Leveraging Deep Learning for Automated Reconstruction of Indoor Unstructured Elements in Scan-to-BIM

Mostafa Mahmoud [a, c], Wu Chen [a], Yang Yang [a], Tianxia Liu [a], Yaxin Li [a, b, *]

[a] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China
[b] Micro Dimension Technology Limited, Hong Kong, China
[c] Public Works Department, Cairo University, Giza, Egypt – mostafa.mahmoud@connect.polyu.hk, wu.chen@polyu.edu.hk, yanyang@polyu.edu.hk, tianxia.liu@connect.polyu.hk, yaxin-roan.li@polyu.edu.hk

Keywords: 3D reconstruction; building information modeling (BIM); point clouds; deep learning.

Abstract

Achieving automatic 3D reconstruction for indoor scenes is extremely useful in the field of scene understanding. Building information modeling (BIM) models are essential for lowering project costs, assisting in building planning and renovations, as well as improving building management efficiency. However, nearly all current available scan-to-BIM approaches employ manual or semi-automatic methods. These approaches concentrate solely on significant structured objects, neglecting other unstructured elements such as furniture. The limitation arises from challenges in modeling incomplete point clouds of obstructed objects and capturing indoor scene details. Therefore, this research introduces an innovative and effective reconstruction framework based on deep learning semantic segmentation and model-driven techniques to address these limitations. The proposed framework utilizes wall segment recognition, feature extraction, opening detection, and automatic modeling to reconstruct 3D structured models of point clouds with different room layouts in both Manhattan and non-Manhattan architectures. Moreover, it provides 3D BIM models of actual unstructured elements by detecting objects, completing point clouds, establishing bounding boxes, determining type and orientation, and automatically generating 3D BIM models with a parametric algorithm implemented into the Revit software. We evaluated this framework using publicly available and locally generated point cloud datasets with varying furniture combinations and layout complexity. The results demonstrate the proposed framework's efficiency in reconstructing structured indoor elements, exhibiting completeness and geometric accuracy, and achieving precision and recall values greater than 98%. Furthermore, the generated unstructured 3D BIM models keep essential real-scene characteristics such as geometry, spatial locations, numerical aspects, various shapes, and orientations compared to literature methods.

1. Introduction

Nowadays, 3D semantic models are valuable and practical for augmented and virtual reality (Xiao and Furukawa, 2014), autonomous interior navigation (Fang et al., 2022), and building maintenance and repairs (Chen and Tang, 2019). However, automatic reconstruction of corresponding 3D shapes and semantic information is hindered by the complexity of indoor surroundings with their inherent point cloud noise and incompleteness, especially for unstructured indoor objects like furniture (Kang et al., 2020). The demand for as-built building information modeling (AB-BIM) models is growing throughout the architecture, engineering, and construction fields. These models are essential for lowering project failure costs since they can be used as guides for planning and modifications. BIM is also extensively utilized in asset management and forms the foundation for assessing existing buildings (Bassier and Vergauwen, 2020). When initial designs for older or converted buildings are unavailable, scan-to-BIM, which turns laser-scanned point clouds into BIM models, is particularly helpful for producing AB-BIM models (Gankhuyag and Han, 2021).

BIM parametric models use geometric shapes and their adjustable features to operate as digital representations of designs. In the maintenance and operation phases, these models contain geometry and characteristics necessary for tasks like replacements, repairs, and performance evaluations (Chen and Tang, 2019). The scan-to-BIM process is usually labor-intensive and costly, requiring manual design efforts for a variety of scene elements. This process's technological and financial constraints prevent AB-BIM models from being widely improved. Recent work has focused on automating the scan-to-BIM process to alleviate the difficulties associated with AB-BIM modeling (Yang et al., 2020). Semantic segmentation accuracy in point clouds is enhanced for reconstruction using deep learning approaches. However, prior approaches give priority to necessary structural elements such as floors, walls, and ceilings (Romero-Jarén and Arranz, 2021) (Tang et al., 2022); separate modeling efforts are still required for various unstructured indoor items, such as chairs and tables.

Despite significant advances in the field, numerous issues persist in the scan-to-BIM reconstruction of indoor elements. Limited automation, difficulties obtaining all indoor features, and difficulty modeling incomplete point clouds due to scanning obstructions are among these problems. Current techniques do not effectively model structural apertures and neighboring items of the same class. They also have trouble taking into account the orientation of indoor elements and managing a variety of shapes within the same class. To overcome these challenges, we propose a new and efficient scan-to-BIM framework using deep learning algorithms and raw point cloud data. This framework automates the creation of 3D models for both structured and unstructured indoor features, aiming for parametric indoor scene representations. For structured elements like walls, doors, and windows, we present a method to accurately reconstruct 3D models for various room layouts, either Manhattan or non-Manhattan structures. We develop a robust workflow for unstructured elements to reconstruct 3D models, even with incomplete 3D point cloud information. This method also uses a parametric algorithm in the Revit platform to automatically generate BIM models, ensuring the models retain essential real-scene characteristics.

This paper is structured as follows: We start with a review of previous studies on indoor unstructured object modeling. Then, we present our proposed framework's methodology and discuss each stage in detail. Finally, we present the experiment design as well as the results and conclude with suggestions for future research.

