforms, decking, roof protection, and structures used in on-site contractor facilities (Tam, 2003). Despite their impact on safety, temporary structures such as scaffolding and temporary stair towers are rarely planned and delineated in building drawings or models (Ratay 1996). The general requirement for the revised scaffold safety standard has mentioned seven sections including the capacity, scaffold platform, criteria for supported scaffolds, criteria for suspension scaffolds, and provide scaffold capable of supporting (29 CFR 1926.451). To describe more detailed information for the second condition for scaffold, each scaffold platform should be fully planked or decked between the front uprights and the guardrail supports (29 CFR 1926.451(b)). So the scaffold is essential for construction safety control.

The development of automation in construction safety management system

The recent developments in robotic systems have led to wide range of automated construction applications that are mostly based on civil infrastructure and house building, for instance, automation of road, tunnel and bridge construction, earthwork or house construction, including building skeleton erection and assembly, concrete compaction, and interior finishing (Balaguer 2008). And nowadays the construction industry has become one of the most important research areas in the field of service robotics. The researchers compared the situation of the construction and automobile industries to find the differences. One of the key factors of any industry's success evaluation is its productivity. And the main reason for this high productivity is the modern manufacturing concept (Rembold et al., 1993), (Rehg,1994). While the house building construction industry continues to be very close to craft work, constructing mostly singular buildings, the automobile industry continuously seeks to reduce the cost of product development. The main research activities in the past decade were divided accordingly to applications into two groups; the first one is civil infrastructure and house building. For example, the typical civil infrastructure robot applications are the automation of road, tunnel and so on. In the group of house construction, researchers implement the building skeleton erection and assembly, concrete compaction, interior finishing and so on (Balaguer 2008). For instance, research utilize the GPS and lase data to control the speed, temperature, layer thickness, travelled distance in civil infrastructures. In the field of earthwork, the researchers implemented the introduction of new coral techniques to existing machinery like excavators, bulldozers. One of the major exponents of this research area is the control by CSIRO of the 100m tall walking crane used in surface coal mining (Corke et al.,2006). An automated excavator that accounts for interaction forces in analyzing the required bucket motion therefore seems promising (Ha et al.,2000). In terms of house building, the most representatives' robots of this type are Japanese ones. There are three examples are presented: the Mighty Hand robot from Kajima, which lifts heavy elements in construction as concrete. After that, due to the number of high-risk buildings increases, more safety issues related to dangerous and repetitive work in high workplaces arise in many constructions site. Steel beam assembly danger of negligent accidents. So, the researchers develop a robotic beam assembly system administered by the robot research group (Hasegawa,2006).

Related Work–Point Cloud

Point cloud data is widely used in the construction industry for 3D reconstruction due to its ability to obtain accurate spatial information about the structure. Efficien t 3D reconstruction of construction sites can be widely applied to construction ma nagement areas such as progress monitoring, safety management, and quality chec king. Especially with the rapid improvement of LiDAR technology, laser scanning point cloud has occupied the mainstream of point cloud-based methodology in co nstruction with its power to perform the more accurate 3D reconstruction. But bec ause high computation costs of laser scanning point clouds have always been an o bstacle to practical usage of point clouds. To overcome this obstacle, many researc hers proposed a method based on simplifying the features of point clouds. Bassier proposed a framework for automatic reconstruction of wall objects in BIM (Buildi ng Information Modeling) based on point cloud data of walls(Bassier 2020). The wall point cloud features are retrieved with machine learning techniques, and unsu pervised reconstruction of lines, arcs, and polylines is performed to fit into the exi sting BIM. Kim et al. proposed a novel method for automatic point cloud 3D regis tration using visual and planer features of construction environments using feature detection algorithms (Kim 2018). Chen et al. used a data-driven deep learning fra mework to identify building elements from the point cloud scene on the constructi on site. The point cloud is simplified into edges, and a data-driven deep learning cl assifier is used to determine the type of building components (Chen 2019). Previo us studies have demonstrated that abstracting the features of point clouds is an effi cient method for 3D reconstruction and data interpretation. However, these metho ds are solely focused on 3D reconstruction and data retrieval. There remains a lack of research on applying the point cloud data directly in construction management, such as safety management or quality checking. So for this study, the abstraction of point clouds was directly applied to safety management, especially for safety ch ecking of scaffolds.

One of the most time-consuming and labor-intensive tasks for point cloud-based 3 D reconstruction on the construction site is the data acquisition process. Using mo bile robots to automate this data acquisition process, Kim et al. proposed a framew ork for automated scaffold segmentation and 3D reconstruction based on mobile L iDAR-based 3D point clouds. Teleoperated robotic dog was used for the acquisitio n of point clouds, and deep learning-based point cloud segmentation was performe d. The segmented scaffold point cloud is reconstructed into the 3D CAD model (K im 2022). Kim et al. proposed a framework for a fully automatic 3D data acquisiti on and registration system using a mobile robot. The author developed an intuitive point cloud 3D registration algorithm and used a mobile robot to reduce the effort and time of the acquisition process based on rule-based navigation (Kim 2018).Ki m et al. proposed a framework using drones and mobile robots to automate the dat a acquisition process efficiently(Kim 2019). These studies have demonstrated that using mobile robots in the data acquisition process can reduce time and labor cons umption by providing repeatability to the process. For this study, a semi-automate d robot dog was used to give repeatability to the proposed framework.

Related Work–Image

With the development of computer vision, a lot of research on automatic monitoring of construction sites using it is being conducted. Computer vision is mainly used as a method for monitoring the site by automatically acquiring data using tools for efficient monitoring of construction site and analyzing the data to prevent the aforementioned accidents. Because image processing has a fast-processing process, it is often used to monitor construction sites in real time or to judge safety.

