2025 European Conference on Computing in Construction
CIB W78 Conference on IT in Construction
Porto, Portugal
July 14-17, 2025

A FRAMEWORK OF PANORAMIC IMAGE-BASED 3D SEMANTIC RECONSTRUCTION FOR BIM ENRICHMENT OF FIREFIGHTING ASSETS

Yutong Qiao [1], Ya Wen [2,*], Chi Chiu Lam [1], Ioannis Brilakis [2], Sanghoon Lee [3], and Mun On Wong [1,*]
[1] University of Macau, Macau SAR, China
[2] University of Cambridge, Cambridge, United Kingdom
[3] University of Seoul, Seoul, South Korea
[*] Corresponding authors: yw710@cam.ac.uk, mowong@um.edu.mo

Abstract

Constructing digital models of firefighting assets is essential for informed decision-making in emergency response and management. However, current practices struggle to efficiently recognize and update these assets in building information models (BIM). This study proposes a framework integrating photogrammetric reconstruction and instance segmentation to enrich BIM. The framework involves creating an expanded firefighting asset dataset and leveraging panoramic images with supervised learning for scene reconstruction and asset segmentation. Real-world evaluation illustrates it performs satisfactorily in both asset recognition and positioning. The framework offers a practical solution for modeling digital twins to support various fire emergency applications.

Introduction and related works

Information of firefighting assets is crucial to support fire risk assessments, code compliance checking (Fitkau & Hartmann 2024), fire dynamic simulation, and real-time emergency response (Ma & Wu 2020). For instance, the quantity, types, and spatial distribution of different firefighting assets are utilized as indicators to objectively measure the risk levels of modern and historical buildings (Qiao, Lam, & Wong, 2024; Qiao et al., 2024). In addition, the locations and statuses of fire service installations are desired by building occupants and first responders to support their escape and fire suppression (Wong & Lee 2022; Wong & Lee 2023).

Despite their importance, the information of firefighting assets is conventionally stored in 2D drawings, which are hardly interpreted in a digital manner to support data-driven analysis. With the emergence of building information modeling (BIM) and digital twin (DT) technologies, the fire safety design and modeling of newly constructed buildings are now able to be managed digitally in BIM models. However, the majority of existing buildings that did not have BIM models or had only simple architectural BIM models may still lack fundamental information of firefighting assets. In other words, how to create and update the firefighting assets in time is still a challenge in facilitating digital fire information management.

In conventional practice, the enrichment of firefighting assets into BIM models relies heavily on manual involvement. Considerable efforts would be made to manually create the BIM models with rich firefighting asset-related information of existing buildings. To tackle this issue, previous studies have developed different solutions that acquire firefighting asset information from existing drawings or built environments. For example, Schönfelder et al. (2024) developed an automated workflow to interpret fire safety equipment information from 2D escape plans, which can successfully extract the symbol locations of firefighting assets and convert them into physical locations in BIM models. Compared to drawing-based enrichment, capturing firefighting assets from reality has advantages in collecting as-is information that reflects the up-to-date status. Despite its potential, existing studies paid little attention to scanning comprehensive firefighting assets and reconstructing their semantic and geometric information into BIM models. Several studies have attempted to reconstruct various assets in indoor scenes, but their scopes are not specialized in fire safety so that the critical firefighting assets have not been comprehensively investigated (Di & Gong 2024; Pan et al. 2022). More specifically, there are several critical gaps in the reality capture-based firefighting asset reconstruction. First, datasets for the recognition of firefighting assets are highly limited. Currently, the most commonly used dataset, FireNet (Boehm, Panella & Melatti 2019), contains 1,452 images and covers eight fundamental classes (Bayer & Aziz 2022). However, other essential firefighting assets, such as hose reels, sprinklers, and emergency lights, are not yet included, which constrains the performance of automated recognition methods for BIM enrichment. Second, there has been scarce exploration of the sole use of panoramic images for reconstructing firefighting assets in BIM models. While existing studies often rely on more accurate but expensive devices such as LiDAR and laser scanners for indoor reality capture (Adán et al. 2018), the capabilities of photogrammetric reconstruction based on images remain underexplored. This gap leaves the potential of image-based approaches for BIM enrichment of firefighting assets unclear.

To address this issue, the study aims to propose a framework of panoramic image-based 3D semantic reconstruction for BIM enrichment of firefighting assets. Instead of normal cameras, panoramic cameras are selected in this study since they have 360-degree coverage of surrounding environments, providing richer visual features for scene reconstruction and forming a more convenient way to capture firefighting assets from all directions. Correspondingly, the objectives of this study are two-fold: (1) to create an expanded firefighting asset recognition dataset that broadens the range of recognition classes and increases the number of images to enhance recognition performance; (2) to develop a generic approach that leverages the advantages of panoramic images for photogrammetric reconstruction while retaining high compatibility with various image inputs for 3D semantic segmentation of firefighting assets. Through the implementation of a real-world case study, the framework is validated as feasible and effective, resulting in good performance in both recognition accuracy (precision of 100% and recall of 74%) and localization accuracy (0.434 meters). It is believed that this study provides a first-of-its-kind solution to facilitate the digital management of fire safety equipment.