2. Literature Review

2.1 Semantic Geometry-Based Approaches

The semantic geometry-based approach targets 3D geometry recovery of indoor objects by incorporating semantic features to improve modeling accuracy for high-level applications. It efficiently provides more straightforward and more meaningful representations. The study of (Shao et al., 2012) proposed a semi-automatic technique for modeling interior settings using sparse Kinect data. During RGB-D image segmentation, users label regions, and an algorithm arranges objects to recreate the scene. However, introducing new scenes or items not in the database may lead to inaccurate results.
Additionally, this approach involves user assistance in object segmentation. The method of (Nan et al., 2012) used a technique that combines classification with region-growing segmentation in laser scans to identify and label unique point clouds. Classification confidence is increased by progressively identifying individual items. Next, a deform-to-fit technique is used to customize 3D models using training data. Nevertheless, their method needs high-quality scans to produce realistic object-based models and precise scene segmentation. It also focuses on creating non-parametric 3D models.

The research of (Kim et al., 2012) proposed a two-stage approach: an offline pre-scanning step for each target object to search for stable primitives, followed by an online modeling step to segment the scene, fit segments with primitives, and match repeated objects hierarchically. Nevertheless, they haven't reconstructed structured components or provided BIM models. Difficulties include incomplete data that results in identification errors and preliminary segmentation constraints that cause problems with grouping or splitting, creating a recovery challenge. The framework of (Poux et al., 2018) presented an integrated technique for 3D geometric semantics reconstruction using a part-to-whole model. The method examines characteristics, object relationships, and contextual data from 3D furniture images stored in the Model-Net10 library. However, no automatic 3D BIM models are available, and point clouds are clustered using voxels, which may cause problems with linked clusters. Rather than using raw measurements, their approach works directly with segmented, not raw point clouds.

2.2 Deep Learning-Based Approaches

Deep learning methods excel in 3D object model reconstruction compared to the semantic geometry-based approach. They independently learn complex features, offer a deeper semantic understanding, and efficiently handle large-scale point cloud datasets (Bapat et al., 2023). Some studies focus on non-structural components (e.g., mechanical and electrical components), and recent reconstruction efforts in deep learning approaches are starting to include indoor object elements (Kufuor et al., 2021). The reference of (Park et al., 2022) presented a technique for indoor object-based point-net deep learning-based 3D model reconstruction. However, their accomplished models need human assistance to model the opening elements. Since they use a density-based clustering approach, it is challenging to identify closed class points independently. This method cannot consider different shapes of the same class and object's orientation.

The study of (Wang et al., 2022) presented a primitive-based reconstruction technique for indoor object retrieval that makes use of semantic segmentation classes and a model library. Nevertheless, neither automatically generated nor manually generated 3D BIM models were provided. Furthermore, they overlooked incorporating structural opening elements in 3D models and using Euclidean clustering for point clouds has limitations with connected clusters. The work of (Kim et al., 2023) used semantic segmentation for geometry from point clouds and materials from panorama images to perform an automated building object reconstruction. However, this method fell short when reconstructing the structural building elements.
Furthermore, consideration of the orientation of indoor objects is not included. Moreover, it is hard to differentiate between independent objects and consider different shapes of the same class. The research of (Mahmoud et al., 2024) proposed a system for reconstructing indoor point clouds and creating automatic BIM models for structured and unstructured features. However, their solution has limitations regarding its ability to handle unstructured features. It is based on bounding box modeling from density clustering, which struggles with connected clusters. Furthermore, their solution ignores objects of the same class with different shapes and requires user intervention to modify element orientations in BIM models.

This research introduces an innovative scan-to-BIM reconstruction framework based on deep learning to address the mentioned limitations. The proposed framework enables the automatic generation of 3D models for structured and unstructured indoor elements, offering parametric 3D BIM models of indoor scenes.

3. Methodology

The proposed overall 3D reconstruction framework, illustrated in Figure 1, comprises three primary stages: point cloud segmentation, structured element reconstruction, and unstructured element reconstruction. The following subsections elaborate on these three stages.

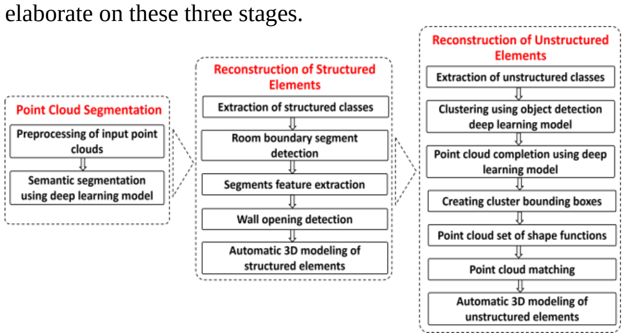


Figure 1. The proposed 3D reconstruction framework.

3.1 Point Clouds Segmentation

Point clouds undergo preprocessing and semantic segmentation in the outlined framework's first phase.

3.1.1 Preprocessing of Input Point Clouds: Building point clouds are often dense and extensive, making direct use of raw data impractical due to computational errors and time constraints. Subsampling and filtering of point clouds are essential to address these challenges. The linear distribution of plane projections within the building remains unaffected by subsampling, enhancing computational performance. Each voxel, produced by dividing the domain into equal segments with a fixed grid spacing, calculates the mean of all the points. A statistical filtering algorithm is applied to deal with outliers, eliminating points deviating farther from their neighbors than the point cloud's average ("Outlier removal," 2023).

3.1.2 Semantic Segmentation-Based Deep Learning Model: Semantic segmentation for 3D point clouds is essential for reconstructing parametric indoor environments because it assigns semantic names to components inside the point clouds. In (Hu et al., 2021), a neural network with an encoding-decoding technique serves as the basis for a large-scale semantic segmentation model. This model uses the S3DIS dataset ("S3DIS," 2023), which has 13 semantic classes, including space-forming (wall, ceiling, floor, column, beam, door, and window), space-occupying (table, chair, sofa, bookshelf, board), and clutter classes. The segmentation results help to implement the succeeding steps of the reconstruction framework.

