Tech Science Press

check for updates

Intelligent Risk-Identification Algorithm with Vision and 3D LiDAR Patterns at Damaged Buildings

Dahyeon Kim[1], Jiyoung Min[1], Yongwoo Song[1], Chulsu Kim[2] and Junho Ahn[1],*

[1]Computer Information Technology, Korea National University of Transportation, Chungju, 27469, Korea
[2]Department of Railway Vehicle System Engineering, Korea National University of Transportation, Uiwang, 16106, Korea
*Corresponding Author: Junho Ahn. Email: jhahn@ut.ac.kr

Abstract: Existing firefighting robots are focused on simple storage or fire suppression outside buildings rather than detection or recognition. Utilizing a large number of robots using expensive equipment is challenging. This study aims to increase the efficiency of search and rescue operations and the safety of firefighters by detecting and identifying the disaster site by recognizing collapsed areas, obstacles, and rescuers on-site. A fusion algorithm combining a camera and three-dimension light detection and ranging (3D LiDAR) is proposed to detect and localize the interiors of disaster sites. The algorithm detects obstacles by analyzing floor segmentation and edge patterns using a mask regional convolutional neural network (mask R-CNN) features model based on the visual data collected from a parallelly connected camera and 3D LiDAR. People as objects are detected using you only look once version 4 (YOLOv4) in the image data to localize persons requiring rescue. The point cloud data based on 3D LiDAR cluster the objects using the density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm and estimate the distance to the actual object using the center point of the clustering result. The proposed artificial intelligence (AI) algorithm was verified based on individual sensors using a sensor-mounted robot in an actual building to detect floor surfaces, atypical obstacles, and persons requiring rescue. Accordingly, the fused AI algorithm was comparatively verified.

Keywords: Three-dimension light detection and ranging; vision; risk identification; damaged building; robot

1 Introduction

Despite the rapid development in technology, buildings collapse and related casualties caused by natural or man-made disasters occur every year. Because buildings that are aged or not regularly maintained are significantly affected by minor external environmental changes, it is difficult to determine the stability of structures over time, which may lead to the collapse of buildings. A report by the American society of civil engineers reported [1] that in the United States (U.S.), water pipes burst every two minutes, 43% of public roads are in inferior conditions, and 7% of the bridges, which are used by millions of commuters each day, are "structurally defective". For example, the Hard-rock hotel in New Orleans

collapsed during construction because of a lack of construction stability. Three construction workers were killed, and dozens injured. In 2022, the floor of a building in Hackney Wick in the United Kingdom (UK) collapsed, causing serious injuries to three people and minor injuries to 13 people [2]. In 2017, a fire broke out in a high-rise building in Tehran, the Iranian capital, which caused the building to collapse, thus killing 20 firefighters [3]. If firefighters are sent in for search and rescue operations without a prior understanding of the disaster situation in a collapsed building, casualties may occur. It may also cause inefficiencies in search and rescue operations. Most of the robots currently used in disaster situations can only perform simple fire suppression and observe the inside of a disaster building using only cameras or existing building drawings. However, it has difficulties in clearly assessing the disaster situation, and the ineffective response at the disaster site may lead to enormous damage in the future. At disaster sites, it is necessary to detect deformed and damaged obstacles and internal debris from building collapses rather than identifiable objects. It is also necessary to allow firefighters to determine whether the space is searchable and to detect rescue requests, which is the most important task at disaster sites. Therefore, this study aims to identify obstacles and rescue requests and to understand disaster sites for efficient and rapid search and rescue operations within buildings that have been destroyed by disasters.

To identify a disaster site, it is necessary to detect the collapsed areas, obstacles, and persons requiring rescue in the disaster site and estimate the location of the detected objects in the disaster-hit building. In such a case, obstacles at the disaster site refer to debris caused by the collapse and atypical obstacles that are difficult to identify as they are deformed from the original shape of objects before the collapse. Research on detecting atypical obstacles usable at disaster sites is necessary owing to the challenges of detecting such obstacles with a conventional supervised learning method. To this end, an algorithm incorporating the fusion of a camera and low-cost three-dimension light detection and ranging (3D LiDAR) [4] is proposed. Although a camera can accurately detect obstacles and persons requiring rescue, it can be difficult to use at actual disaster sites owing to the unique characteristics of a disaster environment, including light and gas. Although 3D LiDAR is relatively less affected by the external environment, it may still be difficult as it is an expensive device costing more than 10 thousand dollars. We conducted this study using a low-cost, short-range 3D LiDAR of less than 2 thousand dollars. The point cloud data collected by the high-cost long-range 3D LiDAR are composed of dense dots; thus, the shape of an object can be recognized as an image. However, it is difficult to recognize the distinct shape of an object from the point cloud data collected by the low-cost short-range 3D LiDAR. This study identified the shape of an object based on the camera and improved the detection accuracy by fusing it with the object detected by 3D LiDAR. Deep learning-based floor segmentation was then applied using the camera, and obstacles were detected by analyzing the segmentation patterns in areas excluding the floor surface. To reduce human casualties or key elements at a disaster site, this study detected obstacles and people through a deep learning-based object detection algorithm. The 3D LiDAR detects clustering-based objects and uses the central point of clustering to find the distance to an object. The camera and 3D LiDAR were installed facing the same direction, and a label was given to the 3D LiDAR cluster by matching it with the nearest cluster. Here, the accuracy of the algorithm for detecting an obstacle and the person seeking rescue was 0.93. This study utilized a low-cost 3D LiDAR, camera, and entry-level robot, thus reducing the replacement and repair cost burden in the field. The detected obstacles and persons seeking rescue will enable safe, efficient, and rapid rescue and search operations when firefighters are sent into disaster-hit buildings.

The remainder of the paper is organized as follows. Section 2 introduces various studies using cameras, two-dimension light detection and ranging (2D LiDAR), and 3D LiDAR to identify disaster sites. Section 3 designs the proposed system and algorithm. Subsequently, a detection algorithm that determines the feasibility of entering a disaster site and detects obstacles and persons in need of rescue is explained. Additionally, the algorithm for estimating the location of a disaster site from the actual detection result by

fusing the detection result with the LiDAR is explained in detail. Section 4 presents the results of experiments conducted by installing sensors on actual robots. Section 5 summarizes the study and describes the direction of future and related research.

