BIM-to-Labelled Point Cloud: Automated Point Cloud Annotation from BIM Models using Bounding Boxes and Solid Geometries

Saad Boudarbala[1,2], Tania Landes[2], Hélène Macher[2], Thibault Bavoux[1]

[1] Futurmap, 90 Rue Paul Bert, 69003 Lyon, France - (saad, thibault)@futurmap.com
[2] Université de Strasbourg, CNRS, INSA Strasbourg, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000 Strasbourg, France - (saad.boudarbala, tania.landes, helene.macher)@insa-strasbourg.fr

Keywords: Point cloud, BIM model, automatic labelling, solid geometry, bounding box

Abstract

This paper presents an automated framework for generating semantically labelled building point clouds from their corresponding BIM models. The proposed methodology aims to facilitate the creation of training datasets for deep learning–based indoor semantic segmentation. Two complementary labelling strategies are introduced. The first relies on bounding boxes (BBX) extracted from BIM elements to efficiently assign labels to points based on volumetric inclusion. The second approach uses solid geometry and a nearest-neighbour principle (SG-NN) to compute distances between BIM object meshes and the point cloud, enabling a more precise spatial correspondence. In addition, a room-based geometric grouping strategy is proposed to structure the annotated point clouds into spatial units compatible with common indoor segmentation datasets. The methods are evaluated through a qualitative analysis on several real building datasets of different typologies and acquisition conditions, as well as through a quantitative evaluation based on a manually segmented reference point cloud. Results show that the SG-NN approach achieves higher performance, with an average Recall of 92% and IoU of 88%, compared to 87% of Recall and %78 of IoU for the BBX approach. While the BBX approach provides faster processing, the SG-NN strategy achieves higher labelling accuracy, particularly for geometrically complex elements. The proposed workflow enables scalable dataset generation from Scan-to-BIM projects while significantly reducing manual annotation effort.

1. Introduction

Recent advances in deep learning have significantly improved the performance of 3D semantic segmentation models, particularly for building-scale indoor point cloud data. However, the success of such models heavily relies on the availability of large, high-quality annotated datasets. Creating reliable training data remains a labor-intensive and time-consuming task, especially as the level of detail and the number of semantic classes increase. In the domain of indoor scene understanding based on point clouds, well-known benchmark datasets such as S3DIS (Armeni et al., 2016) and ScanNet (Dai et al., 2017), as well as synthetic point clouds generated from digital models (Hu et al., 2024), have supported the progress of semantic segmentation research. Nevertheless, there remains a strong need for approaches capable of automatically labelling real-world point clouds based on their corresponding digital building models.

When BIM models are created from point clouds, the semantic information contained in the BIM can be transferred to the point cloud, effectively creating labelled training data. Several recent works have explored automatic labelling of building point clouds using BIM models, relying on spatial matching strategies between BIM elements and laser scans. Other approaches have also investigated combining real labels with pseudo-labels derived from 3D models to improve indoor segmentation datasets. However, these methods often face limitations related to geometric precision, scalability to large datasets, or practical deployment in real industrial contexts. In addition, approaches based on synthetic data generation do not fully reproduce the acquisition characteristics of real scans, such as noise, occlusions, and realistic point densities.

Building on this line of research, this study proposes an automated approach for annotating building point clouds using their corresponding BIM models. Unlike approaches based on synthetic data generation, the proposed approach operates directly on real-world point clouds, preserving their realistic acquisition characteristics. The BIM models used in this process are generated from the point clouds and validated by specialists through multiple quality-control procedures, ensuring reliable semantic information. In many industrial Scan-to-BIM workflows, point clouds are systematically accompanied by such validated BIM models, making the proposed approach particularly suitable for large-scale applications. In this configuration, semantic information can be automatically transferred from the BIM model to the point cloud, enabling the annotation of large datasets with minimal additional manual effort and with a flexible number of semantic classes. The workflow relies on efficient geometric operations and widely available open-source tools, making its implementation accessible and scalable for practical use. Furthermore, the study proposes an approach to perform room-based geometric grouping of the resulting annotated point clouds, ensuring their suitability for existing deep learning-based semantic segmentation models.

2. Related work

Deep learning models for 3D semantic segmentation require large-scale annotated datasets to achieve robust and generalizable performance. Dataset quality is critical not only in terms of data volume, but also with respect to geometric accuracy and the diversity and number of semantic classes. In the context of building point clouds, producing such datasets through manual annotation is labor-intensive and difficult to scale.