Methodology

The proposed framework incorporates three critical stages, i.e., firefighting asset recognition, 3D point cloud reconstruction and segmentation, as well as BIM model enrichment, as shown in Figure 1.

Fire equipment asset recognition

A comprehensive and well-structured training dataset is essential for the successful implementation of image-based fire equipment recognition in the proposed method. Existing image-based object recognition datasets typically contain labels for common objects, while there is a significant lack of datasets specifically focused on firefighting asset recognition. FireNet is one of the most comprehensive databases for indoor fire equipment detection and segmentation (Boehm, Panella & Melatti 2019). It comprises over 1,400 images, encompassing various categories of fire equipment, including fire extinguishers, call points, projective blankets, escape route signs, fire equipment signs, visual alarms, fire alarm sounders, and smoke detectors. Despite these strengths, FireNet has three main limitations that constrain its practical utilization. First, it has a class imbalance issue, which adversely affects the capabilities of the supervised learning models for recognizing those underrepresented classes. Second, the dataset only covers limited types of basic firefighting assets. Thirdly, the majority of images in FireNet were captured in buildings located in the United Kingdom, covering firefighting assets with region-specific features like English-language labels and UK-style symbols. This may affect the generalizability of the dataset to regions with different languages and symbol designs.

To improve the fire equipment recognition performance, this research expands the existing database by collecting additional images from publicly accessible sources. A custom search API was developed to retrieve images systematically from search engines (e.g., Google). The search process was designed with specific criteria to ensure relevance and usability: (1) the use of targeted keywords is related to fire equipment in building-specific scenarios and (2) the permission of images is licensed under Creative Commons Zero (Creative Commons 2019), ensuring these images are freely available for all purposes. Next, a semi-automated image filtering process is also implemented after collecting the data. The filtering process involves an automated removal of image duplication using Duplicate Cleaner (DigitalVolcano Software 2025) and a manual screening to discard low solution images and irrelevant image to enhance the quality of the dataset. As a result, 3,043 images are downloaded from the Google search engine, and 468 images left after the filtering. Together with the original FireNet dataset, the expanded <u>fire</u>fighting <u>a</u>sset <u>r</u>ecognition datase<u>t</u>, which is named as Fire-ART in this research, contains 1,920 images in total.

In addition to adding new images from a broader geographical source, this study also increases the recognition classes to cover 15 different firefighting

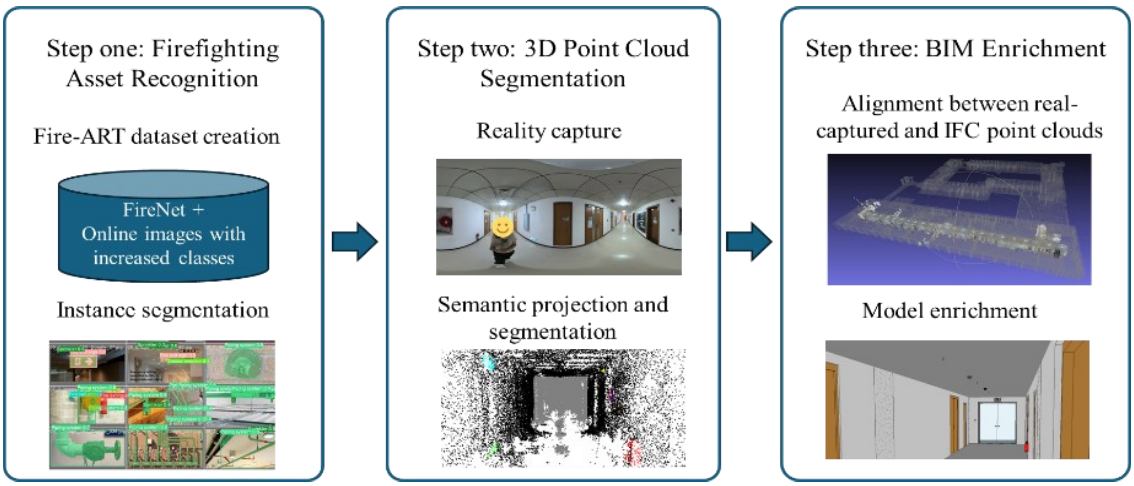

Figure 1: The overview of 3D semantic reconstruction framework for firefighting asset enrichment

assets, i.e., fire extinguishers, fire exit signs, fire door signs, fire alarms, emergency lights, smoke detectors, fire hose reels, piping systems, sprinklers, fire call points, emergency door releases, fire blankets, fire equipment signs, firefighting lift switches, and hidden fire equipment. All the images in Fire-ART are consistently labeled by the authors using the X-AnyLabeling labeling tool (Wang 2023), in which the Segment Anything model (Kirillov et al. 2023) is utilized to help delineate boundaries of objects automatically. Figure 2 shows the instance numbers per class in the Fire-ART and the original FireNet datasets for direct comparison.