2 Related Work

Research [5] is related to detecting an object using the Viola–jones [6] and you only look once version 3 (YOLOv3) [7] algorithms for image data obtained through a camera and tracking the object using the median flow [8] and correlation tracking methods. In the image obtained by the front camera, an object is detected every 10th frame through object detection algorithms, and in other frames, the location of the object detected in the previous frame is tracked by the tracker. Research [9] proposed detecting an object using a stereo camera and finding the distance to the object. It proposed a method of detecting an object using you only look once (YOLO) and finding the distance to the object using the angle between the direction of the object detected by each camera and camera lens axis. Using the proposed method, the distance to the object can be obtained using only a camera and no LiDAR sensor. Research [10] designed a convolutional neural network (CNN) based multi-view smart camera for real-time object detection. Using a CNN utilizing a multi-field programmable gate array (multi-FPGA) accelerator architecture, the recognition rate was shown to be equivalent even if the number of parameters is reduced in the existing CNN. Furthermore, various studies [11–13] have been conducted using deep learning networks to detect various objects by training the model based on the features of objects. Research [14] installed an omnidirectional camera on a drone to detect an object in the air. The research used Ladybug5 [15] as the omnidirectional camera. The processing region from the image acquired by the camera driver module using the Region of interest (ROI) extractor module was extracted. Furthermore, the experimental results showed that the detection performance was higher than that of faster regional convolutional neural network (faster R-CNN) [16] because the Single shot multibox detector (SSD) method [17] was used for object detection. The experimental results showed higher detection performance than faster R-CNN when using the SSD method. Research [18] fused algorithms to capture fast-moving drones. The contiguous outlier representation via online low-rank approximation (COROLA) algorithm [19] was used to detect small moving objects present in the scene. The CNN algorithm was used to enable accurate drone recognition in a wide range of complex scenes. Research [20] used a monocular camera to detect obstacles and estimate the distance to the obstacles. The research separated the bottom surface after image segmentation was completed through the mask regional convolutional neural network (mask R-CNN) [21] algorithm. If the shape of image segmentation was normal, a specific point of the image was designated as a point of interest to find a line for the point, and an obstacle was detected by analyzing the pattern of the line. The height of the detected obstacle was estimated using image processing. The distance to the obstacle was then estimated using the polynomial regression function obtained from the pixel value of the line up to the obstacle and the actual value between the robot's starting point and the obstacle. Research [22] improved the performance of current state-of-the-art single-shot detection networks using red, green, blue (RGB) and depth images fused with depth data on RGB image data for camera-based object detection. A fuse layer architecture was constructed to fuse the two data sources. The research proved that RGB and depth fusion increases both the conventional detection accuracy and positioning performance of object detection regardless of the depth acquisition method.

Various studies have been conducted on the convergence of cameras and 2D LiDAR. There was a study [23] that created a map containing more specific environmental information than a map created using the existing 2D LiDAR by fusing the map obtained through the RGB image and its corresponding depth image (RGB-D) camera and the map obtained through the 2D LiDAR. The map finally generated using this method had similar results to those generated using 3D LiDAR. Another study combining 2D LiDAR and a camera [24] proposed an algorithm to detect lanes in areas obscured by objects such as

vehicles. Objects on the road were detected via the 2D LiDAR. A binary bird's eye view (BEV) image was acquired from the camera data, and a modified BEV image was generated by removing the noise generated by the object. The algorithm increased the lane detection accuracy and robustness. Research [25] conducted road detection by fusing a camera and 2D LiDAR based on a single conditional random field (CRF) framework [26]. A dense result was obtained by upsampling the sparse LiDAR data in the camera image area. Road detection via a camera was performed using previous research on semantic segmentation. Research [27] used a 2D LiDAR and RGB-D camera to semantically detect objects. A grid map was generated and the robot trajectory was obtained using laser-based simultaneous localization and mapping (SLAM). The object point cloud was then fused with the creation result through camera calibration and mask R-CNN based object detection. Research [28] fused a 2D LiDAR and camera to detect an object and display it on a SLAM-generated map. Objects are found by implementing the SSD object detection algorithm to the camera's image data. The 2D LiDAR uses SLAM to create maps. Here, the detected object was added to the map using the map update method proposed by the research. Research [29] used various sensors, such as 2D LiDAR and home cameras, to detect people in abnormal situations. The movement patterns of people in motion were analyzed through 2D LiDAR, and points were detected for both when they were moving and not moving. Additionally, people as objects were detected through YOLOv4 in the image data of home camera.

There are numerous studies on object detection by fusing a camera and 3D LiDAR. Research [30] detected objects by segmenting the point cloud and image values of 3D LiDAR. The research proposed a methodology for detecting a vehicle by fusing an object obtained through RGB-based CNN from an image and a 3D LiDAR point cloud and localizing it through binary BEV point cloud projection. A study [31] combining 3D LiDAR and a camera to detect an object, obtained an object from a point cloud to 3D LiDAR through pointnet++, and fused it with the object information obtained through the camera to achieve final 3D object detection through a region proposal network. Research [32] investigated the unsupervised domain adaptation (UDA) methodology [33] based on the fact that modern detectors can significantly degrade performance between domains. It proposed the semantic point generation (SPG) method to identify the point cloud quality acting as a cause of domain performance degradation during detection. Pointpillars [34], Pointvoxel regional convolutional neural network (PV R-CNN) [35], and SPG were comparatively tested to verify their applicability in various weather conditions. Most LiDAR-based 3D detections are hand-crafted. Hand-crafting is limited by a bottleneck that prevents the effective use of data required for three-dimension (3D) shape and detection work. Research [36] studied cloud 3D detection based on an end-to-end trainable deep network [37] that can solve such a challenge. The Voxelnet methodology proposed in the research can effectively obtain 3D shapes by directly accessing the 3D point. Research [38] studied a more efficient and accurate 3D object detection. It proposed a range sparse net (RSN) model method that predicts foreground points from a range image and detects objects by applying sparse convolutions. Voxelization was performed on the points for the selected object by executing the sparse point feature extraction (SPFE) algorithm. Additionally, research [39] studied 3D LiDAR for object detection in embedded systems. It proposed a visual-LiDAR that fuses a camera and 3D LiDAR for an efficient perceptual system and investigated 3D object detection and tracking. The LiDAR point cloud was measured as a pose value trackable through ground removal and clustering. Localized bounding boxes and class data were simultaneously provided through the YOLOv3 algorithm. This was conducted using the Karlsruhe institute of technology and toyota technological institute (KITTI) datasets [40].

3 Algorithms

This study detects and localizes obstacles and persons requiring rescue at disaster sites using cameras and LiDAR for efficient and rapid rescue and search in disaster situations. This section describes the

entire research system for collecting and processing data. It then describes the detection algorithm for obstacles and victims, such as buried people, and the localization algorithm based on the detection.