3D semantic segmentation of building point clouds supports a wide range of applications, including facility management, BIM model generation, renovation planning, and robotics. Among these domains, Scan-to-BIM workflows particularly benefit from reliable semantic segmentation, as object-level understanding directly supports model reconstruction and updating. This has been demonstrated in several recent studies (Tang et al., 2022; Zhang and Zou, 2023; Campagnolo et al., 2023; Roman et al., 2024). Conversely, BIM models contain rich and structured semantic information that can potentially be transferred back to point clouds. When properly leveraged, this information can serve as a valuable source of annotated data for training deep learning models. In response to the challenges of manual annotation, several research directions have explored automated dataset generation strategies from BIM models. These approaches can be broadly categorized into BIM/CAD to real point cloud label transfer and synthetic data generation from digital models.

2.1 BIM/CAD-to-Point cloud label transfer

A major research direction leverages the semantic richness of BIM or CAD models to automatically annotate their corresponding real point clouds. These approaches assume geometric alignment between the digital model and the scan and transfer object labels through spatial correspondence.

Several works rely on volumetric or bounding-box-based matching. (Yoo et al., 2024) project BIM object bounding boxes into the point cloud space, assigning labels to points located within each volume, optionally after segmentation by floor or room to reduce ambiguity. Similar principles are used in semi-automated dataset generation workflows for indoor scenes. Other studies adopt distance-based or nearest-neighbour geometric matching between scan points and BIM elements. (De Geyter et al., 2022) rely on synthetic point clouds derived from BIM models as an intermediate representation, where each point is labelled according to the nearest BIM object using geometric distance criteria. Extending this idea to industrial environments, (Birkeland and Udnæs, 2024) use geometric Digital Twins (gDTs) - these are geometry-accurate digital models used to visualize, analyze, or simulate a building's real-world behaviours in digital twin applications - to transfer semantic labels to point clouds of complex facilities. These approaches improve assignment precision compared to pure volumetric methods but depend strongly on accurate registration and modelling quality. More advanced strategies combine geometric matching with pseudo-labelling or hybrid supervision. For example, (Mérizette et al., 2025) propose combining real annotations with pseudo-labels derived from 3D models to enhance indoor segmentation datasets. Collectively, these BIM-driven approaches demonstrate that model-to-scan correspondence can significantly reduce manual annotation effort while preserving semantic consistency.

2.2 Synthetic Point Cloud Generation

Another widely explored direction is the generation of synthetic labelled point clouds directly from BIM or parametric models to enlarge training datasets and control scene composition. In this case, labels are inherently available since sampled points inherit the semantic class of their source BIM objects, but the main challenge lies in reproducing realistic acquisition characteristics such as noise, density variations, and occlusions.

Several studies address this issue by introducing increasingly realistic simulation pipelines. (Ma et al., 2020) proposed a practical workflow to convert BIM models from commercial software into synthetic point clouds and showed that combining synthetic and real data improves segmentation performance, while training on synthetic data alone suffers from a domain gap. Similarly, (Huang et al., 2022) generated parametric BIM scenes and produced viewpoint-guided sampled point clouds, introducing Gaussian noise to simulate laser scanning conditions and improve realism. Building on these ideas, (Hu et al., 2024) combined automatic BIM-to-point-cloud generation with domain adaptation techniques. Their framework produces labelled synthetic point clouds using FME and integrates a domain adaptation network (DawNet) to reduce the distribution gap between synthetic and real-world data.

More generally, the synthetic-to-real problem has been widely investigated in point cloud research. Outdoor synthetic datasets such as SynLiDAR by (Xiao et al., 2021) and SqueezeSegV2 by (Wu et al., 2019), as well as translation frameworks such as PCT (Bai et al., 2023), aim to reduce the gap between simulated and real scans by modeling differences in appearance, density, and sparsity.

While synthetic data generation provides scalable labelled datasets, the remaining gap between simulated and real acquisitions often limits direct applicability to real-world scans. In contrast, when a point cloud is associated with its corresponding CAD or BIM model, the semantic information embedded in the digital model can be directly transferred to the data. Building on this principle, this study proposes two automated labelling strategies that exploit the geometric and semantic correspondence between the digital model and the real point cloud. Unlike synthetic generation approaches, the proposed approaches operate directly on real scans, preserving their inherent acquisition characteristics such as noise, occlusions, and density variations, while automatically transferring semantic labels from the model to the point cloud.