3.2 Workflow of Structured Elements Reconstruction

The second stage proficiently identifies structured components within indoor buildings through point cloud analysis, as shown in Figure 1. Subsequent sections detail the steps for reconstructing these structured elements in 3D.

3.2.1 Room Boundary Segment Detection: This work detects Room boundary segments using subsampled points from the wall and column classes. The robust Random Sampling and Consensus (RANSAC) algorithm is used for this task. It iteratively estimates each detected surface's inliers, outliers, and plane parameters based on an input of predetermined thresholds. Additional steps employ the principal component analysis algorithm to separate inlier points and list line segments. Figure 2 shows an example of room line segment detection.

3.2.2 Segments Feature Extraction: Finding intersecting points that form the walls for modeling is the process of line feature extraction. On the other hand, it is not always accurate to sort detected lines according to their orientation or angle, which can result in inaccurate corner points. We introduce a novel line-sorting strategy for different layouts to address this issue. There are further steps in this process. It starts with the detected lines' endpoints and concludes that the lowest left line is the first line. Following that, parallel lines are removed, and the distances between the centers of intersecting lines and the current line are calculated. The line with the shortest distance is then picked as the following line, which is removed from the list, and the procedure is continued until all lines have been handled. This systematic approach groups lines based on spatial relationships and lengths, allowing for accurate room corner detection. This method facilitates the extraction of corner locations where lines consecutive intersect, as illustrated in Figure 2.

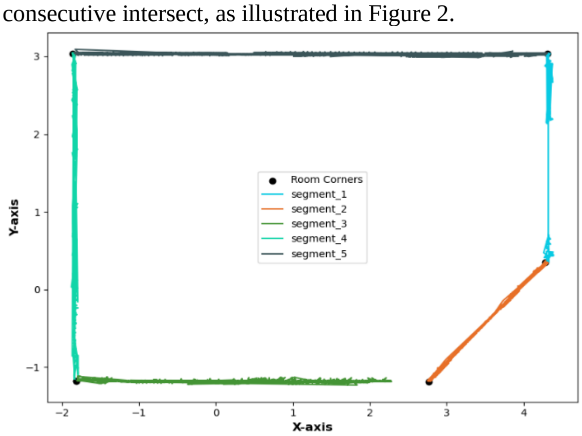


Figure 2. An illustration for wall segment detection with feature extraction

3.2.3 Wall Opening Detection: Our four-step approach reliably detects and localizes building openings using 3D point clouds. Firstly, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is used to extract and cluster door and window classes. Erroneous segmented clusters are eliminated using statistical analysis. Next, possible opening clusters are found by applying certain conditions. Next, using point cloud coordinates, the line identification technique retrieves critical characteristics for each cluster, such as its height, length, and center. This information is essential for the last stage, when opening points are fine-tuned to guarantee precise alignment with the underlying wall segment lines. The steps involved in refining are finding the closest line, projecting the opening points onto it, and exporting projected coordinates for modeling. In BIM technology, this refinement is vital for considering the spatial relationship of these elements.

3.2.4 Automatic 3D Modeling of Structured Elements: Parametric models for 3D BIM representations are generated by the automatic 3D modeling of structural elements. Information regarding opening points and room corners is exported for automated modeling. In Revit, dynamic visual programming replaces manual element manufacturing and streamlines the process while increasing efficiency. This algorithm's workflow is composed of groups, links, and nodes. In order to automatically create the wall, floor, ceiling, and opening components, input data is taken from a file, assigned to distinct coordinates, and utilized. Suitable creator nodes and Python scripts are then applied to model and edit structured elements. An important aspect missing from many earlier modeling approaches is the algorithm's ability to modify element dimensions based on processed dimensions for structured and unstructured elements. The results section presents the 3D BIM models of all aspects for different datasets.

3.3 Workflow of Unstructured Elements Reconstruction

In the third stage of the proposed framework, unstructured components within indoor buildings are reconstructed using point cloud analysis.

Extracting and Clustering of Indoor Classes: Indoor classes are retrieved and clustered for the individual item modeling in indoor unstructured element modeling. Based on our semantic segmentation results from the S3DIS dataset, these classes include bookcases, boards, sofas, tables, and chairs. A density clustering technique of DBSCAN has been attempted; however, it suffers from related items and needs parameter changes. In order to solve this, we propose applying the TR3D object detection deep learning network (Rukhovich et al., 2023). This network produces reliable results by giving indoor items' bounding box dimensions, which are then used to extract the point clouds.

3.3.2 Point Cloud Completion and Bounding Boxes Creation: Point cloud completion fills in the empty regions or gaps in a 3D point cloud that are brought about by occlusions, random features, and sensor limitations to produce a better and complete representation of an object or scene. In order to overcome this issue in point clouds, we used Snowflake-Net (Xiang et al., 2023), which enhances the incomplete point clouds by encoding and decoding layer structures. This neural network makes use of a transformer architecture, which is renowned for its capacity to use self-attention techniques to identify complicated patterns in sequential input. The three primary sections are feature extraction, seed generation, and point creation modules. For the creation of 3D models, obtaining the

https://doi.org/10.5194/isprs-archives-XLVIII-1-2024-479-2024

object positions and dimensions from the point cloud clusters is crucial. This can be achieved by using the orthogonal bounding boxes to the three principal axes. Spatial relationships and element connections provide essential benefits in BIM technology. Modeled window and door elements are connected to the wall element through Revit modeling, and walls are attached to floors and ceilings. Maintaining spatial connections with chairs, sofas, bookcases, and tables, the floor cluster serves as the host element. Precise bounding box measurements are required for indoor elements modeling. For further analysis, these bounding boxes establish each cluster's location and measurements (height, width, and length).