3.1 Data Collection and Processing

A smartphone and LiDAR sensors are installed on a robot. Data are then collected using these sensors when the robot moves. Fig. 1 shows the overall data flow and processing procedure of the system.
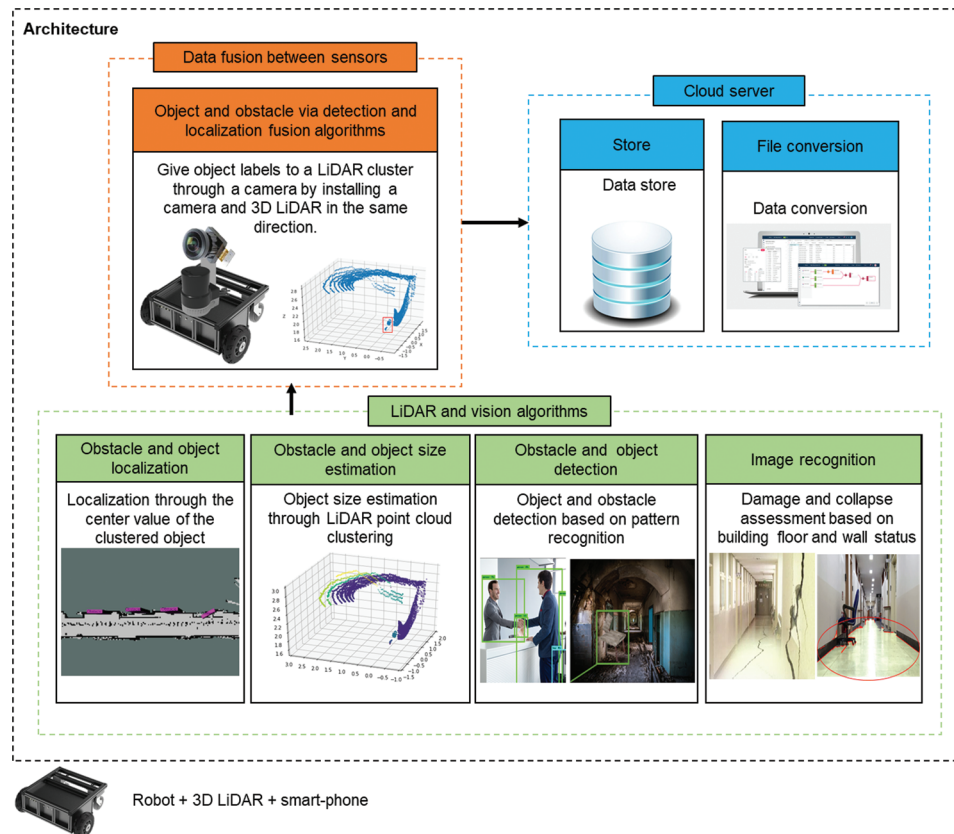

Figure 1: Overall system configuration of the proposed methodology

The shape of objects and obstacles can be identified as the image data collected via smartphone has RGB image values. The collapse and damage status of the floor and wall using the image data can also be identified. Fig. 2 shows the algorithm used for the 3D LiDAR that estimates the size and location of detected objects and obstacles.

The algorithm is fused after parallel processing with two algorithms that process camera and 3D LiDAR data. Point cloud data collected through 3D LiDAR carry x-axis, y-axis, and z-axis 3D vector values. Objects are clustered and classified per object after separating them from spaces by clustering the point values of 3D LiDAR. Subsequently, the center point of the classified clusters is determined, and the range value of the point closest to the center point is stored. The camera provides image-based RGB image frame data while two algorithms are processed in parallel. Here, a deep learning-based object detection algorithm is utilized for rapid rescues, which is the most important factor in a disaster situation. Objects at disaster sites are usually deformed into abnormal or difficult-to-recognize shapes. Furthermore, debris and collapsed structures caused by the collapse are present within the building. A study was conducted to

define such atypical objects as obstacles and to detect obstacles in image data. The floor surface is first detected to determine whether the space can be entered. If the space is deemed accessible, floor patterns are analyzed to detect obstacles. If the space is deemed inaccessible, its data is directly sent to the server to notify the disaster situation control center of its inaccessibility. When a person is detected through an obstacle and object detection algorithm based on the analyzed pattern, the camera and 3D LiDAR are installed on the robot to face an identical direction to match the cluster's center point extracted from the 3D LiDAR. The closest point in the 3D LiDAR cluster is matched following the position and size orders of the image's bound box. Labels are given to the point cloud cluster through the matching result based on the camera's object detection result. A person and obstacle can be localized through the range value. Here, the 3D LiDAR can detect a wide range from floor to ceiling. It is possible to check accessibility, as well as the damage and collapse status of floors, walls, and ceilings by fusing the damage and collapse situation determined by the camera and the detection result from the 3D LiDAR. The detected results are transmitted to the main cloud server controlling the disaster situation. Such data enable the identification of disaster sites within buildings and rapid deployment of rescuers for search and rescue operations.
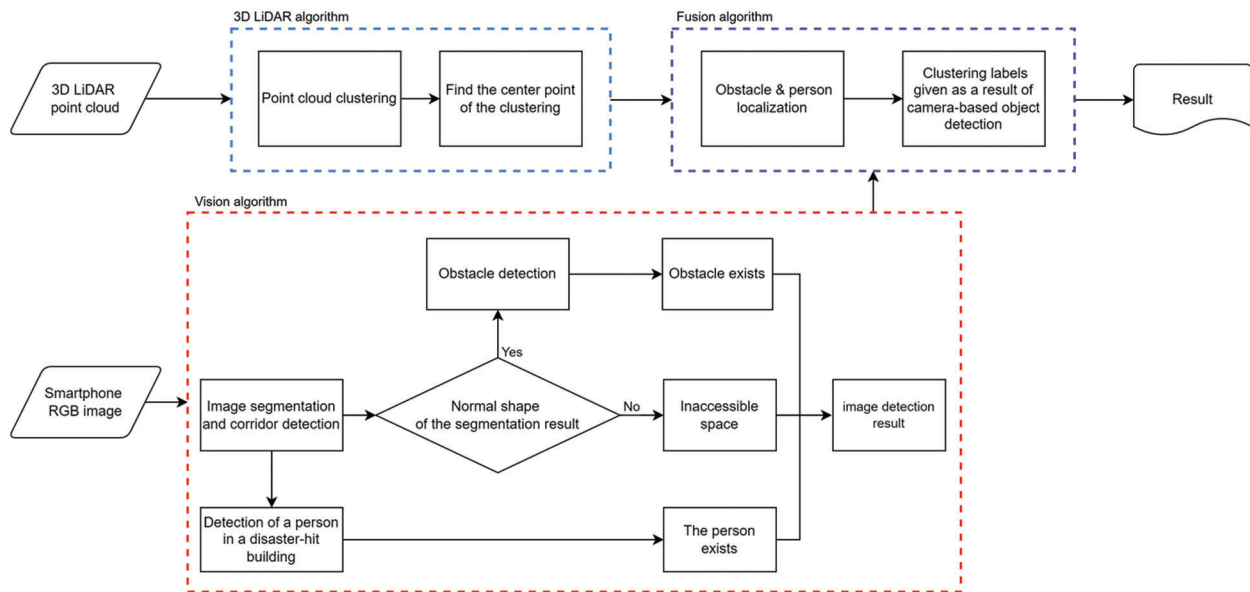

Figure 2: Flow chart of the proposed algorithm

3.2 Obstacle and Object Detection

Using cameras, two algorithms are used to detect obstacles and persons requiring rescue, such as those buried in disaster-hit buildings. An algorithm for detecting obstacles and determining accessibility was studied in the previous research [20]. Fig. 3 shows the overall flow diagram of the detection algorithm.