3. Methodology

Our work focuses on the automation of point cloud labelling using semantic labels extracted from corresponding digital building models. To this end, we developed two different approaches. The first approach relies on the BBX of building components, while the second is based on solid geometry and a nearest-neighbour principle to achieve a more detailed correspondence.

3.1 Bounding box based approach

The first approach uses bounding boxes derived from the digital building model. This geometric approach allows each BIM object to be represented by its spatial extent, providing a simple and efficient way to identify the points belonging to that object. To ensure both datasets share the same coordinate system, transformation parameters applied to the point cloud during model creation are determined. After aligning the point cloud and BIM elements, the labelling process is then performed. Each point within a BBX inherits the object's semantic label. Points that do not belong to any of the targeted semantic classes are assigned to a residual "clutter" class, which groups all remaining elements of the point cloud. This approach is efficient and simple. However, it may include irrelevant points when the bounding box does not tightly fit the actual object geometry. As object shapes become more complex, this approximation becomes less precise. Nevertheless, the approach remains fast and practical for large-scale point cloud labelling. The technical implementation of this approach is illustrated in Figure 1 and detailed below.

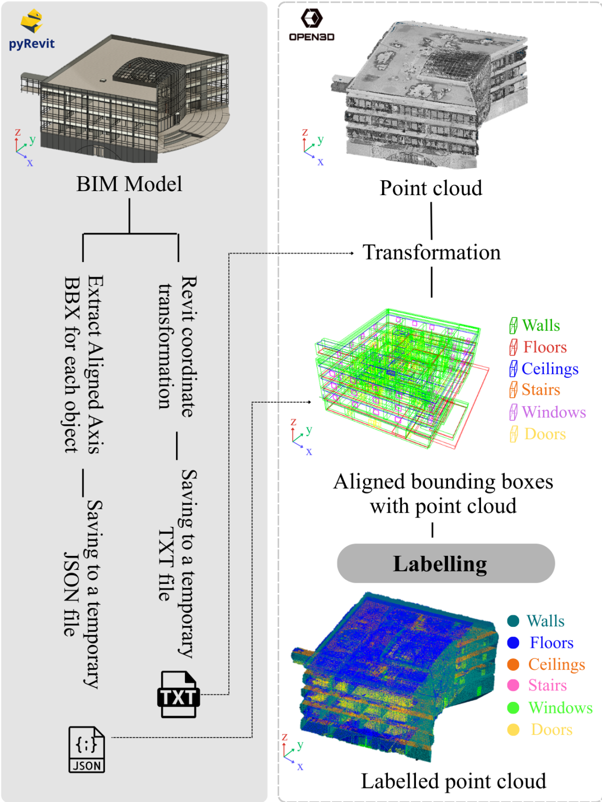

Figure 1. Bounding-box-based method for automated labelling of building point clouds

To implement this approach, a dedicated workflow was developed to extract and process the geometric information from the BIM model. The BIM data were accessed through the Revit API using PyRevit, allowing direct manipulation of the native RVT models without exporting them to intermediate formats such as IFC. A PyRevit plugin was developed to extract the BBX of the targeted elements together with the transformation parameters of the Revit coordinate system, which were exported in JSON format.

Due to PyRevit limitations (based on IronPython) regarding external libraries, point cloud processing was performed in a separate Python environment using Open3D. The extracted transformation matrices were applied to align the point cloud with the BIM coordinate system, and each point located within a bounding box was assigned the semantic label of the corresponding BIM object. However, Revit BBX are axis-aligned with the global coordinates rather than the true object orientation, which can include irrelevant points for non-orthogonal elements. To reduce such errors, geometric constraints were introduced to specific classes, enforcing verticality for walls and horizontality for floors or ceilings.