Image detection model

Object detection is an important technique in machine learning that classifies instances and indicates their location (Kulchandani et al. 2015). A common type of deep learning object detection algorithm is the 2-stage-obejct detection model. They generate a large number of candidate boxes and then use a pre-trained classifier to classify these boxes. The two-stage object detection model achieved the highest performance in terms of accuracy in several benchmarks. A representative example is a region-based convolutional neural network (R-CNN), which uses a support vector machine (SVM) to extract image features and classify candidate boxes, and for this, a convolutional neural network is employed. This two-step detection method has been advanced to high performance by other techniques such as region-of-interest (ROI) pooling (Girshick 2015) and anchor boxes (Ren et al. 2017), and is much faster than before. achieved However, in order to use the model in real time, a higher speed was required, which resulted in a 1-stage detection model. The most famous 1-stage object detection model is YOLO (You Only Look Once) (Redmon et al. 2016), and by storing the resulting prediction in the last convolution layer, the algorithm was processed at a high speed to achieve real-time speed.

Image detection in construction

We investigated studies incorporating object detection using image data at construction sites. Park et al. proposed a detection framework to recognize construction workers by using various features such as motion (by background subtraction), histogram of gradient(HOG) shape and color(Park,2012). Son et al. proposed a method to accurately and quickly detect construction workers in a changing background of various poses and image sequences with Faster R-CNN using ResNet152, a very deep residual networks. Memarzadeh et al. developed a novel descriptor by combining the histograms of HOG and Hue-Saturation-Value (HSV) and the trained SVM (Memarzadeh, 2013). Son et al. proposed an architecture that detects and tracks construction workers using an image sensor of Complementary Metal-Oxide Semiconductor (CMOS) and a Siamese network based on YOLO (Son, 2019). Kim et al. integrated computer vision with a fuzzy inference to monitor and assess the safety of people performing their tasks in the vicinity of plant (Kim, 2016). There are a lot more studies using computer vision and DNN for automate visual site monitoring tasks such as safety monitoring (Fang et al.2018a,b) progress monitoring(Braun et al.2020;Lei et al.2019).

In order to improve the performance of a model in computer vision, a large amount of dataset is required. However, since the construction site image dataset is alwa ys insufficient, there are many studies on various image augmentation techniques. Bang et al**.** proposed an image augmentation method by leveraging Generative Ad versarial Network (GAN) to develop a large-size dataset for improving constructio n resource detection (Bang, 2020). Baek et al proposed a novel data augmentation method that interates a conditional GANs framework with a target classifier (Baek, 2022).

METHODOLOGY

SLAM

SLAM is short for Simultaneous Localization and Mapping, which is a method to enable a robot to estimate its current waypoint, position, and orientation as well as a 3D map in the construction site. The point cloud generation can divide into two categories, the first one is the laser-measurement point cloud and the second one is the photogrammetry point cloud, and in this experiment, we utilize laser to measu re the distance, orientation, and the map coordinates. And using a lidar- based SL AM is the main method in this research. To compare the differences and find the a dvantages of different types of SLAM, the researchers found the proposed method of the lidar-based SLAM has the known absolute scale, high resolution map, 3D man and RGB-mapped points (Shang 2017).

YOLO v5

YOLO is a state-of-the-art, real time object detection system which is short for "Y ou Look only once". Since 2016 YOLOv1 was firstly invented. The processing im ages with YOLO is simple and straightforward. And then YOLOv2, YOLOv3, Y OLOv4, Tiny-YOLOv3 are invented. Among the YOLOs, YOLOv3 is the most re presented one. YOLOv3 is extremely fast and accurate. In mAP measured at .5 Y OLOv3 is on par with Focal Loss but about 4x faster(https://pjreddie.com/darknet/ yolo/). Moreover, the researchers can easily tradeoff between speed and accuracy simply by changing the size of the model. It is no retraining required. The structur e of YOLOv4 and YOLOv5 is very similar. And the implementation of YOLOv5 i nto embedded devices is very simple, which is only required the torch installation and python libraries. So the final cost will be reduced because of the lower require ments of YOLOv5. YOLOv5 can infer video feeds, batch imagers, individual ima ges, and webcam ports. In this paper, we utilized the YOLOv5 in object detection field (Adibhatla 2021).