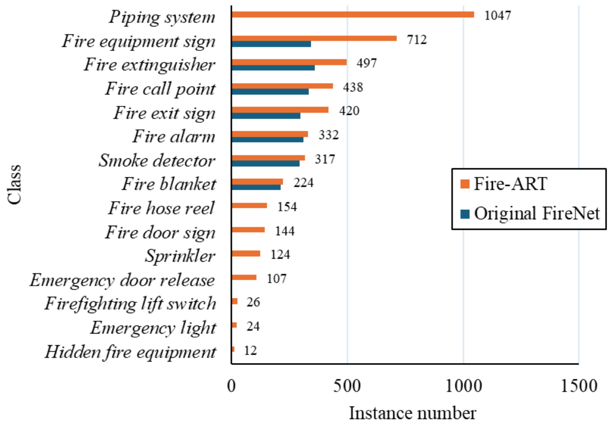

Figure 2: Instance distribution per class: Fire-ART vs. original FireNet dataset

On the basis of the Fire-ART dataset, this study further leverages supervised learning models for firefighting assets' recognition within images. Specifically, the off-the-shelf instance segmentation model, i.e., YOLOv8 (Jocher, Chaurasia & Qiu 2023), is selected since it enables rapid inference while maintaining good segmentation accuracy. To conduct training, the Fire-ART dataset is split into a train, validation, and test subsets in line with a ratio of 0.70:0.15:0.15. As a result, the training subset has 1344 images, while both validation and test sets contain 288 images. The training is performed using a desktop computer, with AMD Ryzen 5 5600X CPU, NVIDIA GeForce RTX 4070 Ti GPU and 12GB VRAM GPU memory. Consequently, the best-performing model with the least validation loss is selected for real-world deployment to predict instances of firefighting assets within images. The prediction results on the test subset are summarized in Table 1.

Table 1: Performance metrics of the YOLOv8 model on the Fire-ART test subset

Class	Instances	Precision	Recall	mAP^{50}	$mAP^{50\text{-}95}$
Fire extinguisher	67	0.912	0.927	0.948	0.542
Fire exit sign	59	0.791	0.814	0.801	0.15
Fire door sign	22	0.937	0.68	0.785	0.346
Fire alarm	62	0.901	0.883	0.937	0.648
Emergency light	2	1	0	0	0
Smoke detector	35	0.788	0.8	0.902	0.531
Fire hose reel	19	0.908	0.684	0.814	0.639
Piping system	141	0.603	0.333	0.333	0.163
Sprinkler	16	0.745	0.312	0.32	0.194
Fire call point	69	0.941	0.921	0.94	0.653
Emergency door release	13	0.741	0.615	0.728	0.485
Fire blanket	39	0.98	0.974	0.985	0.815
Fire equipment sign	109	0.891	0.881	0.871	0.651
Firefighting lift switch	8	0.442	0.25	0.158	0.0721
Hidden fire equipment	2	1	0	0.0421	0.0421
All	738	0.839	0.605	0.638	0.419

3D point cloud reconstruction and segmentation

To collect firefighting asset data, this study adopts a 360-degree panoramic camera to record videos and capture detailed representations of indoor environments. The panoramic camera is selected since it can provide horizontally elongated fields of view, which enables comprehensive capture from all angles and accelerates the data acquisition process with abundant features to benefit the photogrammetric reconstruction. The recorded panoramic videos are subsequently processed using FFmpeg (FFmpeg 2019) to extract individual frames, serving as sequential inputs for point cloud reconstruction. Next, an off-the-shelf photogrammetry software named Metashape (Agisoft 2019) is utilized to perform structure from motions and multi-view stereo calculations. As a result, a dense point cloud of the scene as well as the relative camera trajectory are obtained.

After creating the 3D point cloud structure, the next step is to identify and segment the point clusters representing different firefighting assets. This step is critical for determining the spatial locations and quantities of firefighting assets within the 3D scenes, which serve as fundamental data for enriching BIM models. To do so, this study conducts deep learning-based instance segmentation on 2D images and further projects the segmentation masks onto the 3D point cloud structure using spherical camera projection. Figure 3 illustrates the conceptual workflow of the semantic projection and segmentation.

In regard to image-based instance segmentation, this study leverages the trained YOLOv8 model to recognize and delineate firefighting assets within images. Specifically, panoramic images are typically displayed using the equirectangular projection. This projection represents spherical views onto a rectangular image plane, which may introduce distortions and make the panoramic images significantly different from the standard perspective images. The appearance of the firefighting assets in the equirectangular images would be stretched or compressed disproportionally and thus affect the visual patterns for recognition. On the other hand, since the Fire-ART dataset consists exclusively of standard perspective images for model training, the heterogeneity between the training data and the panoramic images may lead to suboptimal model performance when applied directly. To mitigate this issue, a pre-processing conversion is carried out to split panoramic images into cube maps.