3.2 Solid geometry and nearest-neighbour based approach

The second approach directly annotates point clouds using the solid geometries of digital model objects, ensuring correspondence between points and elements while minimizing noise, i.e., points that are incorrectly labelled or do not belong to a given element. In this context, solid geometries refer to the mesh representations of BIM elements, which explicitly define the surfaces and volumes of each object. The approach relies on computing orthogonal distances between the point cloud and these model surfaces to assign labels according to predefined thresholds. For example, thresholds of 5 cm for walls and slabs and 10 cm for other classes were used. These values were selected to reflect typical modelling accuracy while accounting for the geometric characteristics of different object categories. Structural elements such as walls and slabs generally correspond to large planar surfaces with high geometric fidelity, allowing the use of a smaller tolerance. In contrast, elements such as doors, windows, or stairs are often represented as loadable BIM families whose parametric geometries may differ slightly from the scanned shapes. A larger tolerance (10 cm) is therefore introduced for these objects to accommodate modelling simplifications. The process is fully automated through a modular pipeline that exports object geometries by category (e.g., walls, floors, ceilings) in a CAD format, aligns the point cloud to the same coordinate system, computes cloud/surface distances and assigns labels to points that fall within the predefined minimum distance thresholds. Similarly to the BBX approach, points not associated with the targeted semantic classes are assigned to a clutter category representing the remaining elements of the point cloud.

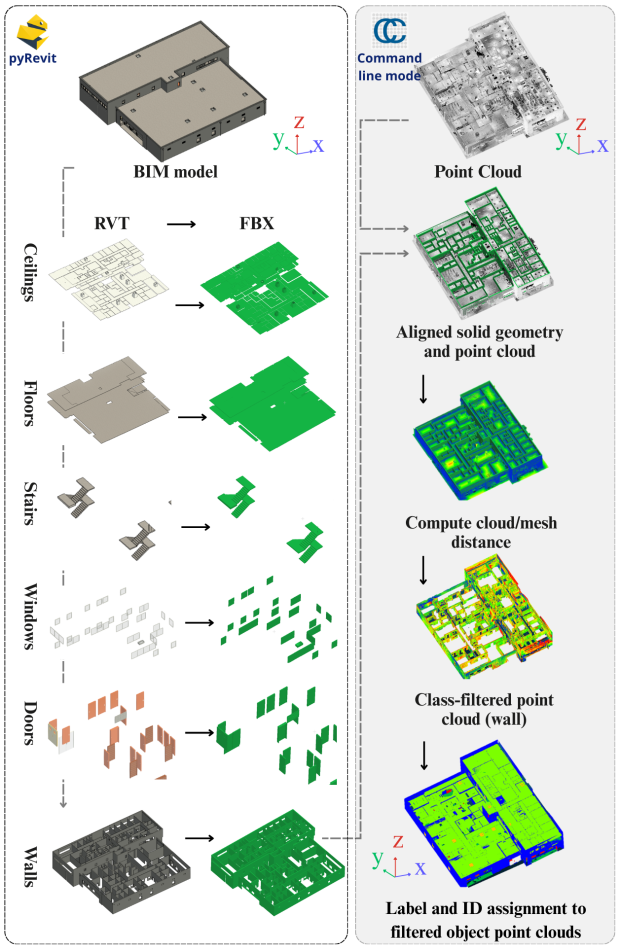

Figure 2. Solid geometry and nearest-neighbour (SG-NN) based approach

To technically implement this approach, custom plugins based on the PyRevit framework (left bloc of Figure 2) were first developed to export the solid geometries of BIM objects by category (e.g., walls, floors, ceilings) as mesh models in FBX format. The FBX format was selected because it preserves mesh geometries and transformations while remaining lightweight and compatible with 3D processing tools such as CloudCompare, used in the subsequent processing stages. As no native Revit tool supports this type of extraction, one plugin was developed to retrieve object meshes from the digital model, while another aligns the point cloud with the coordinate system of the exported geometries.

Once the geometries and point cloud were prepared, the distance computation and filtering stages were performed using CloudCompare through command-line scripts (right bloc of Figure 2). The mesh–cloud distance tool was used to compute the distance between each object mesh and the point cloud. Points located within the predefined distance thresholds were retained and assigned the semantic label of the corresponding object category. The resulting class-specific point clouds were then merged to produce the final labelled dataset. These operations were fully automated through a batch-processing pipeline that sequentially processes all object categories.

3.3 Room-based geometric grouping strategy

Beyond the labelling process, this study also introduces an approach for the geometric grouping of annotated point clouds for deep learning applications. Since many segmentation models require spatially coherent subsets, an approach was adopted inspired by datasets like S3DIS, organizing point clouds into room-based semantic units. Starting from labelled point clouds and their digital models, room volumes are generated by extruding floor-plan surfaces between adjacent slabs to ensure geometric consistency. Each room volume serves as a spatial filter to isolate its corresponding point cloud through automated cloud-to-mesh distance calculations. A modular pipeline processes all rooms and classes, producing sub-datasets organized by room and object type, facilitating balanced and context-rich data for model training and evaluation. Figure 3 presents the technical architecture of the approach.