3.3.1 Point Cloud Set of Shape Functions: Semantic labeling quickly identifies the model type, but within the same class, such as tables, there are various shape variations, like dining tables or office tables. Feature matching compares distinct features from point clouds to determine the most similar indoor point cloud object. The set of shape function (SSF) descriptor is adequate for this, capturing unique shape characteristics. It comprises ten sets of 64-dimensional shape-function histograms (total dimensionality: 640) (Wohlkinger and Vincze, 2011), analyzing shape characteristics like angles, areas, and distances.

$$S(C_i, C_j) = \sqrt{\sum_{K=1}^{N} \left(C_{ik} - C_{jk}\right)^2} \tag{1}$$

Where S is the similarity, N is the number of dimensions, C_{ik} is the target point cloud SSF histogram values, and C_{jk} is the model point cloud SSF histogram values.

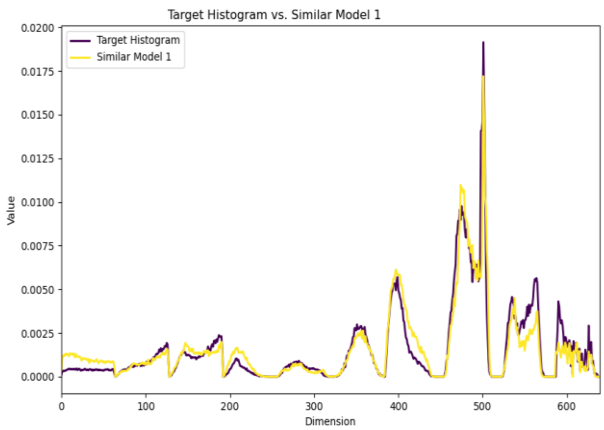


Figure 3. Set of shape function histograms for target and model point clouds.

This step aims to find the most similar indoor object point cloud for each cluster target file. We have created a model library using the Model-Net dataset (Wu et al., 2015) to categorize each point cloud cluster type, aligning with those from the S3DIS dataset in our framework. This library contains various shapes representing indoor classes, such as tables, sofas, chairs, and bookcases. Each indoor class includes tens of point cloud shapes, which are subsampled before calculating SSF histogram descriptors for comparison with the new target point cloud to determine its type. Figure 3 shows differences in SSF descriptor histograms between model and target point clouds. Various similarity metrics distinguish between the target and models, extracting the most similar model for each target point cloud. We used the Euclidean distance metric to compute the similarity between the histograms of the target and each library chair model. The most similar model is the one with the smallest Euclidean distance. Equation 1 measures the similarity index between target and library point cloud models of the same class.

3.3.2 Point Cloud Matching: Point cloud matching aligns two-point clouds in 3D space to find the optimal transformation for alignment. The target point cloud is registered to the model most similar to the previous step. For each target and its corresponding model, a five-step matching algorithm determines their orientation. Subsampling is the first step in this approach, which lowers processing costs and preserves essential points in both the target and model point clouds. Scale adjustment is then performed by computing size ratios along the bounding box's principal directions to align the model with the target. Then, to streamline the registration procedure, the translation is computed to align the centers of the target and model point clouds.

The following steps involve applying initial and final matching. The RANSAC technique is used for robust estimation, coarsely registering based on priors, and choosing the candidate with the lowest registration error. The fast Point Feature Histogram (FPFH) algorithm defines local point cloud shapes for initial matching. Using the iterative closest point (ICP) approach, we ensure fine registration in the final matching step by fine-tuning the coarse registration produced by initial matching. For vertically oriented indoor furniture alignment, the matching is limited to revolve mainly around the Z-axis. ICP reduces the sum of squared alignment errors between corresponding points (target (Q) and model (P)) using translation and rotation matrices described in Equations 2.

$$RMSE(t, R) = \sqrt{\frac{1}{N_p} * \sum_{i=1}^{N_p} \left\| p_i - (R * q_i + t) \right\|^2} \tag{2}$$

Where $RMSE$ is the matching room mean squared error, p_i is the reference point cloud, q_i is the corresponding target point cloud, t is the translation vector, R is the rotation matrix, and Np is the number of points.

3.3.3 Automatic 3D Modeling of Unstructured Elements: We combine information from previous steps: point cloud coordinates, dimensions, model type, and orientation for each object to perform automatic parametric modeling of indoor point cloud data. For 3D BIM modeling, we have established a BIM object library to enhance indoor scene representation. We then align these BIM files' characteristics for accuracy with the point cloud files in the SSF library. A Dynamo algorithm is implemented during the development of the BIM model. The process involves importing three-dimensional input data and assigning coordinates, dimensions, types, and orientations. Standardized items from the BIM library are then fitted to model unstructured components, with dimensions and orientation automatically modified to achieve the desired outcome.