To determine accessibility, floor segmentation is analyzed through mask R-CNN in camera-based image data. The segmentation of the unobstructed, searchable, and accessible floor surface has a normal triangular shape, whereas that of an obstructed, unsearchable, and inaccessible floor has an abnormal shape, as shown in Fig. 4.

If the corridor is determined to be accessible, then the presence of obstacles can be determined by analyzing the edge of the segmentation through the line segment detector [41]. If the obstacle pattern is present on the floor, it is determined that obstacles are present, and their sizes are estimated through the camera. The area with the obstacle is obtained for the estimation. The area without obstacles is

determined based on the farthest point among the points obtained through the floor pattern. The obstacle characteristics are clearly expressed through the morphology operation [42], where the obstacles are present on the other side. The morphology operation is a technique that analyzes the objects' shape or structure from image data. This method extracts the main obstacles characteristics by performing opening and then closing computations. The uppermost characteristic point among the main characteristics is determined as the height of the obstacle. With this, obstacles are detected, and their shapes are determined. The second algorithm detects the buried persons through the you only look once version 4 (YOLOv4) algorithm [43] for rescue operations, which is the most important factor at disaster sites. The YOLOv4 model and Common Object in Context (COCO) datasets [44] trained with human labels detect even small indications of people in the image with high accuracy. However, the accuracy of using a camera to determine accessibility and the presence of obstacles and persons in need of rescue varies depending on the disaster environment. In particular, accurate detection cannot be performed with image data alone in a dark situation owing to a lack of light or clouds caused by gas. Because this study assumes a disaster situation, there is a high probability that the building's power system will fail and there will be no lighting; however, disaster situations can occur at any time of the day or night. Additionally, the image data cannot make accurate detections if there is dust caused by the building collapse or fire smoke. Therefore, accurate detection must be achievable even without light. To solve such challenges and to more accurately localize obstacles and persons seeking rescue, a LiDAR fusion algorithm is proposed.
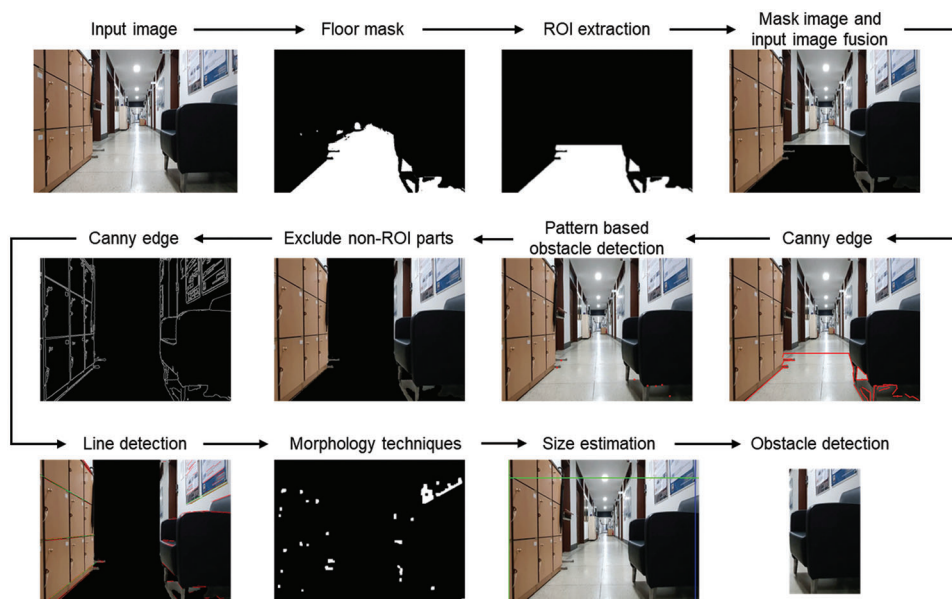

Figure 3: Flow diagram of obstacle detection algorithm

3.3 Obstacle and Object Localization

3.3.1 2D LiDAR and Camera Fusion

We investigated the fusion of a 2D LiDAR and camera [45] to accurately localize obstacles. The obstacles were detected, and disaster site maps were generated using the 2D LiDAR based on the proposed algorithm. The 2D LiDAR utilized the hector SLAM methodology [46] to locate the robot and generate a map, as shown in Fig. 5.

Figure 4: Example of detecting a normal corridor and a corridor blocked by obstacle through floors in an indoor space

Figure 5: Result of the hector SLAM via 2D LiDAR

Hector SLAM is a method of performing real-time matching based on the scanned point cloud that allows for the creation of a map using 2D LiDAR. In this case, the map update algorithm proposed by segmentation SLAM [28] was used to display the detected obstacles on the map. The map update algorithm displays the object on the map by matching the 2D LiDAR-based scan result on the current map with the bounding box obtained by the object detection algorithm. The size of the detected obstacle was estimated and updated on the map. A disaster site map was created by indicating the accessibility of a passage through the collapsed floor, as shown in Fig. 6. The red line is the scan result, and the gray part is the mapping result.

The estimation accuracy of obstacle localization using 2D LiDAR is higher than when using only the camera. Nevertheless, existing methods still depend solely on the camera for obstacle detection, making them vulnerable to the environmental factors in disaster-hit buildings, such as light and gas. Additionally, size estimation using a robot moving on uneven surfaces at disaster sites may yield inaccurate results as estimating the obstacle size is also based on image data. To solve challenges induced by the specificity of disaster situations, this study detects objects by clustering point cloud data of 3D LiDAR rather than relying only on a camera when estimating obstacles.
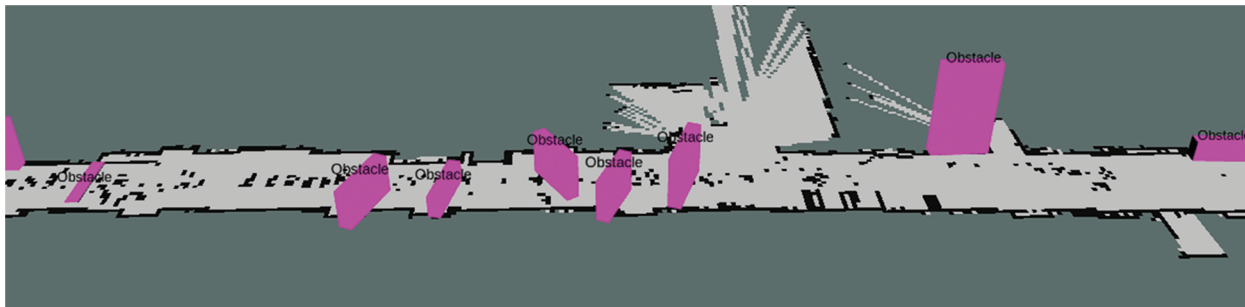
Figure 6: Result of maps