To generate the room volumes used for geometric grouping, a custom plugin was developed using the PyRevit framework. The plugin extrudes room surfaces defined in the building floor plans to create closed 3D volumes. The extrusion height is automatically computed from the vertical distance between two adjacent slabs (not just Revit levels). When room surfaces are not already defined in the BIM model, they are automatically generated using Revit’s room placement tool. The resulting room geometries are exported as OBJ meshes, providing the spatial references required to isolate the corresponding portions of the point cloud. The extraction of room-level point cloud subsets was automated using a CloudCompare command-line script organized in a batch-processing workflow. The processing sequence consists of the following steps as demonstrated in Figure 3: (1) Loading the room volume meshes (OBJ) together with the semantic class point clouds (TXT) corresponding to building elements (walls, floors, ceilings, etc.). (2) Computing the cloud-to-mesh distance between each class-specific point cloud and the corresponding room mesh. (3) Filtering points based on the computed distances using predefined thresholds [-10 m (inside) to 0.10 m (outside)] to retain only points belonging to the room volume.

As a result, the proposed workflow enables the automatic extraction of room-level labelled point cloud subsets, producing structured datasets that are directly compatible with existing deep learning frameworks for indoor semantic segmentation.

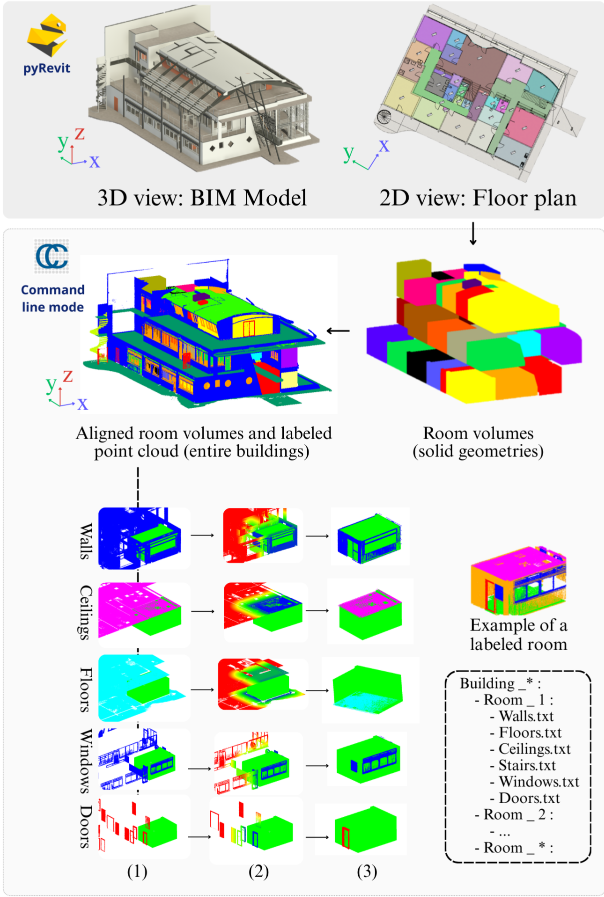

Figure 3. Room-based geometric grouping strategy

4. Experimentation

This section presents the experimental setup and the evaluation of the proposed methodology. First, the datasets used to test the developed approaches are described. Then, the results obtained at each stage of the processing pipeline are presented. Finally, both quantitative and qualitative evaluations are conducted to assess the performance and reliability of the proposed methods.

4.1 Description of the experimental datasets

The experiments were carried out on a diverse dataset including educational, office, residential, and single-family buildings (BIM models and corresponding point clouds), comprising around twenty buildings in total. These building types typically present regular and repetitive structural layouts, making them suitable for analysis under the Manhattan World assumption, according to which most structures are aligned with three orthogonal axes (X, Y, Z). The corresponding digital building models were created at Level of Development 300 (LOD300-detailed design), ensuring sufficient geometric detail for reliable correspondence between BIM elements and point cloud data.