4. Results And Discussion

4.1 Experimental Dataset

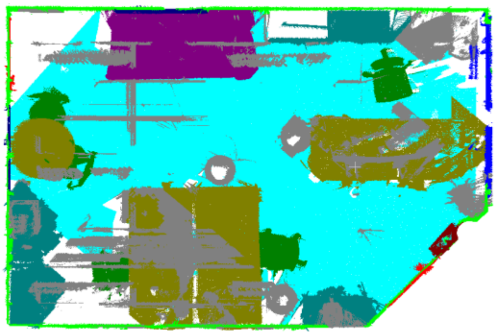
Two datasets—locally acquired and publicly accessible—were used to assess the effectiveness and precision of the suggested framework. These datasets showcase a variety of furniture configurations and different room layouts. The research dataset ("UZH," 2023) is the source of the publicly accessible datasets, as depicted in Figure 4 (a). Using the Nav-Vis M6 mobile 3D scanner, the locally collected dataset was observed on the campus of Hong Kong Polytechnic University. This dataset represents a lecture room, as shown in Figure 4 (b).

(a)

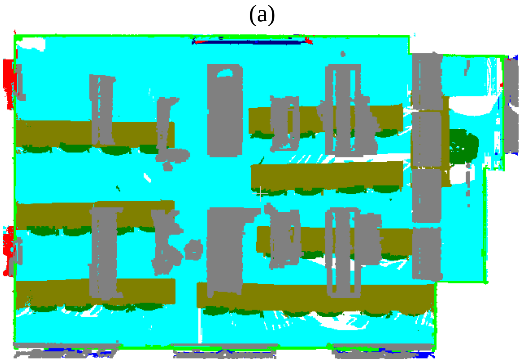
(b)

Figure 4. The datasets utilized to assess our methodology: (a) publicly accessible dataset (dataset 1) and (b) locally acquired dataset (dataset 2).

4.2 Semantic Segmentation Results



(a)



(b)

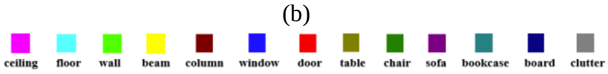


Figure 5. The results of the semantic segmentation: (a) dataset 1 and (b) dataset 2.

The first stage of our methodology is point cloud segmentation, which involves preprocessing and semantic segmentation. Initially, all data undergoes preprocessing, including filtering and down-sampling with specific parameters. RGB values and point cloud coordinates serve as input for semantic segmentation. Figure 5 displays semantic segmentation outcomes on publicly accessible and locally acquired datasets using our encoding-decoding deep learning model. The segmentation output classifies semantic categories into structured (e.g., floor, ceiling, wall, beam, column, window, door) and unstructured (e.g., chair, bookcase, sofa, table, board) classes, indicating suitability for reconstruction purposes in both structured and unstructured stages.

4.3 3D models Reconstruction of Structured Elements

Our framework's second stage involves reconstructing structured elements utilizing room boundary segment detection, feature extraction, wall surface object detection, and automatic 3D modeling. Indoor structured classes are first determined. The iterative RANSAC technique is used for line detection and corner extraction tasks based on wall and column classes. Our sorting algorithm uses detected lines as input, guaranteeing that sets of lines are arranged logically for various layouts. The corners of the room are defined by subsequent corner point detection. Wall surface object recognition is used to identify window and door openings, providing dimensions and insertion center locations for automatic modeling. Opening refinement addresses missing the host element problem in BIM modeling and guarantees accurate 3D modeling. The excellent accuracy of our opening identification method is validated by Figure 7, which shows the precise recognition of door and window elements aligned with actual numbers for datasets 1 and 2. Information about opening elements and detected room corners are gathered and exported for 3D modeling of structured elements.

As shown in Figure 7, the last stage of reconstruction entails automatic 3D modeling in Revit using our Dynamo method. The resultant models accurately depict structural components and wall surface objects. Our method accommodates a wide range of space configurations, from simple to intricate ones. The generated BIM models include walls, floors, ceilings, and wall openings; the ceiling has been removed for clarity. These models show the high detection accuracy of the structured elements compared to real input point clouds. The structured 3D models were evaluated using the method outlined by the ISPRS research group (Khoshelham et al., 2021), which calculates precision (P) and recall (R). These metrics, defined in equations (3–4), provide a comprehensive assessment and comparison of the generated models against reference models.

$$P = \frac{Area_{pred\ model} \cap Area_{ref\ model}}{Area_{pred\ model}} \quad (3)$$

$$R = \frac{Area_{pred\ model} \cap Area_{ref\ model}}{Area_{ref\ model}} \quad (4)$$

Precision measures the correctness of anticipated model areas by calculating the intersection area of accurate and reconstructed models over the total reconstructed area. Recall measures accuracy by calculating the ratio of the intersection area to the total actual area. Table 1 shows a complete evaluation of each model, including precision and recall measures. The results show high accuracy, precision, and recall levels greater than 98%. This implies a high degree of resemblance between the produced models and the hand-crafted reference models.

Dataset	Intersection area (m2)	Predicted area (m2)	Reference area (m2)	P	R
Dataset 1	24.355	24.748	24.498	0.984	0.994
Dataset 2	42.860	43.245	43.164	0.991	0.993

Table 1. Evaluation of the reconstructed BIM models based on predicted areas.

4.4 3D Models Reconstruction of Unstructured Elements

The third stage involves recreating unstructured parts. This comprises clustering indoor classes, completing point clouds, generating bounding boxes, applying shape functions, matching orientations, and automating 3D modeling.