3.3.2 3D LiDAR and Camera Fusion

The results of the point cloud data based on 3D LiDAR are shown in Fig. 7.
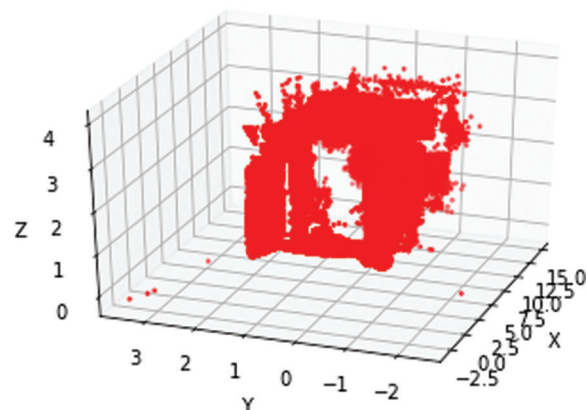

Figure 7: Example of a 3D LiDAR point cloud scan

Of the 1,003,006 points collected, only 3000 of the corresponding data are coordinates walls are separated from objects by clustering 3D LiDAR-based point cloud data through the density-based spatial clustering of applications with noise (DBSCAN) algorithm [47]. Here, the DBSCAN clustering is applied to the point cloud classified as an object to classify the separated results once again. Fig. 8 shows the result of classifying the point cloud determined as an object. The epsilon value which is the hyper parameter value of DBSCAN is 0.8. Then, the center point of the point cloud clustering result categorized per object is determined. The distance to the object is estimated based on the value of the point closest to the x-axis, y-axis, and z-axis coordinates of the corresponding center point. This study used a low-cost, short-range, single-channel 3D LiDAR costing less than $1,077 rather than a long-range scan-capable, multi-channel 3D LiDAR costing tens of thousands to several tens of thousands of dollars. This study was conducted using a low-cost 3D LiDAR to reduce the cost burden as the 3D LiDARs installed on robots may be damaged in disaster sites, or multiple 3D LiDARs may need to be used to cover a larger disaster area. However, unlike high-cost 3D LiDAR, low-cost 3D LiDAR makes it difficult to accurately identify the shape of an object. Therefore, fusion with a camera is proposed because it is necessary to distinguish between an obstacle and a person.
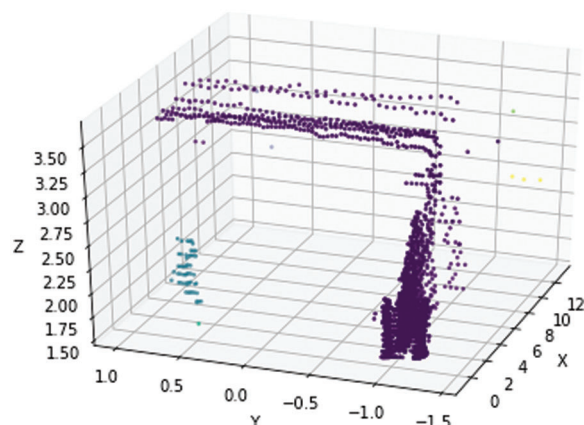
Figure 8: Example of object clustering in a 3D point cloud

When installing the 3D LiDAR on the robot, it should be placed on a vertical line with the camera so that both can detect and localize the same object based on the direction of movement of the robot. If the camera determines that there is an object, such as an obstacle, the 3D LiDAR checks the presence of an object cluster in that direction. In this case, the result of detecting the obstacle and the requestor for rescue in the camera is shown in Fig. 9. If both the camera and 3D LiDAR determine that there is an object, the distance to the object is estimated using the x-axis value among the point coordinates of the cluster's center point. The cluster closest to the center of the image is then selected and labeled as an obstacle or a person in need of rescue, based on the camera's estimation. This study can directly specify the initial position of the robot and the starting point inside the building. This study was conducted based on previous research [48] to reset the error caused by the robot's movement. The current robot position is estimated based on the distance the robot has moved to localize obstacles and persons seeking rescue in the building, and their locations can be obtained by fusing the LiDAR's spatial scan results.
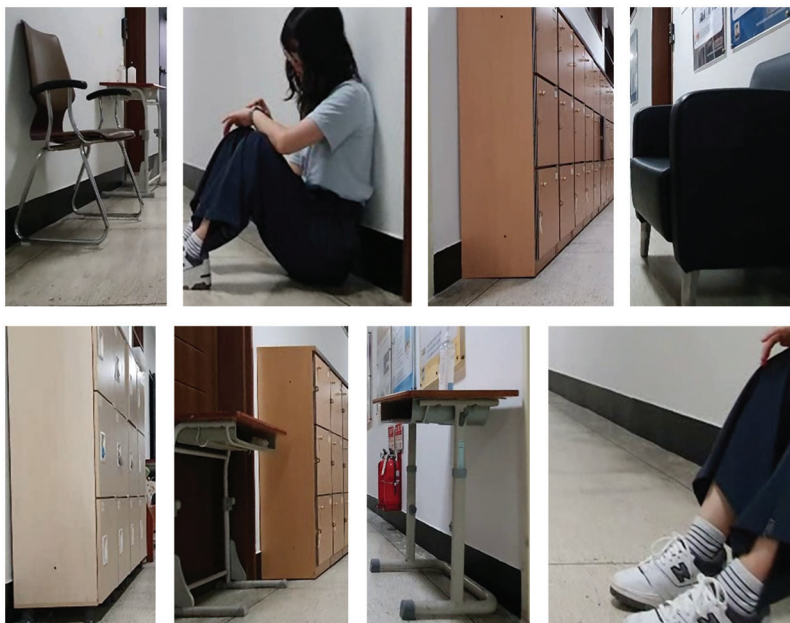

Figure 9: Result of detecting obstacles and person in need of rescue using pattern analysis and resulting images for estimating the height of obstacles and person in need of rescue

4 Experiments and Results

As shown in Fig. 10, sensors were installed on the actual robot to detect and localize obstacles. The camera used here is Samsung's S10 [49], and the 3D LiDAR is the Yujin Robot's YRL3-v10 model [4]. The robot used in the experiment is the OMO Robot's R1 model [50].