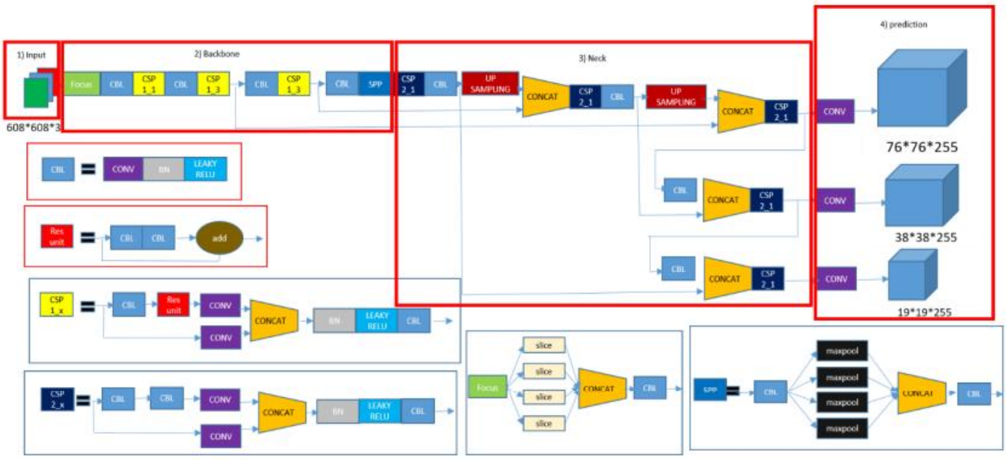


Figure 1. Structure of YOLO-v5

Semi-Automated Scanning Platform

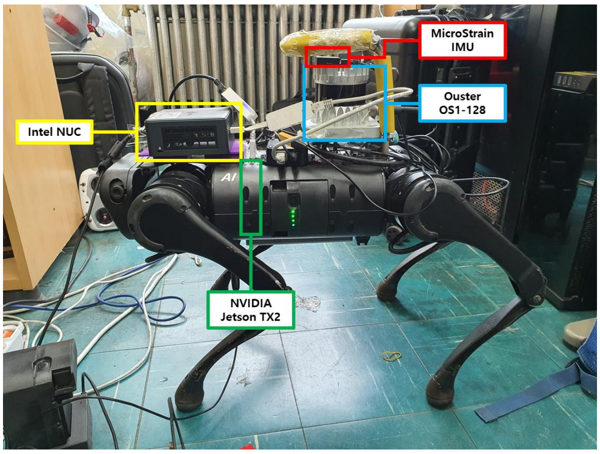


Figure 2. Scanning platform

Semi-automated mobile robot dog was used as a scanning platform in this study. Ouster OS1-128 mobile LiDAR, Microstrain IMU (Inertial Measurement Unit) was deployed on the robot. 2 computers, NVIDIA Jetson TX2 and Intel NUC 11 were used to control the whole system. In the area of point cloud data acquisition, a robot dog can have two significant advantages over a wheeled mobile robot. First, a robot dog can walk stably through rough terrains and small obstacles. Because construction sites generally have rough and uneven surfaces compared to refined workspaces, a wheeled robot's bump could affect the results of mobile laser scanning. Second, a robot dog has more DoF (Degree of Freedom) compared to wheeled mobile robots. This can easily increase the FOV (Field of View) without additional actuators. All the components equipped in the robot have been wired, connected, and communicate based on ROS (Robot Operating System).

The robot was controlled via TeamViewer, an internet-based software package for desktop sharing between computers. Because desktop sharing program like Team Viewer displays an exact copy of the output of the graphics card, headless systems without a graphics display like our system cannot use this method right away. To solve this problem, a dummy HDMI plug, a device that sends fake graphics information to the computer, was connected to the graphics card. By using TeamViewer, internet-based teleoperation is possible. KT 5G egg was attached to the robot hardware to provide internet consistently to the robot system. This teleoperation methodology has three substantial advantages: First, because the communication is based on the internet, the range of the teleoperation is limitless. Second, because the security system is based on the ID and password inside the desktop sharing program instead of the network itself, the teleoperation can operate very flexibly on many computers. Third, because the system mirrors the same screen inside the computer, even people unfamiliar with technology, such as real industrial workers, can operate it intuitively.

Semi-automation has been applied to the robot dog to give repeatability to the scanning process. The scanning process can be divided into two parts: Teleoperation and automation. In the teleoperation part, the robot was teleoperated through the course that the user wanted to give repeatability. While the robot is being teleoperated, the robot automatically records the trajectory as waypoints. After the teleoperation part has finished, the automation part is performed: the robot repeatedly moves along the designated waypoints. This way, the controller can initiatively plan the optimal path for the robot and provide repeatability to the planned route, enabling practical semi-automation to the robot dog scanning system.

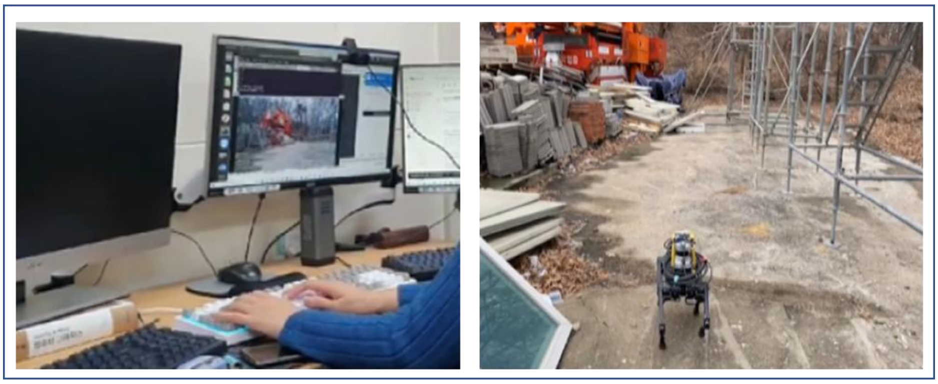


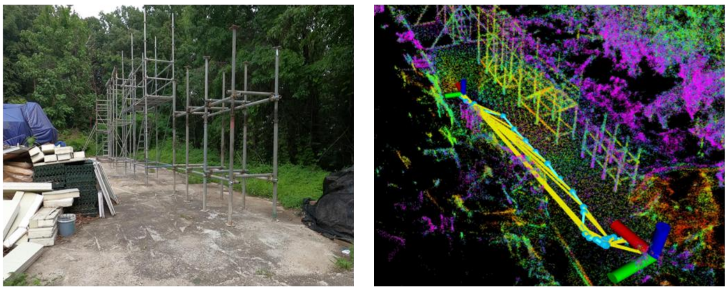
Figure 3. Internet-based long-range teleoperation of robot

3D SLAM based Bird's Eye View Projection

3D SLAM algorithm is used for localization of the robot dog and registration of point cloud data into a 3D point cloud map. Localization of the robot dog is used for path planning and automated navigation, and a 3D point cloud map is abstracted to detect safe scaffolds and unsafe scaffolds in real-time.

LIO-SAM (Shan 2020), a 3D SLAM algorithm that uses LiDAR and IMU data, was used for this study. LIO-SAM receives LiDAR point cloud data and IMU sensor data as input. The estimated motion based on IMU data de-skews point clouds to calculate an initial guess for LiDAR odometry. The bias of IMU is calculated based on the LiDAR odometry solution. Scan-matching is performed on a local scale to speed up one calculation by marginalizing old lidar scans for pose optimization.