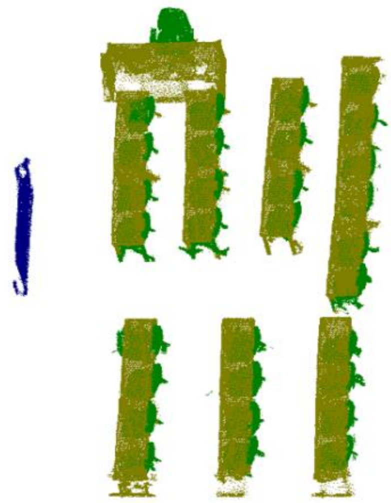
4.4.1 Indoor Classes Extracting and Clustering Results: Indoor classes are extracted for unstructured element construction, including sofas, chairs, boards, tables, and bookcases. The DBSCAN algorithm is then used to cluster objects based on point cloud density. However, accuracy is hindered due to interconnected point cloud objects, as illustrated in Table 2, which shows the dataset clustering results. The detection accuracy of derived indoor classes ranges between 60% and 100%. We use the TR3D object identification deep learning model to address this issue, which achieves 100% accuracy when extracting all indoor classes compared to density clustering. Applying this method guarantees that the number of modeled elements is identical to the actual existing elements.

Dataset	Number of extracted objects (Detected/Actual)				
	Sofa	Bookcase	Board	Table	Chair
Dataset 1	1/1	3/5	2/2	3/4	4/5
Dataset 2	-	-	1/1	8/8	29/31

Table 2. Detection accuracy of extracting indoor classes using density clustering method for all datasets.

4.4.2 Point Cloud Completion and Bounding Box Creation Results: Point cloud completion is critical for filling in missing components of discovered clusters and introducing better representation shapes. Existing reconstruction algorithms frequently struggle with incomplete point clouds due to scanning obstructions and inherent point cloud features, resulting in the loss of crucial indoor scene elements. We use the Snowflake-Net deep learning model to address this issue in point cloud completion. Figure 6 depicts the completion outcomes for indoor classrooms from datasets 1 and 2. These results better affect the next steps of type and rotation determination. After identifying indoor clusters and completing the point clouds, bounding boxes are created to characterize each cluster's position and dimensions. The bounding box computations mainly consider the floor class to provide correct space modeling and maintain spatial relationships among indoor objects.

(a)



(b)

Figure 6. Point cloud completion results: (a) dataset 1, and (b) dataset 2.

4.4.3 Point Cloud Set of Shape Functions and Matching Results: In this stage, we use a feature matching technique, specifically the SSF descriptor, because of its ability to capture distinct shape characteristics. This approach successfully recognizes objects with different geometric patterns by generating histograms for each model in the library and target point cloud clusters. Using the Euclidean distance metric, we compute a similarity index for each cluster, keeping only the most similar model types for parametric BIM modeling. These results, which include the model type for each cluster and the associated selected point clouds, are then applied in the matching step.

Each target point cloud is matched with the corresponding similar model during the point cloud matching procedure. Subsampling reduces costs and identifies significant points, followed by scale equivalence and translation to improve alignment. The matching algorithm then employs FPFH and ICP algorithms to provide reliable results. RMS matching errors are determined to assess the alignment between computed orientation values and the actual orientation of indoor point clouds, which impacts 3D modeling accuracy.

4.4.4 Automatic 3D Modeling Results of Unstructured Elements: The final step incorporates the results of the preceding stages into the 3D modeling process. Input data comprises detected indoor cluster positions, dimensions, model types, and orientations. The Dynamo method in Revit is then used to perform automatic parametric calculations. This procedure uses imported information to reconstruct 3D BIM models for all elements, as Figure A1 in Appendix A shows. It shows a screenshot of the parametric BIM Dynamo method (left) and the automatically generated 3D BIM models (right) in the Revit platform, with reconstruction results for datasets 1 and 2. The automatically created unstructured 3D BIM models are highly accurate in element locations, dimensions, model types, and orientations, demonstrating the efficiency of our methodology. Tables 3 and 4 measure the determination accuracy of model types and orientations, respectively. Table 3 shows that all modeled element types were adequately identified, with just one chair misclassification in dataset 1. The average accuracy of type determination across all datasets is 100% for sofas, bookcases, boards, and tables and 90% for chairs.

Furthermore, Table 4 shows that the orientations of most modeled components are precisely acquired, resulting in excellent modeling accuracy. After modeling, a few items needed to be rotated slightly: one chair and one table for datasets 1 and 2. The average accuracy of orientation determination across all datasets is 100% for sofas, bookcases, and boards and 94% and 90% for tables and chairs, respectively. Figure 7 shows the refined 3D BIM models for all datasets. These models were slightly refined according to Tables 3 and 4, with small rotations and model-type adjustments made using the BIM algorithm. These final 3D BIM models demonstrated the high accuracy of our proposed reconstruction framework.

Dataset	Number of reconstructed objects remaining unchanged in type (Detected/Total)				
	Sofa	Bookcase	Board	Table	Chair
Dataset 1	1/1	5/5	2/2	4/4	4/5
Dataset 2	-	-	1/1	8/8	31/31

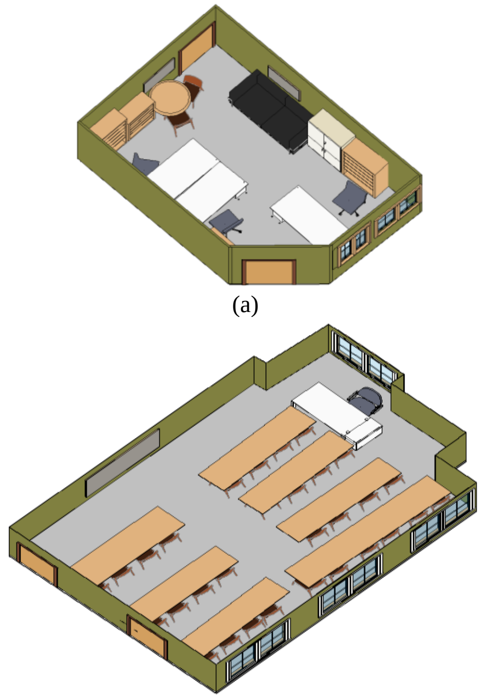
Table 3. Automatic 3D BIM determination accuracy of object model type.