Figure 10: Setup of robot for experiments

The experiments were carried out to detect obstacles and rescue requests and estimate their locations using a camera, 2D LiDAR, the combination of camera and 2D LiDAR, 3D LiDAR, and the combination of camera and 3D LiDAR while walking around real buildings on the university campus. In the building, there were general objects such as chairs, desks, and lockers, as well as blackboards, sofas, water purifiers, and newsstands. In order to simulate a disaster situation, the objects were knocked down and placed in unusual shapes. Table 1 shows the results of the experiment.

Table 1: Obstacles per sensor and accuracy of the detection of persons requiring rescue

Algorithm	Modal	Accuracy
Obstacle and object detection	Camera	0.87
	Camera + 2D LiDAR	0.56
	Camera + 3D LiDAR	0.93

The 3D LiDAR and camera fusion has the highest obstacle detection accuracy of 0.93. In the past, only tomographic spatial data could be obtained using a 2D LiDAR and camera. However, in the case of a 3D LiDAR-based point cloud, as shown in Fig. 11, the space of the current single floor can be identified in detail up to the ceiling, and it is possible to predict even other sites because the ceiling of the current floor becomes the floor of the upper floor.

Assuming various situations within the disaster site, the possibility of detection and estimation is shown in Table 2. Obstacles and persons requiring rescue were marked with 'O' if detection and localization were possible and 'X' otherwise.

Using a 3D LiDAR and camera, it was possible to detect and localize obstacles and persons requiring rescue in more situations than other models.
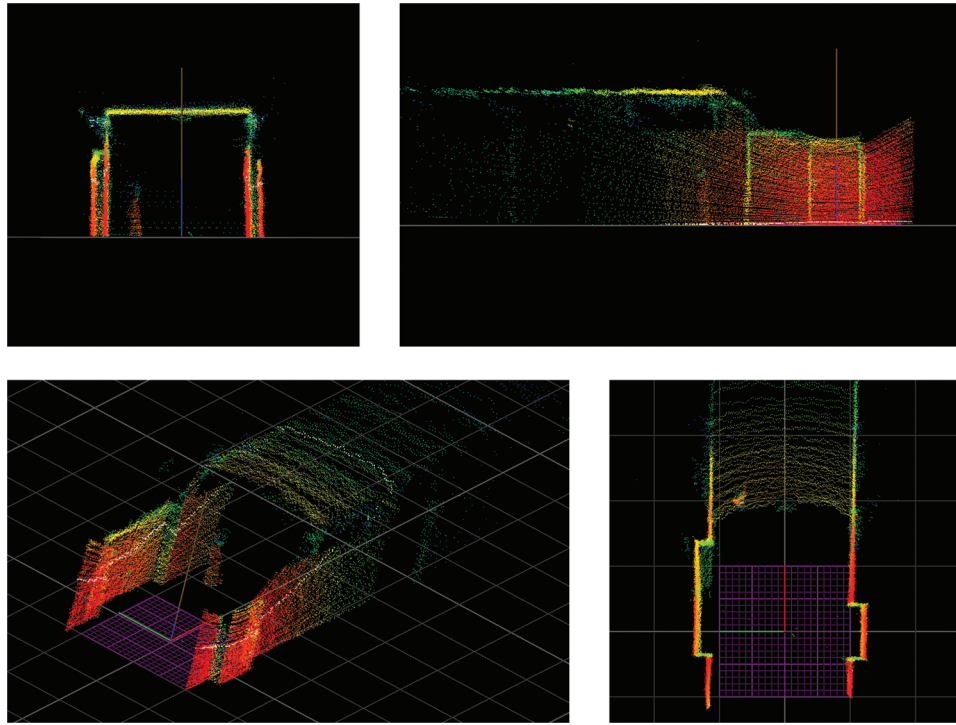
Figure 11: Three-dimensional LiDAR point cloud results

Table 2: Detection and localization of obstacles and persons requiring rescue per sensor based on the situation

Situation/State	Algorithm	Modal	Availability
Building collapse disaster situation that occurred during the day	Obstacle and object detection	Camera	O
		Camera + 2D LiDAR	O
		Camera + 3D LiDAR	O
	Obstacle and object localization	Camera	O
		Camera + 2D LiDAR	O
		Camera + 3D LiDAR	O
Building collapse disaster situation that occurred during the night	Obstacle and object detection	Camera	X
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O
	Obstacle and object localization	Camera	X
		Camera + 2D LiDAR	O
		Camera + 3D LiDAR	O
Building interior full of smoke due to fire	Obstacle and object detection	Camera	X
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O
	Obstacle and object localization	Camera	X
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O

(Continued)

Table 2 (continued)

Situation/State	Algorithm	Modal	Availability
Multiple overlapping atypical obstacles due to the collapse	Obstacle and object detection	Camera	O
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O
	Obstacle and object localization	Camera	O
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O
The ceiling inside the building collapsed due to the collapse of the building	Obstacle and object detection	Camera	O
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O
	Obstacle and object localization	Camera	O
		Camera + 2D LiDAR	X
		Camera + 3D LiDAR	O

5 Future Work

This study was intended to show the exact location of the disaster situation on the building drawing by detecting and determining the location of obstacles in space, the possibility of entry, and rescue requests. However, there are old buildings for which no drawings are available, and there are cases where existing drawings have become obsolete as the topography of the buildings has been changed due to disasters. Therefore, it is intended to use a 3D LiDAR-based SLAM algorithm in the future to detect and map the topography or obstacles of a disaster site to create a 3D map of the disaster site without existing drawings. There are various studies on the way to detect and map the surrounding topography or obstacles using the 3D LiDAR-based SLAM algorithm based on the information obtained by the camera and 3D LiDAR. Research [51] proposed a LiDAR-based SLAM loop closing methodology for measuring distances using 3D LiDAR sensors. After measuring the distance between objects based on the 3D LiDAR range image method, the overlap amount was obtained using the degree of overlap or function loss between the two values. Overlapnet using loop closing was proposed afterward in which the higher the value, the higher the coincidence rate. Research [52] used a surfel-based registration methodology to register the measured values scanned using the 3D LiDAR on the map based on the hierarchical refinement back-end [53]. The research combined the measurements into a multi-resolution map for localization. The measurements were refined using a continuous-time trajectory representation [54] for graph optimization. Research [55] used directed geometry point (DGP) and sparse frames to increase the localization accuracy. The research performed data optimization by iteratively learning with the scan-to-match module after extracting geometric points using the DGP extraction module. Research [56] proposed efficient 3D mapping with segment group matching. The research performed real-time loop detection using geometric segments extracted from LiDAR points. Robustification techniques were used to enhance the precision. Based on this, the segment-group matching (SGM) and SGM with pose-based method (SGM-PM) algorithms were proposed. Research [57] aimed to improve the 3D LiDAR large-area mapping accuracy and efficiency. Real-time distributed cooperative SLAM (RDC-SLAM), which transmits distributed data and fits the narrow bandwidth and limited range environment, is proposed. The research has the advantage of using the descriptor-based registration method to increase the efficiency in calculation time and data transmission. The distributed graph optimization module also connects data

between different robots and consistently monitors their global states. The above-mentioned 3D LiDAR-based SLAM will be designed in the future to create a 3D disaster site map that includes the disaster site situation.