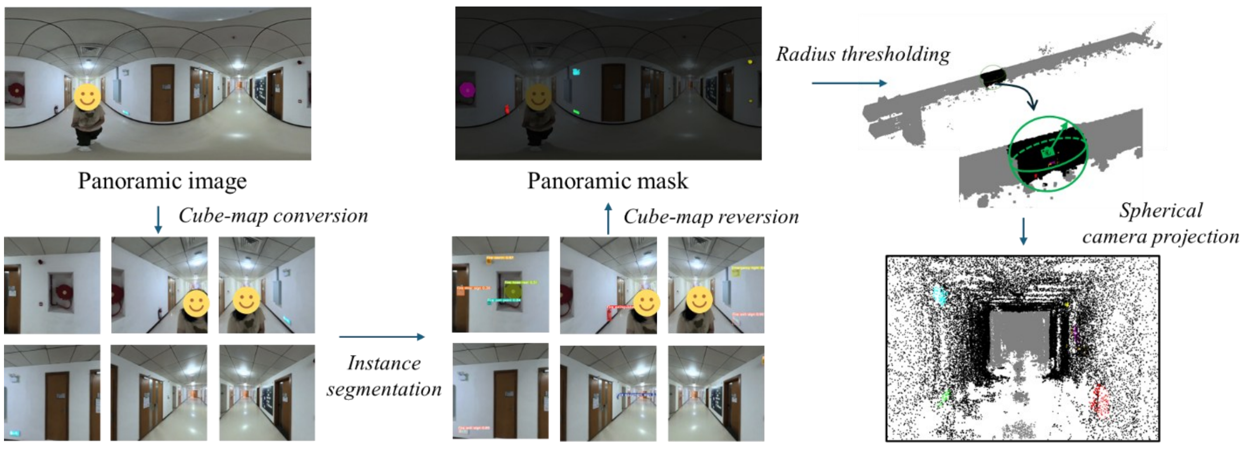

Figure 3: Workflow of the 3D point cloud segmentation of firefighting assets

Particularly, a single panoramic image is decomposed into six horizontal rectilinear face images, where each face image contains a 30-degree of overlap between its adjacent images, as shown in Figure 4. This overlap is intentionally designed to ensure that firefighting assets remain intact within at least one face image, thereby preventing the appearance of assets from being split fragmentally across multiple images, which could break their recognizable features and reduce segmentation accuracy.

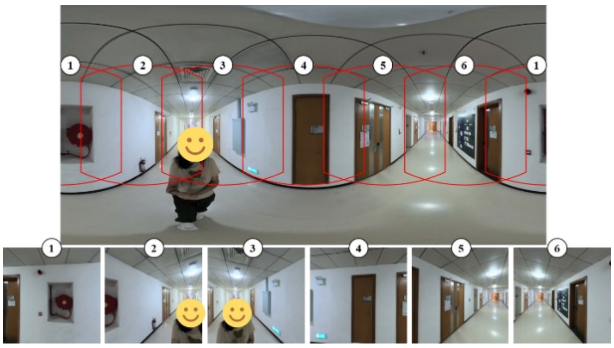

Figure 4: Cube-map conversion

Afterward, the face images are input into the trained models to perform instance segmentation, resulting in prediction masks of firefighting assets. For each set of six face images, the corresponding masks are then combined to generate a panorama mask for semantic projection.

Given that there are overlaps between two adjacent face images, inconsistent prediction results may occur in these overlapping regions due to the variations in the capturing perspectives. To handle this situation, this study prioritizes the classes of firefighting assets over the class of background. Additionally, for the conflicts between two recognized firefighting asset classes, the confidence levels of the predictions are leveraged to determine priority. This ensures that the panoramic mask incorporates more firefighting asset instance annotations with higher confidence levels.

Building upon the panoramic masks, semantic projections are further conducted to associate the image masks with the photogrammetric point clouds. The mechanism of the semantic projection primarily utilizes the spherical camera model to project the 3D point cloud onto the spherical surfaces constituted by the singular panoramic masks. Specifically, each point in the cloud is mapped to a pixel in the panorama mask, from which the semantic class (either background or different types of firefighting assets) is acquired and assigned back to the corresponding point. Given that the conventional spherical camera model would project all the points onto the spherical surface along the directions from the points to the camera position regardless of their distance, wall penetration issues may occur in indoor environments. This issue may incur incorrect classification, where some of the points are supposed to be blocked by other occlusions but would still be assigned to certain semantic classes indiscriminately. To address this issue, a radius threshold is established to limit the projection process to points $\boldsymbol{p_i}$ within a specified distance r from the camera position $\boldsymbol{c_j}$ while excluding points farther away. By setting this threshold, the projection effectively focuses on the point cloud objects that are close to the camera positions, which are typically associated with higher segmentation accuracy. Equations (1) – (3) illustrate the mapping between the cloud points (x_i, y_i, z_i) and the corresponding image pixels (u_i, v_i) using the modified spherical camera model.