Dataset	Number of reconstructed objects remaining unchanged in orientation (Detected/Total)				
	Sofa	Bookcase	Board	Table	Chair
Dataset 1	1/1	5/5	2/2	4/4	4/5
Dataset 2	-	-	1/1	7/8	31/31

Table 4. Automatic 3D BIM determination accuracy of object model orientation.

5. Conclusion And Future Work

BIM models are critical for lowering project execution costs and are utilized in planning and renovations. Traditional scan-to-BIM systems usually employ manual or semi-automatic techniques that require significant time and effort. These approaches concentrate on modeling key structural objects while disregarding other indoor unstructured elements due to difficulties in modeling incomplete point clouds and capturing detailed indoor scenes. They also lack accuracy, fail to consider indoor objects' various shapes and orientations and use erroneous clustering approaches. This effort aims to automate the scan-to-BIM process by reconstructing indoor scenes from input point clouds. We present an efficient framework using deep learning techniques to produce 3D models of structured and unstructured indoor elements. Our method effectively reconstructs precise 3D models of structured elements with varied layouts while ensuring the correct reconstruction of unstructured elements to preserve real-world properties. The proposed framework's efficiency and precision were assessed using publicly accessible and locally acquired datasets that depicted various furniture arrangements and layout difficulties. The results demonstrate the framework's ability to reconstruct indoor structured elements with completeness and geometric accuracy. Furthermore, the created 3D models accurately depict real-world elements, including geometry, spatial locations, and different shapes for unstructured objects. Future work will include growing the BIM model library, increasing the point cloud library, and incorporating actual dataset samples to improve feature matching, point cloud registration, and overall model accuracy.



(a)

(b)

Figure 7. Final Resulted in 3D BIM structured and unstructured models: (a) dataset 1 and (b) dataset 2.

References

Bapat, H., D. Sarkar, R. Gujar, 2023. Application of multi-criteria decision-making for evaluation of key performance indicators of integrated project delivery and BIM model for an infrastructure transportation project in Western India. Int. J. Constr. Manag., 23 (12), 2077–2086. doi: 10.1080/15623599.2022.2040077.

Bassier, M., M. Vergauwen, 2020. Unsupervised reconstruction of Building Information Modeling wall objects from point cloud data. Autom. Constr., 120, 103338. doi: 10.1016/j.autcon.2020.103338.

Chen, C., L. Tang, 2019. BIM-based integrated management workflow design for schedule and cost planning of building fabric maintenance. Autom. Constr., 107, 102944. doi: 10.1016/j.autcon.2019.102944.

Fang, Q., X. Xu, X. Wang, Y. Zeng, 2022. Target-driven visual navigation in indoor scenes using reinforcement learning and imitation learning. CAAI Trans. Intell. Technol., 7 (2), 167–176. doi: 10.1049/cit2.12043.

Gankhuyag, U., J.-H. Han, 2021. Automatic BIM Indoor Modelling from Unstructured Point Clouds Using a Convolutional Neural Network. Intell. Autom. Soft Comput., 28 (1), 133–152. doi: 10.32604/iasc.2021.015227.

Hu, Q., B. Yang, L. Xie, S. Rosa, Y. Guo, 2021. Learning semantic segmentation of large-scale point clouds with random sampling. IEEE Trans. Pattern Anal. Mach. Intell., 44 (11), 8338–8354. doi: 10.1109/TPAMI.2021.3083288.

Kang, Z., J. Yang, Z. Yang, S. Cheng, 2020. A review of techniques for 3D reconstruction of indoor environments. ISPRS Int. J. Geo-Information, 9 (5), 1–21. doi: 10.3390/ijgi9050330.

Khoshelham, K., H. Tran, D. Acharya, L.D. Vilariño, Z. Kang, 2021. Results of the ISPRS benchmark on indoor modeling. ISPRS Open J. Photogramm. Remote Sens., 2 (August), 100008. doi: 10.1016/

j.ophoto.2021.100008.

Kim, S., K. Jeong, T. Hong, J. Lee, J. Lee, 2023. Deep Learning–Based Automated Generation of Material Data with Object–Space Relationships for Scan to BIM. J. Manag. Eng., 39 (3), 4023004. doi: 10.1061/jmenea.meeng-5143.

Kim, Y.M., N.J. Mitra, D.-M. Yan, L. Guibas, 2012. Acquiring 3d indoor environments with variability and repetition. ACM Trans. Graph., 31 (6), 1–11. doi: 10.1145/2366145.2366157.

Kufuor, J., D.D. Mohanty, E. Valero, F. Bosché, 2021. Automatic MEP component detection with deep learning. International Conference on Pattern Recognition., Springer, 373–388. doi:10.1007/978-3-030-68787-828.

Mahmoud, M., W. Chen, Y. Yang, Y. Li, 2024. Automated BIM Generation for Large-Scale Indoor Complex Environments Based on Deep Learning. Autom. Constr., 162, 105376. doi:10.1016/j.autcon.2024.105376.

Nan, L., K. Xie, A. Sharf, 2012. A search-classify approach for cluttered indoor scene understanding. ACM Trans. Graph., 31 (6), 1–10. doi: 10.1145/2366145.2366156.