6 Conclusion

This study proposes a method of using 3D LiDAR for firefighters and persons requiring rescue in case of a disaster. A low-cost, short-range 3D LiDAR and camera fusion algorithm is proposed, which is then implemented on an entry-level robot to identify disaster sites. Atypical obstacles, persons requiring rescue, and accessibility inside the building are determined based on the camera's image data, and the distance to the detected object is estimated based on the 3D LiDAR's point cloud data. The accuracy of detecting obstacles and persons requiring rescue is 0.93. The detection is achieved in various situations. This study is effective in 3D LiDAR and camera fusion, which can aid in quick and efficient search and rescue operations in disaster situations. Furthermore, using a camera and 3D LiDAR fusion to localize collapsed areas of the ceiling and floor surfaces can ensure the safety of firefighters who are dispatched to disaster situations.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. Brady, A. Denecke, E. Feenstra, R. H. Gongola and M. Hight *et al.,* "American society of civil engineers," (2021, Mar, 3), *Report card for america's infrastructure.* [Online]. Available: https://infrastructurereportcard.org/wp-content/uploads/2020/12/National_IRC_2021-report.pdf.

[2] E. Torres, in *New Orleans Sues Hard Rock Hotel Developers Over Fatal Collapse*, Manhattan, NY, USA: ABC News, 2020. [Online]. Available: https://abcnews.go.com/US/orleans-sues-hard-rock-hotel-developers-fatal-collapse/story?id=72621896.

[3] S. Bozorgmehr and L. S. Spark, in *Iran: 'More than 20 Firefighters Dead' in Tehran Building Collapse*, Atlanta GA, USA: CNN World, 2017. [Online]. Available: https://edition.cnn.com/2017/01/19/middleeast/iran-tehran-building-fire-collapse/index.html.

[4] Yujinrobot, (2020, Aug, 26), *YRL3-10*, [Online]. Available: https://yujinrobot.com/ko/.

[5] J. Ciberlin, R. Grbic, N. Teslic and M. Pilipovic, "Object detection and object tracking in front of the vehicle using front view camera," in *IEEE on 2019 Zooming Innovation in Consumer Technologies Conf. (ZINC)*, Novi Sad, Serbia, pp. 27–32, 2019.

[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR)*, HI, USA, pp. I-I, 2001.

[7] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[8] Z. Kalal, K. Mikolajczyk and J. Matas, "Forward-backward error: Automatic detection of tracking failures," in *ICPR Int. Conf. on Pattern Recognition*, Istanbul, Turkey, pp. 23–26, 2010.

[9] B. Strbac, M. Gostovic, Z. Lukac and D. Samardzija, "YOLO Multi-camera object detection and distance estimation," in *IEEE 2020 Zooming Innovation in Consumer Technologies Conf. (ZINC)*, Novi Sad, Serbia, pp. 26–30, 2020.

[10] J. Bonnard, K. Abdelouahab, M. Pelcat and F. Berry, "On building a cnn-based multi-view smart camera for real-time object detection," *Microprocessors and Microsystems*, vol. 77, pp. 103177, 2020.

[11] Y. Fan, Y. Li, Y. Shi and S. Wang, "Application of YOLOv5 neural network based on improved attention mechanism in recognition of thangka image defects," *KSII Transactions on Internet and Information Systems*, vol. 16, no. 1, pp. 245–265, 2022.

[12] Y. Hou, J. He and B. She, "Respiratory motion correction on pet images based on 3d convolutional neural network," *KSII Transactions on Internet and Information Systems*, vol. 16, no. 7, pp. 2191–2208, 2022.

[13] A. Jabbar, X. Li, M. M. Iqbal and A. J. Malik. "FD-stackGAN: Face de-occlusion using stacked generative adversarial networks," *KSII Transactions on Internet and Information Systems*, vol. 15, no. 7, pp. 2547–2567, 2021.

[14] M. Hirabayashi, K. Kurosawa, R. Yokota, D. Imoto, Y. Hawai *et al.,* "Flying object detection system using an omnidirectional camera," *Forensic Science International: Digital Investigation*, vol. 35, no. 6, pp. 301027, 2020.

[15] Teledyne flir, (2018), *Ladybug5+*, [Onlilne]. Available: https://www.flirkorea.com/products/ladybug5plus/.

[16] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed *et al.,* "SSD: Single shot multibox detector," in *Computer Vision–ECCV 2016*, vol. 9905, Cham, Switzerland: Springer, pp. 21–37, 2016.

[17] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[18] A. Sharjeel, S. A. Z. Naqvi and M. Ahsan, "Real time drone detection by moving camera using color and cnn algorithm," *Journal of the Chinese Institute of Engineers*, vol. 44, no. 2, pp. 128–137, 2021.

[19] M. Shakeri and H. Zhang, "COROLA: A sequential solution to moving object detection using low-rank approximation," *Computer Vision and Image Understanding*, vol. 146, pp. 27–39, 2016.

[20] D. Kim, C. Kim and J. Ahn, "Vision-based recognition algorithm for up-to-date indoor digital map generations at damaged buildings," *Computers, Materials & Continua*, vol. 72, no. 2, pp. 2765–2781, 2022.

[21] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," in *IEEE Int. Conf. on Computer Vision (ICCV)*, Venice, Italy, pp. 2961–2969, 2017.

[22] J. Wu, L. Zun and L. Yong, "Double-constraint inpainting model of a single-depth image," *Sensors*, vol. 20, no. 6, pp. 1797, 2020.

[23] L. Leyuan, H. Jian, R. Keyan, X. Zhonghua and H. Yibin, "A LiDAR-camera fusion 3d object detection algorithm," *Information*, vol. 13, no. 4, pp. 169, 2022.

[24] Y. Yenıaydin and K. W. Schmidt, "Sensor fusion of a camera and 2d lidar for lane detection," in *Signal Processing and Communications Applications Conf. (SIU)*, Sivas, Turkey, pp. 1–4, 2019.