$$\boldsymbol{P_r} = \{\boldsymbol{p_i} \mid \|\boldsymbol{p_i} - \boldsymbol{c_j}\|_2 \leq r\} \tag{1}$$

$$\begin{cases} u_i = 0.5w + f \times \tan^{-1}(x_i/z_i) \\ v_i = 0.5h + f \times \tan^{-1}\left(y_i/\sqrt{{x_i}^2 + {z_i}^2}\right), \end{cases} \tag{2}$$

$$\text{where } \boldsymbol{p_i} = (x_i, y_i, z_i) \in \boldsymbol{P_r}$$

$$f = w/(2\pi) \tag{3}$$

where $\boldsymbol{P_r}$ is the point set containing the points within the distance threshold r, w and h are the width and height of the panoramic image, and f is the focal length of the image.

The entire semantic projection is conducted iteratively for each camera position using its associated panoramic mask. As a result, the 3D segmentation is obtained by collectively considering all the projections across the scanned areas. To aggregate the results, a weighted majority voting approach is implemented to determine the final classes of points by counting the most frequently occurring class across the projections. In this way, point clusters of firefighting assets are accurately identified and segmented within the 3D point cloud structure.

BIM enrichment

Since the photogrammetric point cloud and BIM model have two separate coordinate systems, accurate registration between the two 3D structures is required. In this study, a semi-automated alignment is conducted. The alignment progress starts with converting the BIM model into a 3D point cloud. Specifically, this study employs IFC 4×3 schema (buildingSMART 2024) to export the BIM model into an IFC data model. Next, the geometric information from the IFC model is extracted to acquire the polygonal surfaces of different architectural elements, including walls, floors, ceilings, and doors. Random point sampling is applied to generate numerous points within these surfaces proportional to their surface areas, resulting in a point cloud of the BIM model for alignment purposes. To align the photogrammetric point cloud with the BIM-derived point cloud, CloudCompare (Girardeau-Montaut 2025) is leveraged for pairwise point picking. The point-picking process involves the manual selection of corresponding point pairs, which enables the photogrammetric point cloud being transformed into a consistent coordinate system of the BIM-derived point cloud. On this basis, point cloud clustering is performed using the DBSCAN algorithm (Ester et al. 1996) to group the firefighting asset points into distinct clusters, with each cluster representing an individual firefighting asset. Subsequently, the locations and types of each firefighting asset object are obtained and utilized to enrich the BIM model. The enrichment process at present relies on manually defining and inserting the relevant families of firefighting assets into the BIM model based on the calculated locations and types. Specifically, predefined families from the Glodon digital component dock Revit family library (Glodon 2025) are leveraged to streamline the process. The BIM objects with the most similar appearance are manually selected from the family library and inserted into the model. Currently, the approach still requires modelers to infer the spatial constraints from the segmented point cloud and replicate the relationships in the model. Consequently, the BIM model is enriched with firefighting asset information to support general building inventory management and emergency risk assessment.

Case study

This study selected a corridor area of an institutional building at the University of Macau as the experimental site. The corridor areas have standardized fire safety equipment configuration, including ten types of firefighting assets such as fire extinguishers, fire exit signs, fire alarms, emergency lights, smoke detectors, fire hose reels, piping systems, sprinklers, fire call points, and emergency door releases. The BIM model of the institutional building with a level of details at 200 was available in advance, which contains fundamental building components like walls, floors, ceilings, doors, and windows while lacking the firefighting equipment.

Data collection was conducted using an Insta360 X4 panoramic camera, mounted at a height of 1.75 meters consistently to maintain stable capture perspectives. The acquisition path followed a predefined movement route that fully covered the corridor area, as shown in Figure 5. During the data preprocessing phase, the research team extracted 830 panoramic images from the raw video at a sampling rate of 3 frames per second and converted them into 14,940 rectilinear face images using the OpenCV library. The panoramic images were further imported into Metashape for point cloud reconstruction, and the face images were processed using the trained YOLOv8 model for firefighting asset segmentation. Subsequently, the semantic masks are merged to conduct the spherical camera projection to assign the firefighting asset classes to the cloud points. As a result, the segmented point cloud was acquired, in which the locations and numbers of various firefighting assets are calculated via point cloud clustering. Finally, these firefighting asset data were leveraged to enrich the BIM models by aligning the point cloud to the IFC data model as well as inserting firefighting asset objects in line with the calculation results.

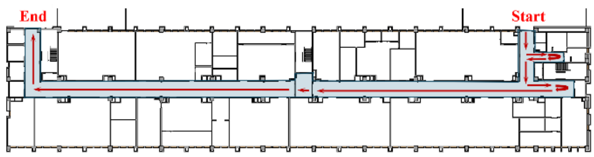

Figure 5: Experimental site and data collection route in the institutional building corridor

For comparison, this study also generated the ground truth BIM model by manually measuring and recording the firefighting assets. Consequently, a total of 134 instances of fire safety equipment were created.