Figure 4 (b) shows the registration of point clouds using LIO-SAM at the scaffold testbed at Yonsei University.



(a) (b)

Figure 4. Scaffold testbed in Yonsei University: (a) Photogrammetry (b) LIO-S AM registered point clouds

To apply the registered 3D point cloud map to scaffold safety management, deep l earning-based object detection is used to understand the information in the map au tomatically. Compared to image-based object detection, point cloud-based detecti on has more accurate spatial information. Even if image detection identifies the co mponents of a structure, the post-processing process is crucial for understanding t he exact spatial information of the structure. Point cloud-based detection has the p ower to capture the geometry of detected components accurately, so the result for object detection can be directly used for the next step, such as safety checking. Als o, because point cloud data is mainly based on spatial information instead of visua l information, it is easier to generalize the detection result to other environments. But applying deep learning-based object detection to point cloud map has some m ajor weaknesses. First, point cloud-based detection has a very high computational cost than image-based detection. To overcome this problem, some studies attempt ed to project raw point cloud data into BEV (Bird's Eye View) images and applied an image-based object detection algorithm to the BEV image (Beltran 2018). This idea can perform image detection using 3D point clouds in real-time, but this lead s to another problem: Point cloud has relatively sparse visual information compare d to the image, and abstracting the features such as BEV projection makes the visu al information sparser. This limits the performance of the detection model. To solv e this problem, this study applied 3D SLAM to BEV projection to make the visual information denser. By using this method, a 3D point cloud map could be directly applied to scaffold safety management using deep learning-based object detection. Figure 5. shows the framework in this study.

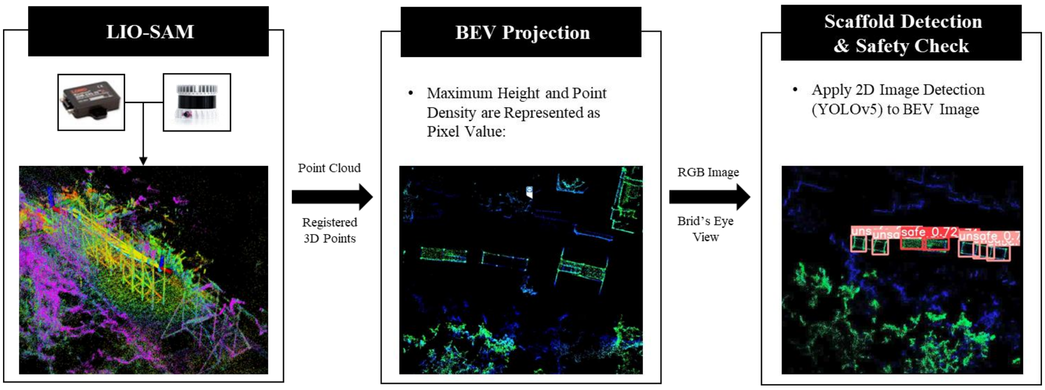


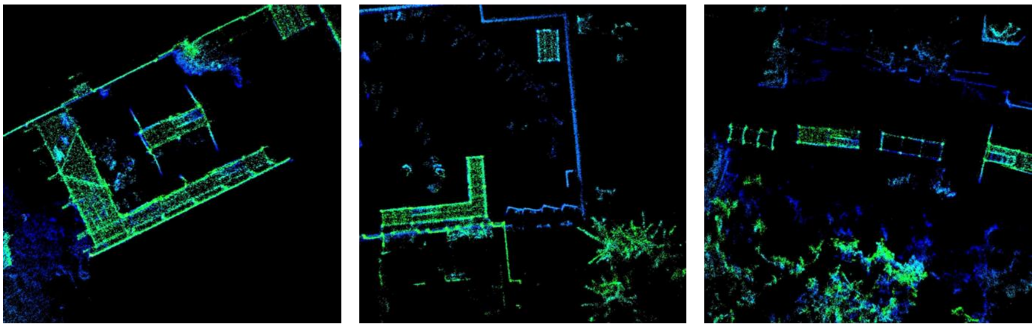
Figure 5. Scaffold detection & Safety checking using SLAM-BEV data

Previous BEV projection-based methods such as (Beltrán 2018), usually preserve maximum height, maximum intensity, and point cloud density as RGB pixel value s. But for this study, the intensity value was discarded to make more distinct visua lization of the maximum height value. This is because to classify whether if the sc affold has an appropriate platform or stairs, the difference in height should be clea rly distinguishable. So green and blue pixel value marks the maximum height of th e point cloud map, and the red pixel value marks the density of the point cloud ma p. The maximum height was set to 2.3m in point cloud map based on the scaffoldi ng regulation. The image only shows up to second floor of scaffold but can be gen eralized to other floors by stacking multiple BEV images with the same range, diff erent value of heights. Figure 7 shows the example of BEV images in training and testing dataset.

Because it was impossible to get data of unsafe scaffolds with missing stairs or pla tforms, synthetic unsafe scaffold data was generated. By filtering out the heights o f appropriate platforms and stairs, point cloud data of platform-less scaffolds are g enerated. Figure 8 shows the example of synthetic unsafe scaffold data.