[25] P. Krähenbühl and V. Koltun, "Efficient inference in fully connected CRFs with Gaussian edge potentials," in *Conf. on Neural Information Processing Systems*, Granada, Spain, pp. 109–117, 2011.

[26] S. Gu, Y. Zhang, J. Tang, J. Yang and H. Kong, "Road detection through CRF based lidar-camera fusion," in *IEEE 2019 Int. Conf. on Robotics and Automation (ICRA)*, Montreal, QC, Canada, pp. 3832–3838, 2019.

[27] X. Qi, W. Wang, Z. Liao, X. Zhang, D. Yang *et al.,* "Object semantic grid mapping with 2d lidar and RGB-D camera for domestic robot navigation," *Applied Sciences*, vol. 10, no. 17, pp. 5782, 2020.

[28] O. Tslil, A. Elbaz, T. Feiner and A. Carmi, "Representing and updating objects' identities in semantic SLAM," in *Conf. on Information Fusion (FUSION)*, Rustenburg, South Africa, pp. 1–7, 2020.

[29] D. Kim and J. Ahn, "Intelligent abnormal situation event detections for smart home users using lidar, vision, and audio sensors," *Journal of Internet Computing and Services*, vol. 22, no. 3, pp. 17–26, 2021.

[30] R. Barea, C. Pérez, L. M. Bergasa, E. L. Guillén, E. Romera *et al.,* "Vehicle detection and localization using 3d lidar point cloud and image semantic segmentation," in *Int. Conf. on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, pp. 3481–3486, 2018.

[31] X. Junyao, Y. Dejun, W. Zijian and F. Zezheng, "A simultaneous localization and mapping technology based on fusion of radar and camera," in *Proc. 3rd Int. Conf. on Mechanical Instrumentation and Automation (ICMIA)*, Sanya, China, pp. 012029, 2022.

[32] Q. Xu, Y. Zhou, W. Wang, C. R. Qi and D. Anguelov, "SPG: Unsupervised domain adaptation for 3d object detection via semantic point generation," in *IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, Montreal, QC, Canada, pp. 15446–15456, 2021.

[33] W. G. Chang, T. You, S. Seo, S. Kwak and B. Han, "Domain-specific batch normalization for unsupervised domain adaptation," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 7346–7354, 2019.

[34] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang *et al.,* "Pointpillars: Fast encoders for object detection from point clouds," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 12689–12697, 2019.

[35] S. Shi, C. Guo, L. Jiang, Z. Wang, J. Shi *et al.,* "PV-RCNN: Point-voxel feature set abstraction for 3D object detection," in *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 3481–3486, 2020.

[36] Z. Yin and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3d object detection," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, pp. 4490–4499, 2018.

[37] H. Seong, J. Hyun and E. Kim, "FOSnet: An end-to-end trainable deep neural network for scene recognition," *IEEE Access*, vol. 8, pp. 82066–82077, 2020.

[38] P. Sun, W. Wang, Y. Chai, G. Elsayed, A. Bewley *et al.,* "RSN: Range sparse net for efficient, accurate lidar 3d object detection," in *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, pp. 5725–5734, 2021.

[39] M. Sualeh and G. W. Kim, "Visual-lidar based 3d object detection and tracking for embedded systems," *IEEE Access*, vol. 8, pp. 156285–156298, 2020.

[40] A. Geiger, P. Lenz, C. Stiller and R. Urtasun, "Vision meets robotics: The KITTI dataset," *International Journal of Robotics Research (IJRR)*, vol. 32, no. 11, pp. 1231–1237, 2013.

[41] R. G. Gioi, J. Jakubowicz, J. Morel and G. Randall, "LSD: A line segment detector," *Image Processing on Line (IPOL)*, vol. 2, pp. 35–55, 2012.

[42] R. M. Haralick, S. R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 4, pp. 532–550, 1987.

[43] A. Bochkovskiy, C. Wang and H. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.

[44] Common Objects in Context, (2017), *COCO Dataset 2017 Test Images*, [Online]. Available: http://images.cocodataset.org/zips/test2017.zip.

[45] D. Kim and J. Ahn, "Intelligent SLAM algorithm fusing low-cost sensors at risk of building collapses," *Computers, Materials & Continua*, In press, April 2022.

[46] S. Kohlbrecher, O. Stryk, J. Meyer and U. Klingauf, "A flexible and scalable SLAM system with full 3d motion estimation," in *IEEE Int. Symp. on Safety, Security, and Rescue Robotics (SSRR)*, Kyoto, Japan, pp. 155–160, 2011.

[47] M. Ester, H. P. Kriegel, J. Sander and X. Xu, "A Density-based algorithm for discovering clusters in large spatial databases with noise," *AAAI Press*, vol. 96, no. 34, pp. 226–231, 1996.

[48] J. Ahn and R. Han, "Rescueme: An indoor mobile augmented-reality evacuation system by personalized pedometry," in *IEEE Asia-Pacific Services Computing Conf.*, Jeju, Korea, pp. 70–77, 2011.

[49] Samsung, (2019, Apr, 5), *Galaxy S10*, [Onlilne]. Available: https://www.samsung.com/sec/smartphones/all-smartphones/.

[50] OMOROBOT, (2021, Oct, 19), *OMO R1*, [Online]. Available: https://omorobot.com/docs/omo-r1/.

[51] C. Xieyuanli, L. Thomas, M. Andres, R. Timo, V. Olga *et al.,* "Overlapnet: Loop closing for lidar-based SLAM," in *Robotics: Science and Systems, Virtual Conference*, Corvalis, OR, USA, 2020.

[52] D. Droeschel and S. Behnke, "Efficient continuous-time SLAM for 3d lidar-based online mapping," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Brisbane, Australia, pp. 5000–5007, 2018.

[53] W. Guangming, Z. Jiquan, Z. Shijie, W. Wenhua, L. Zhe *et al.,* "3D hierarchical refinement and augmentation for unsupervised learning of depth and pose from monocular video," arXiv preprint arXiv:2112.03045, 2022.

[54] E. Mueggler, G. Gallego and D. Scaramuzza, "Continuous-time trajectory estimation for event-based vision sensors," in *Robotics on Science and Systems*, Rome, Italy, 2015.

[55] S. Liang, Z. Cao, C. Wang and J. Yu, "A novel 3d lidar SLAM based on directed geometry point and sparse frame," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 374–381, 2021.

[56] M. Tomono, "Loop detection for 3d lidar SLAM using segment-group matching," *Advanced Robotics*, vol. 34, no. 23, pp. 1530–1544, 2020.

[57] X. Yuting, Z. Yachen, C. Long, C. Hui, T. Wei *et al.,* "RDC-SLAM: A real-time distributed cooperative SLAM system based on 3d lidar," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 1–10, 2021.