To evaluate the recognition performance of the proposed framework, this study compares the recognized instances with the ground-truth instances. Specifically, precision and recall metrics were utilized to quantify the accuracy, as summarized in Table 2. Figure 6 illustrates the spatial distribution of recognized instances and their comparison with ground-truth instances on the BIM floor plans. Overall, an average precision of 100% and an average recall of 74% were achieved, indicating a moderately high level of recognition performance. Moreover, the recognition accuracy varied notably across different types of firefighting assets. Fire extinguishers, piping systems, and fire alarm points obtained 100% in both precision and recall rates, suggesting that all these relevant components were recognized and pinpointed successfully. In contrast, emergency lights, smoke detectors, sprinklers, and emergency door releases had relatively lower recall rates, at 54%, 44%, 56%, and 29%, respectively. Further investigation revealed two primary factors contributing to these lower recall rates. First, certain firefighting assets, such as emergency lights and emergency door releases, had limited training data, which constrained the model's recognition capabilities. Second, this present study only considered horizontal rectilinear face images, in which ceiling-mounted assets, such as sprinklers and smoke detectors would appear relatively small within these face images, making them more difficult to be detected accurately.

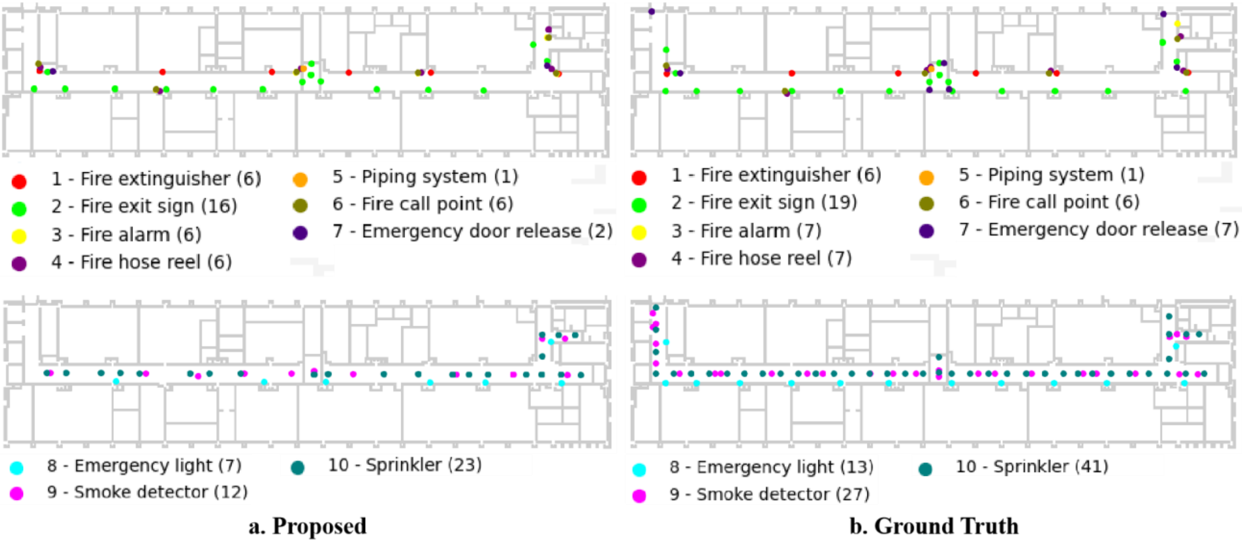

Figure 6: Spatial distribution of recognized and ground-truth firefighting assets on the BIM floor plans

In addition to the recognition accuracy, this study measured the distance errors between the ground-truth instances and the recognized instances. The results are summarized in Table 2, which indicates that all firefighting assets had a similar positioning performance, reaching a distance error of 0.434 meters on average. Empirically, these errors could stem from the photogrammetric reconstruction as well as the semi-automated alignment between the point cloud and the IFC BIM model. While there is space for further improvement, the result should be sufficient to provide room-level localization accuracy for many real-world applications, such as asset inventory management, risk assessment, and information provision for emergency response.

Table 2: Recognition and localization accuracy results

Firefighting asset	GT	TP	FP	FN	Precision	Recall	Distance error (meter)
Fire extinguisher	6	6	0	0	100%	100%	0.357
Fire exit sign	19	16	0	3	100%	84%	0.438
Fire alarm	7	6	0	1	100%	86%	0.419
Emergency light	13	7	0	6	100%	54%	0.483
Smoke detector	27	12	0	15	100%	44%	0.434
Fire hose reel	7	6	0	1	100%	86%	0.678
Piping system	1	1	0	0	100%	100%	0.378
Sprinkler	41	23	0	18	100%	56%	0.342
Fire call point	6	6	0	0	100%	100%	0.387
Emergency door release	7	2	0	5	100%	29%	0.425
Average					100%	74%	0.434

Note: GT, TP, FP, and FN respectively represent the numbers of ground-truth, true-positive, true-negative, and false-negative instances.