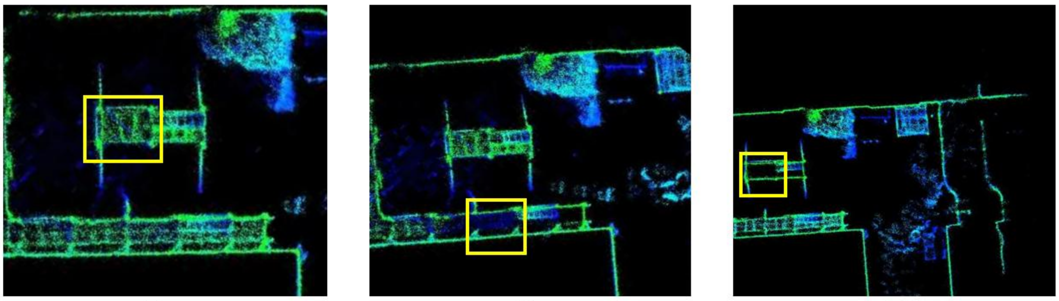
(a) (b) (c)

Figure 6. Example of data acquisition location. (a) Ansung Testbed-Training (b) ChungAng University Testbed-Training (c) Yonsei University Testbed-Testing



(a) (b) (c)

Figure 7. Example of BEV images in training and testing dataset. (a) Ansung Te stbed-Training (b) ChungAng University Testbed-Training (c) Yonsei University Testbed-Testing



(a) (b) (c)

Figure 8. (a) Original data (b) Synthetic data without platforms (c) Synthetic dat a without stairs

Safe Detection of SLAM-BEV images of scaffold.

Image set

A total of 280 train images and 88 validation images were used for model train. T he train and validation images were randomly divided. The images used for train a nd validation are data acquired from two sites, 'ChungAng University Testbed and 'Ansung Testbed'. Both data were acquired by Robot Dog. The 'Yonsei Universit y Testbed' was used as the test dataset. Sites are shown in Figure 6.

Image Labeling Standards

There are two types of classes used in image detection: 'safe' and 'unsafe'. Image detection. Units are divided based on the vertical of the scaffold. 'Two platforms' or 'one platform and one stairs' are exist in one unit. If the scaffold platform or sta irs exist, the points are uniformly taken in the 3D point cloud BEV image. Sparse data is not used because the safety detection is performed after generating the 3D map for the first time. Based on the guard border, it was judged that both platform s or stairs exist if the points are uniformly stamped. The color of the point cloud d ata used varies according to the height, and the green dot means high height. Ther efore, the 2nd floor platform is shown in green, and the stairs are shown in blue an d green as a gradient. The green point and gradation point are uniformly taken means that both the 2nd floor platform or stairs exist. So it is classified as 'safe'.(Figure 8(a)). If there is only a 1st floor platform without 2nd floor, it is expressed in blue color points because it has a low height. And it is classified as 'unsafe'(Figure 8.(b)). If a part of the unit is empty without any points means there has missing platform or stairs. It is classified as 'unsafe'(Figure 8.(c))

Image Detection with YOLOv5

Yolov5 was used to distinguish 'safe' and 'unsafe' scaffold in 3D SLAM-BEV images. YOLOv5 was adopted because it shows high performance in both FPS and mAP, since is is a 1-stage image detection model. Also YOLO is divided into small, medium, and large size according to the size of the model. Since we run the model on the intel NUC connected to the Robot dog which has insufficient memory, so being able to select a model size is one of the advantage. In this paper, yolov5s is used for smooth model execution.

Model training was performed for 90 epochs, bat-size = 4 and fliplr = 1. The GPU used is one 2080Ti and model training took about an 1 hour. As for validation performance of the model is shown in Table 1.

	Precision	**Recall**	**mAP (0.5)**
Safe	0.898	0.868	0.937
Unsafe	0.971	0.754	0.849
All	0.934	0.811	0.893

Table 1. Result of validation set (88 images)

Model Test at Yonsei University Testbed

The experiment was conducted in the following way.

A 3D map is formed using SLAM while rotating a scaffold with a robot dog using teleoperation. At this time, the robot's trajectory is automatically saved every 4 seconds. It becomes a waypoint so the same trajectory can be repeated by itself.

Based on the created map and the set waypoint, the robot dog repeats the patrol with the same trajectory. The robot dog moves to the target point with a range of radius of 0.6m of the waypoint coordinates. If the radius of the waypoint is too small, the radius of 0.6m is appropriate because the robot dog rotates around to arrive at the correct coordinates. Detailed code is given in Figure 9(a).

At the same time, YOLOv5 started to detection. scaffold safety is detected using yolov5. Since the point cloud data is saved as a BEV image of 640*640 pixels, the pixel information of the image must be converted into m units. (Figure 8,b) At this time, the origin is the position where the robot dog first started making the map. yolo detects safe and unsafe based on the BEV image, and the result is output as shown in Figure 9.0 means safe and 1 means unsafe, and the coordinates are the coordinates of the center of the box.

```
if distance_to_goal > 0.6:
    if abs(angle_diff) > dist_angle_th(distance_to_goal):
        if angle_diff > 0:
            speed.linear.x = 0.0
            speed.angular.z = 0.2
        if angle_diff < 0:
            speed.linear.x = 0.0
            speed.angular.z = -0.2
    else:
        speed.linear.x = 0.4
        speed.angular.z = 0.0
else:
   print("moving to next goal")
   speed.linear.x = 0.0
   speed.angular.z = 0.0
   if i<len(goallist):
       i = i+1
   if i==len(goallist):
       i = 0
```

(a)

```
#Pixel value to cartesian
def pix2cart(xp,yp,sc_class,x0,y0):
    xc = x0 + 0.03125*(xp - 320)
    yc = y0 + 0.03125*(320 - yp)
    cart = np.array([xc, yc, sc_class])

    return cart
```

(b)

Figure 9. code for (a) waypoint range control (b) Cartesian coordinate transfor mation

```
scaffoldcart:  [[      11.85     -5.7344         1]
 [     12.944     -6.2657          1]
 [     14.038     -6.8907          1]
 [     19.444     -9.3594          0]
 [     15.896     -11.452          0]
 [     5.2626     -3.6643          1]
 [     7.4501     -10.289          0]
 [     7.1376     -3.6643          0]
 [     8.1376     -4.0393          0]]
already sent!
```