Outlier removal, 2024. https://www.open3d.org/docs/release/tutorial/geometry/pointcloud_outlier_removal.html (accessed 7 August 2023).

Park, J., J. Kim, D. Lee, K. Jeong, J. Lee, 2022. Deep learning–based automation of scan-to-BIM with modeling objects from occluded point clouds. J. Manag. Eng., 38 (4), 4022025. doi: 10.1061/(asce)me.19435479.0001055.

Poux, F., R. Neuville, G.A. Nys, R. Billen, 2018. 3D point cloud semantic modeling: Integrated framework for indoor spaces and furniture. Remote Sens., 10, 1–35. doi: 10.3390/rs10091412.

Romero-Jarén, R., J.J. Arranz, 2021. Automatic segmentation and classification of BIM elements from point clouds. Autom. Constr., 124, 103576. doi: 10.1016/j.autcon.2021.103576.

Rukhovich, D., A. Vorontsova, A. Konushin, 2023. Tr3d: Towards real-time indoor 3d object detection. arXiv Prepr, arXiv2302.02858. doi: 10.48550/arXiv .2302.028 58.

S3DIS, 2023. https://buildingparser.stanford.edu/dataset.html (accessed 12 August 2023).

Shao, T., W. Xu, K. Zhou, J. Wang, D. Li, 2012. An interactive approach to semantic modeling of indoor scenes with an rgbd camera. ACM Trans. Graph., 31 (6), 1–11. doi: 10.1145/ 2366145.2366155.

Tang, S., X. Li, X. Zheng, B. Wu, W. Wang, 2022. BIM generation from 3D point clouds by combining 3D deep learning and improved morphological approach. Autom. Constr., 141, 104422. doi: 10.1016/j.autcon.2022 .104422.

UZH, 2024. https://www.ifi.uzh.ch/en/vmml/research/datasets.htm (accessed 11 August 2023).

Wang, T., Q. Wang, H. Ai, L. Zhang, 2022. Semantics-and-Primitives-Guided Indoor 3D Reconstruction from Point Clouds. Remote Sens., 14 (19), 4820. doi: 10.3390/rs14194820.

Wohlkinger, W., M. Vincze, 2011. Ensemble shape functions for 3D object classification. IEEE International Conference on Robotics and Biomimetics, ROBIO 2011. IEEE. 2987–2992. doi: 10.1109/ROBIO. 2011.6181760.

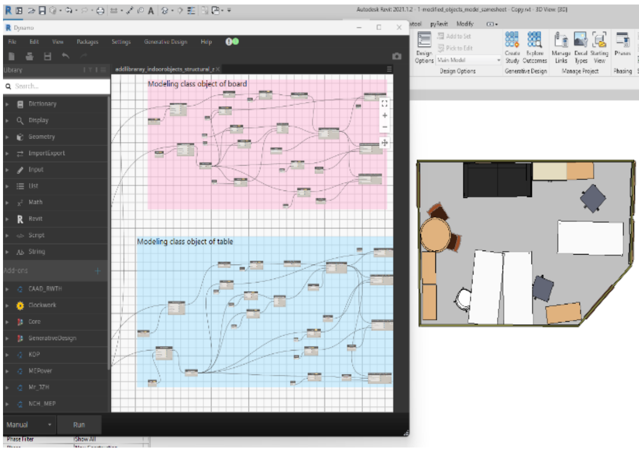
Wu, Z., S. Song, A. Khosla, F. Yu, L. Zhang, 2015. 3D ShapeNets: A deep representation for volumetric shapes. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 07-12-June, 1912–1920. doi: 10.1109/ CVPR.2015.7298801.

Xiang, P., X. Wen, Y.S. Liu, Y.P. Cao, P. Wan, 2023. Snowflake Point Deconvolution for Point Cloud Completion and Generation With Skip-Transformer. IEEE Trans. Pattern Anal. Mach. Intell., 45 (5), 6320–6338. doi: 10.1109/TPAMI.2022.3217161.

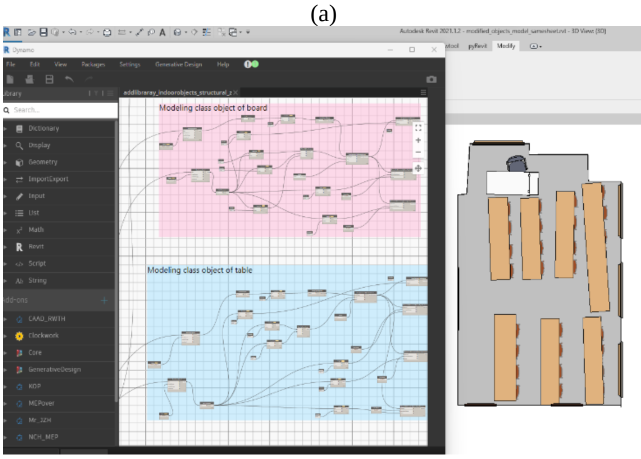
Xiao, J., Y. Furukawa, 2014. Reconstructing the world's museums. Int. J. Comput. Vis., 110, 243–258. doi: 10.1007/s11263-014-0711-y.

Yang, L., J.C.P. Cheng, Q. Wang, 2020. Semi-automated generation of parametric BIM for steel structures based on terrestrial laser scanning data. Autom. Constr., 112, 103037. doi: 10.1016/j.autcon.2019. 1030 37.

Appendix A



(a)



(b)

Figure A1. 3D BIM modeling using Dynamo Revit (left) and automatically created obtained models (right): (a) dataset 1 and (b) dataset 2.