Discussion

This study develops the Fire-ART dataset and a cost-effective panoramic image-based 3D reconstruction approach to empower the enrichment of BIM models for firefighting assets. Specifically, the implications and limitations are elaborated as follows.

First, this research expands the dataset beyond FireNet through a systematic collection of publicly available images. The expanded Fire-ART dataset not only increases the number of images, but also covers new categories of fundamental firefighting assets, such as fire hose reels, sprinklers, and emergency lights. Consequently, the Fire-ART dataset contains 4,578 instances in total and retains 2.38 instances per image on average, providing richer content and more data for supervised learning and benchmarking compared to the existing FireNet dataset (2,154 instances and 1.48 instances per image). Besides, the addition of publicly available images includes various indoor environments across different countries, which helps reduce regional biases and enhances the generalizability of our dataset.

Second, this study introduces a panoramic image-based generic approach for indoor scene reconstruction and component segmentation to empower BIM enrichment of firefighting assets. The proposed approach exerts the strengths of both panoramic and perspective images: (1) it utilizes panoramic images for 3D scene reconstruction, where the 360-degree captures provide much more details for visual feature matching and thus are significantly instrumental in reconstructing 3D point cloud structures completely, and (2) it flexibly takes perspective images as inputs for instance segmentation, ensuring good compatibility with the widely used datasets and models in existing research. Particularly, the approach involves a cube-map conversion and reversion to split panoramic images into horizontal face images for firefighting asset predictions, then further merges the results into panoramic semantic masks for spherical camera projections. Through this approach, commonly available images can be utilized as training datasets, and the challenges of the scarcity of 360-degree training datasets can be effectively addressed. Together with the Fire-ART dataset, the proposed approach achieved relatively good performance in both firefighting asset recognition and localization, with a precision rate of 100%, a recall rate of 74%, and a deviation error of 0.434 errors.

The limitations of current research are also identified and should be addressed in future studies. First, the accuracy of recognition and segmentation varies across asset categories and needs to be improved to ensure the applicability and reliability in practice. For instance, emergency lights and sprinklers exhibit lower recall rates due to their less distinctive features, insufficient training materials, and ceiling-mounted positions. To overcome this, the Fire-ART dataset should be further expanded, and the preprocessing techniques should be refined to tailor the characteristics of ceiling-mounted objects, such as including additional tilt-angle captures in the cube-map conversion. Moreover, resolutions could potentially affect the recognition and reconstruction performance, where advanced super-resolution techniques may be instrumental in enhancing the results for objects with small appearances.

Second, the BIM alignment and enrichment process remains partially manual, particularly for registering the photogrammetric point cloud with the BIM-derived point cloud. Automating the alignment of photogrammetric point clouds with BIM models is essential for improving workflow efficiency. Methods such as feature-based matching or AI-driven identification and alignment of architectural features will be developed in future research. Furthermore, integrating predefined firefighting asset families into BIM software, e.g., open-source IFC modeling tools like BonsaiBIM (Blender community 2025), to automate the generation of enriched models will be also considered. Several open research areas, such as (1) automated selection of appropriate firefighting asset families based on their semantic/geometric information and (2) spatial relationship reasoning among the recognized assets and the established building components, deserve further systematic investigation.

Last, the approach should be validated with additional real-world case studies. Diverse space types in various buildings across different countries are desired to quantify the performance of the firefighting asset recognition model and further evaluate the generalizability of the proposed approach.

Conclusion

This study develops an innovative framework for panoramic image-based fire asset recognition and 3D reconstruction to support BIM model enrichment for fire safety management. First, the study develops the Fire-ART dataset by increasing the image numbers and covering more types of firefighting assets. 15 critical assets with more than 4,000 labeled instances are contained in the dataset, making it a valuable resource for supervised learning and benchmarking of fire safety equipment image recognition. Second, a generic and cost-effective approach is proposed to leverage panoramic images for comprehensive 3D scene reconstruction and point cloud segmentation. Specifically, cube-map conversion and reversion techniques are developed to flexibly perform firefighting asset predictions as well as spherical camera projections, seamlessly mapping semantic classes from 2D rectilinear face images onto 3D point cloud structures. Through a real-world case study, the proposed framework is validated as effective in recognizing and pinpointing firefighting assets in 3D reality captures to empower BIM model enrichment. Consequently, the framework achieved a recognition precision of 100%, a recall rate of 74%, and a localization accuracy of 0.434 meters. The framework forms a technical solution to enhance the automation of digital management of fire safety assets, thereby facilitating the development of data-driven asset inventory management, fire risk assessment, and on-site emergency response. Future work will focus on enhancing the diversity of the dataset, refining the cube-map conversion to improve asset segmentation accuracy, developing fully automated BIM model enrichment of firefighting assets, as well as validating the refined approach with more diverse real-world scenarios. To facilitate future research, the Fire-ART dataset is available to download from: https://github.com/UMRATE/Fire-ART.