Figure 10. Result of scaffold detection

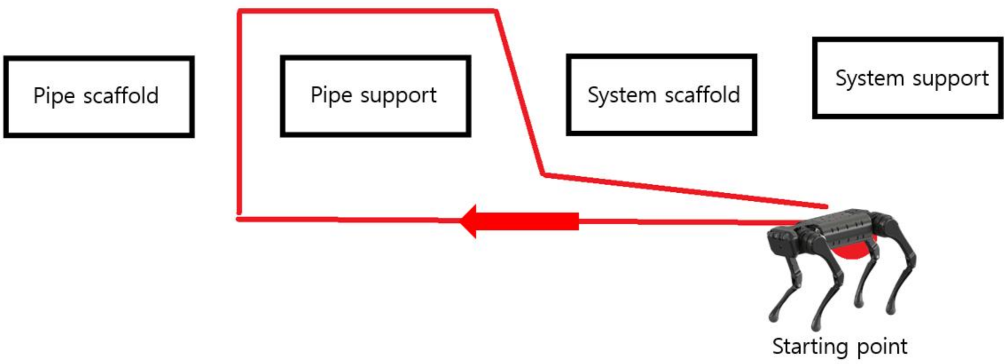


Figure 11. Path of Robot Dog

FINDINGS AND CONCLUSIONS

The trajectory we set is shown in Figure 11(BEV path). All data of support, syste m scaffold, and pipe scaffold can be acquired, and the path of the robot dog can be secured. The real-time detection result with yolo is shown in Figure . 9. The result of detection is shown in once per second. Figure 9 shows that a total n scaffold u nits were detected, of which n are safe and n are unsafe. The first and second colu mns in Figure 10 mean the center coordinates of the detection bounding box. It m eans how far away from the starting point of the robot dog. It was calculated by C artesian Coordinate transform. Detailed code is given in Fig 9 (b). In this study, co ordinates and detection results were output in an array format so that the robot dog can use the output results immediately in the future work. The step of visualizing the output coordinates in real time was omitted.

	Precision	Recall	mAP (0.5)
Safe	0.887	0.887	0.927
Unsafe	0.828	0.457	0.574
All	0.857	0.672	0.750

Table 2. Result of the test set (92 images)

In Figure 7(c), 'system support 1', 'system scaffold 2', 'pipe support 3', and 'pipe s caffold 4' are sequentially from the left. To confirm the quantified detection result, 92 frames were extracted from the stored point cloud BEV map video. During the 46 seconds that the robot dog automatically moved, two frames per second were uniformly extracted. The results for the 92 frames are shown in Table 2. Overall, t he recall is lower than the precision, and the performance for the unsafe class is lo w. The reason is that the unit of scaffold was not properly detected. Figure 12 sho ws the ground truth and prediction of the image. Unlike the ground truth in Figure 12(a), it can be seen that the system support is not properly recognized in Figure 1 2(b). The reason is that the system support is the smallest support. Because it is s mall, it is difficult to detect by unit, and density of point cloud data is low, so pixel information is stored differently from other scaffolds. Therefore, the left support c annot be properly judged, and this affects the performance of the unsafe class. An other reason for the low performance is shown in Figure 12(c). As the robot dog c ontinues to acquire points, the points become denser. So model classifies not only the scaffold but also the surrounding objects as part of the scaffold, and consequen tly affects the precision.

Also in the train data, there were 1,083 labeled safe and 486 unsafe labeled. It occ urs the class imbalance and also affected the poor performance of unsafe classes.

For images in which scaffolds or supports are fully visible, the results of correctly detecting each support and vector are shown in table 3. Since the path was set with the system scaffold as the main object, there were not many frames that the other structures be fully captured.

In conclusion, if the entire scaffold appears in the frame, it can be seen that the det ection success rate is very high.

	System support	System scaffold	Pipe support	Pipe scaffold
# of frames that fully captured	65	92	68	47
# of frames detected corectly	41	92	66	47
Accuracy (%)	63.1	100	97.1	100

Table 3. Percentage of correct detections in frames where the structure is perfect ly visible