Acknowledgments

The authors acknowledge funding support from The Science and Technology Development Fund, Macao S.A.R (FDCT project no. 0034/2024/RIB1), University of Macau (Conference Grant – FST, CG2025-FST), and funding from the Horizon Europe UKRI Underwrite Innovate, under the grants G115919 – BuildSpace, University of Cambridge.

References

Adán, A, Quintana, B, Prieto, SA & Bosché, F 2018, 'Scan-to-BIM for "secondary" building components', Advanced Engineering Informatics, vol. 37, pp. 119–138.

Agisoft 2019, Agisoft Metashape, Agisoft.com. Available at: https://www.agisoft.com/.

Bayer, H & Aziz, A 2022, Object Detection of Fire Safety Equipment in Images and Videos Using Yolov5 Neural Network, in Proceedings of 33. Forum Bauinformatik, Forum Bauinformatik, Munich, Germany, pp. 62–69.

Blender community 2025, BonsaiBIM. Available at: https://bonsaibim.org/index.html.

Boehm, J, Panella, F & Melatti, V 2019, 'FireNet', figshare, University College London. Available at: https://doi.org/10.5522/04/9137798.v1.

buildingSMART 2024, IFC 4.3.2 Documentation. Available at: https://ifc43-docs.standards.buildingsmart.org.

Creative Commons 2019, Creative Commons — CC0 1.0 Universal. Creativecommons.org. Available at: https://creativecommons.org/publicdomain/zero/1.0/.

Di, C & Gong, J 2024, ‘An AI-based approach to create spatial inventory of safety-related architectural features for school buildings’, Developments in the Built Environment, vol. 17, pp. 100376.

DigitalVolcano Software 2025, Duplicate Cleaner - Remove duplicate files. Duplicatecleaner.com. Available at: https://www.duplicatecleaner.com.

Ester, M, Kriegel, H-P, Sander, J & Xu, X 1996, A density-based algorithm for discovering clusters in large spatial databases with noise, in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231.

FFmpeg 2019, FFmpeg, Available at: http://ffmpeg.org/.

Fitkau, I & Hartmann, T 2024, ‘An ontology-based approach of automatic compliance checking for structural fire safety requirements’, Advanced engineering informatics, vol. 59, pp. 102314–102314.

Girardeau-Montaut, D. 2025, CloudCompare - Open Source project. Cloudcompare.org. Available at: http://www.cloudcompare.org.

Glodon 2025, Glodon Digital Component Dock Revit Family Library, Available at: https://www.goujianwu.com/.

Jocher, G., Chaurasia, A. and Qiu, J. 2023, YOLOv8 by Ultralytics. GitHub. Available at: https://github.com/ultralytics/ultralytics.

Kirillov, A, Mintun, E, Ravi, N, Mao, H, Rolland, C, Gustafson, L, Xiao, T, Whitehead, S, Berg, AC, Lo, W-Y, Dollár, P & Girshick, R 2023, ‘Segment Anything’, arXiv, Cornell University.

Ma, G & Wu, Z 2020, ‘BIM-based building fire emergency management: Combining building users’ behavior decisions’, Automation in Construction, vol. 109, pp. 102975.

Pan, Y, Braun, A, Brilakis, I & Borrmann, A 2022, ‘Enriching geometric digital twins of buildings with small objects by fusing laser scanning and AI-based image recognition’, Automation in Construction, vol. 140, pp. 104375.

Qiao, Y, Lam, CC & Wong, MO 2024, ‘A GIS-Based Framework of Fire Risk Assessment and Emergency Route Planning for Heritage Buildings’, Lecture notes in operations research, Springer International Publishing, pp. 1721–1731.

Qiao, Y, Lam, CC, Wong, MO & Xu, Y 2024, ‘A Framework of Integrating HBIM and GIS for Automated Fire Risk Assessment of Heritage Buildings’, Proceedings of the 41st International Symposium on Automation and Robotics in Construction (ISARC), pp. 714-721.

Schönfelder, P, Aziz, A, Frédéric B & König, M 2024, ‘Enriching BIM models with fire safety equipment using keypoint-based symbol detection in escape plans’, Automation in construction, vol. 162, pp. 105382–105382.

Wang, W 2023, GitHub - CVHub520/X-AnyLabeling: Effortless data labeling with AI support from Segment Anything and other awesome models., GitHub, Available at: https://github.com/CVHub520/X-AnyLabeling.

Wong, MO & Lee, S 2022, ‘IFC-based information exchange for multi-agency response to indoor fire emergencies’, Automation in Construction, vol. 144, pp. 104623.

Wong, MO & Lee, S 2023, ‘Indoor navigation and information sharing for collaborative fire emergency response with BIM and multi-user networking’, Automation in Construction, vol. 148, pp. 104781.