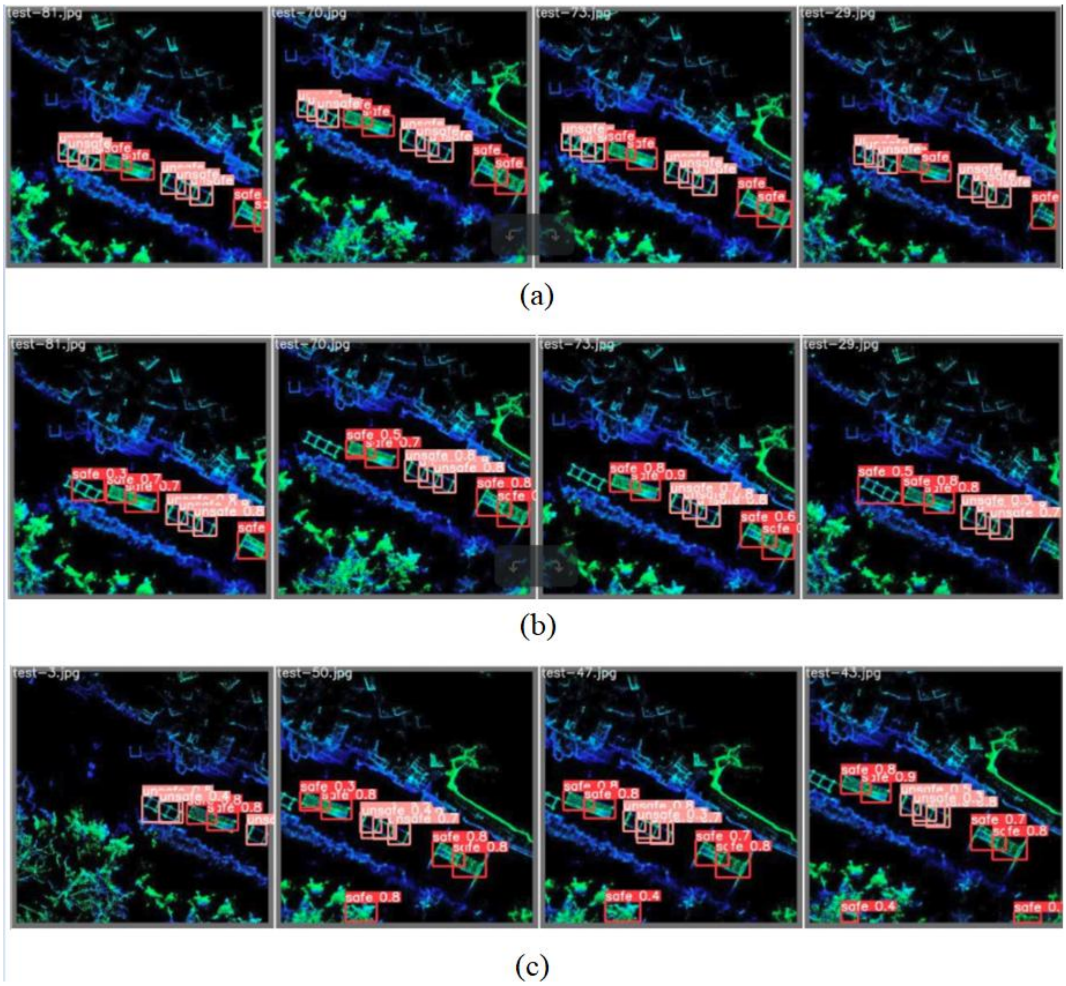


Figure 12. (a) ground truth (b) Incorrect detection of support (c) Incorrect detect ion of surroundings

In this study, We proposed a model that automatically acquires point cloud data us ing a robot dog and do safety checking of the scaffold using the BEV projection of 3D SLAM. Two sites were used for the training(i.e., Chungang University Testbe d, Ansung Testbed), and the Yonsei University testbed was used to test the model. We made a map using LIO-SAM and set the patrol so that the robot dog can auto matically acquire data. Scaffold safety was judged by the presence of platforms an d stairs. If the points are uniformly marked and filled with platform or stairs data, i t is considered safe. If there is an empty space without a point, there is no platform or stairs, so it is classified as unsafe. Unsafe scaffold images used for the train we re made by synthetic data extracted without a platform or stairs based on the heigh t information. Image detection was performed using YOLOv5. In the result, our m odel showed good performance with good accuracy.

This research has three contributions. First, automatically acquire data using a rob ot dog. The robot dog successfully created a 3D map using LIO-SAM through tele operation. Based on the trajectory of this path, it automatically creates a waypoint and repeats the patrol with the same trajectory and automatically acquires data. By automating data acquisition, it is possible to efficiently monitor construction sites.

Second, by converting point cloud data into images, scaffold safety can be judged in real-time. Point cloud data is not affected by light, contains more spatial inform ation than images, and is accurate. However, it takes a lot of time to process due t o the large amount of computation. In this study, image detection was attempted b y changing the point cloud to BEV images. By changing to an image with a fast pr ocessing speed, we were able to judge the safety of the scaffold in real-time.

Last, our proposed model is a general that can be applied to various sites. Since po int cloud data contains spatial information, it has certain characteristics even if the type of site changes, so it is easier to train a general model than using an image. B ecause scaffolds are standardized, scaffold units exist in other sites as well. Theref ore, our model can maintain satisfactory performance in various fields.

RECOMMENDATIONS

There are some limitations in this study: First, because there were no real unsafe s caffolds, all unsafe scaffold dataset was synthetic. Because synthetic datasets were generated by filtering out the cartesian heights of platforms or stairs, a slight grou nd-level difference makes the synthetic data noisy with partially removed platform s. This makes the labeling ambiguous, which confuses the model. Second, because construction sites have uneven surfaces, the optimal maximum/minimum height v aries on the surface's height. Because this study uses a limited range in the point c loud map for generalization between each floor, platforms on higher ground level might be omitted. Third, it is hard to tell whether it's an unsafe scaffold or looks u nsafe just because there's not enough point cloud collected.

To solve this limitation, the first thing to do is to divide the detection algorithm int o two stages: Detection of scaffolds and classification of safety levels. By dividing the detection algorithm, the dataset could be more properly used and the limitatio ns of using synthetic data would only affect the classification part. Also, re-trainin g with real unsafe scaffold data would make the model more exact. Making a visu alization tool to provide more intuitive information for the users would also be nee ded.

ACKNOWLEDGEMENTS

This research has been supported by the literature and theoretical materials of excellent researchers in the field of construction automation and Construction Automation course offered by Prof. HyoungKwan Kim.

REFERENCES

Adibhatla, V. A., Chih, H. C., Hsu, C. C., Cheng, J., Abbod, M. F., & Shieh, J. S. (202 1). Applying deep learning to defect detection in printed circuit boards via a newest m odel of you-only-look-once.

Balaguer, C., & Abderrahim, M. (Eds.). (2008). Robotics and automation in constructi on. BoD–Books on Demand.

Bassier, M., & Vergauwen, M. (2020). Unsupervised reconstruction of Building Infor mation Modeling wall objects from point cloud data. Automation in Construction, 12 0, 103338.

Beltrán, J., Guindel, C., Moreno, F. M., Cruzado, D., Garcia, F., & De La Escalera, A. (2018, November). Birdnet: a 3d object detection framework from lidar informatio n. In 2018 21st International Conference on Intelligent Transportation Systems (ITS C) (pp. 3517-3523). IEEE.

Braun, A., S. Tuttas, A. Borrmann, and U. Stilla. 2020. “Improving progress monitorin g by fusing point clouds, semantic data and computer vision.” Autom. Constr. 116

Chen, J., Kira, Z., & Cho, Y. K. (2019). Deep learning approach to point cloud scen e understanding for automated scan to 3D reconstruction. Journal of Computing in Civil Engineering, 33(4), 04019027.

Corke, P., Winstanley, G., Dunbabin, M., & Roberts, J. (2003). Dragline automation: Experimental evaluation through productivity trial. In Field and service robotics (pp. 4 59-468). Springer, Berlin, Heidelberg.

Fang, Q., H. Li, X. Luo, L. Ding, H. Luo, T. M. Rose, and W. An. 2018a. “Detecting no n-hardhat-use by a deep learning method from far-field surveillance videos.” Autom. Constr. 85 (Jan): 1–9.

Fang, W., L. Ding, H. Luo, and P. E. Love. 2018b. “Falls from heights: A computer vis ion-based approach for safety harness detection.” Autom. Constr. 91 (Jul): 53–61.

Fathi, H., Dai, F ., & Lourakis, M. (2015). Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenge s. Advanced Engineering Informatics, 29(2), 149-161.

Francis Baek, Daeho Kim, Somin Park, 2022 “Conditional Generative Adversarial Ne tworks with Adversarial Attack and Defense for Generative Data Augmentation”, volu me 36, issue 3

Girshick, R. 2015. “Fast R-CNN.” In Proc., 2015 IEEE Int. Conf. on Computer Vision (ICCV), 1440–1448. New York: IEEE.

Ha., Q. P., Nguyen Q. H.; Rye D. C. & Durrant-Whyte, F. (2000). Impedance control o f a hydraulically-actuated robotic excavator, Journal of Automation in Construction, V ol. 9, No 5, 421-435, ISSN 0926-5805

Hasegawa, Y. (2006). Construction Automation and Robotics in the 21th century, 23r d International Symposium on Robotics and Automation in Construction (ISARC’06), Japan, October 2006, Tokyo

Hongjo Kim, Kinam Kim, and Hyoungkwan Kim,2016, “Vision-based object-centric saf ety assessment using fuzzy inference: monitoring struck-by accidents with moving obj ects”, Journal of Computing in Civil Engineering, Volume 30 Issue 4

Hyojoo Son, Hyunchul Choi, Hyeonwoo Seong, and Changwan Kim, 2019, “Automati on in Construction”, volume 99, 27-38

Kim, P., Chen, J., & Cho, Y. K. (2018). Automated point cloud registration using vis ual and planar features for construction environments. Journal of Computing in Civi l Engineering, 32(2), 04017076.

Kim, P., Chen, J., & Cho, Y. K. (2018). SLAM-driven robotic mapping and registrati on of 3D point clouds. Automation in Construction, 89, 38-48.

Kim, P., Park, J., Cho, Y. K., & Kang, J. (2019). UAV-assisted autonomous mobile r obot navigation for as-is 3D data collection and registration in cluttered environmen ts. Automation in Construction, 106, 102918.

Kim, J., Chung, D., Kim, Y., & Kim, H. (2022). Deep learning-based 3D reconstructi on of scaffolds using a robot dog. Automation in Construction, 134, 104092.

Kulchandani, J. S., and K. J. Dangarwala. 2015. “Moving object detection: Review of recent research trends.” In Proc., 2015 Int. Conf. on Pervasive Computing (ICP C), 1–5. New York: IEEE.

Lei, L., Y. Zhou, H. Luo, and P. E. Love. 2019. “A CNN-based 3D patch registration a pproach for integrating sequential models in support of progress monitoring.” Adv. En g. Inf. 41

Man-Woo Park, Ioannis Brilakis,2012, "Construction worker detection in video frames for initializing vision trackers" Automation in Construction,Volume 28, 15-25.

Milad Memarzadeh, Mani Golparvar-Fard, Juan Carlos Niebles, 2013 “Automated 2D detection of construction equipment and workers from site video streams using histog rams of oriented gradients and colors” Automation in Construction, volume 32, 24-37

Radej, J., & Guid, N. (1993). Reviews by. Journal of computing and information techn ology, 1(3), 225-229.

Ratay, R. (1996). Handbook of Temporary Structures in Construction: Engineering St andards, Designs, Practices and Procedures, McGraw-Hill, New York.

Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. 2016. “You only look once: Unifi ed, real-time object detection.” In Proc., 2016 IEEE Conf. on Computer Vision and Pa ttern Recognition (CVPR), 779–788. New York: IEEE.

Rehg, J. A. (1995). Computer-integrated manufacturing. Prentice-Hall, Inc..

Ren, S., K. He, R. Girshick, and J. Sun. 2017. “Faster R-CNN: Towards real-time obj ect detection with region proposal networks.” IEEE Trans. Pattern Anal. Mach. Intell. 39 (6): 1137–1149.

Seongdeok Bang, Francis Baek, Somin Park, Wontae Kim, and Hyoungkwan, Kim, 2 020, “Image augmentation to improve construction resource detection using generati ve adversarial networks, cut-and-paste, and image transformation techniques”, Auto mation in Construction, volume 115

Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., & Rus, D. (2020, October). Li o-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 5135-5142). IEEE.

Shang, Z., & Shen, Z. (2017). Real-time 3D reconstruction on construction site using visual SLAM and UAV. arXiv preprint arXiv:1712.07122.

Tam, C. M., Zeng, S. X., & Deng, Z. M. (2004). Identifying elements of poor constructi on safety management in China. Safety science, 42(7), 569-586.

Weili Fang, Botao Zhong, Neng Zhao, Peter E. D. Love, Hanbin Luo, Jiayue Xue, and Shuangjie Xu, 2019 “A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network”, Advanced Engineering Informatics, volume 49, 170-177

시스템 비계 안전작업 지침(KOSHA GUIDE C-32-2020)