The point clouds were acquired using terrestrial laser scanning through different acquisition modes, including static and mobile scanning, and in some cases a combination of both for indoor–outdoor datasets. The data were spatially sampled at 1 cm resolution, representing a compromise between data size and geometric fidelity. Although intensity and RGB attributes are available, their quality is often degraded and not radiometrically corrected. In particular, intensity values are strongly influenced by material properties, distance, and incidence angle, as shown in (Boudarbala et al., 2024), while color information is often of limited quality. Since these raw datasets were originally acquired for manual Scan-to-BIM modelling workflows, radiometric information is generally not critical and is therefore not systematically optimized. Six main architectural classes were considered: walls, floors, ceilings, stairs, doors, and windows. They represent key structural components. These datasets were used to test and evaluate the proposed labelling approaches under realistic building conditions.

4.2 Results and discussion

In this section, results obtained at each stage of the developed methodology are analysed, from the BBX and SG-NN based approaches to the room-based geometric grouping. Additionally, a qualitative / quantitative evaluation has been conducted.

4.2.1 Bounding box based approach: One of the main advantages of this approach lies in its computational efficiency, as it relies on the BBX of BIM objects rather than their full complex geometries. These bounding volumes, together with the point clouds, can be efficiently processed using the Python library Open3D, which is well suited for geometric operations on large 3D datasets. In most building projects, architectural elements are typically aligned with the XY axes of the project coordinate system to facilitate interpretation in design deliverables. However, this assumption does not always hold. Certain elements may deviate from this orientation or exhibit non-orthogonal geometries. In such cases, the used axis-aligned bounding box may not tightly fit the actual object shape, which can introduce labelling inaccuracies despite the additional geometric constraints applied to refine the results. Such constraints can partially correct simple geometries, as illustrated in Figure 4. Nevertheless, their effectiveness remains limited when dealing with more complex shapes, such as circular, triangular, or irregular objects.

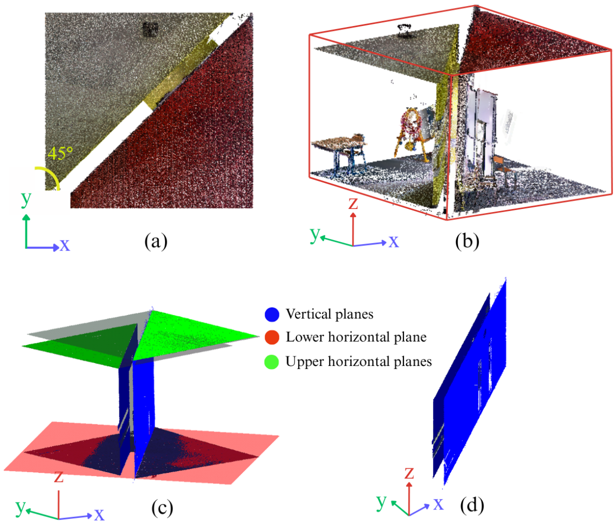

Figure 4. Misalignment between object orientation and axis-aligned BBX. (a) Top view of a point cloud from a wall BBX oriented 45° to the Y-axis. (b) Axis-aligned BBX of the wall. (c) Principal planes extracted using RANSAC. (d) Two dominant vertical planes of the BBX.

As illustrated in the previous example, a significant part of the limitation arises from the use of axis-aligned bounding boxes, which do not necessarily follow the actual orientation of the objects. As a result, the bounding volume may include points that do not belong to the target element, especially for objects that are not aligned with the global coordinate axes. In future work, it would therefore be beneficial to explore alternative bounding representations that better conform to object orientation and geometry, such as oriented bounding boxes (OBB) or convex hull. Such approaches could preserve the computational efficiency of BBX-based methods while reducing labelling inaccuracies.

4.2.2 Solid geometry based approach: The first finding from applying this approach to our datasets is that the nearest-object-based method is relatively slower than the BBX-based approach. This difference in performance stems from the fact that the method relies on computing distances between the exact solid geometry of BIM objects and the point cloud using mesh–cloud distance calculations. As the geometric complexity of objects increases, such as for stairs, windows, or doors, the number of mesh faces involved in the computation also increases, leading to significantly longer processing times. In addition, each semantic class must be processed sequentially against its corresponding object geometry, which further contributes to the overall computational load of the method.

Conversely, because this approach uses the exact solid geometry of BIM objects instead of simplified geometric approximations such as bounding boxes, it provides a more precise spatial correspondence between the model elements and the scanned points. This allows points to be assigned to their respective objects with fewer geometric ambiguities, particularly for elements with complex shapes or orientations. As a result, the approach is expected to produce more accurate labelling results, as confirmed by the quantitative evaluation presented in subsection 4.2.4.

4.2.3 Room-based geometric grouping: By applying one of the previously developed labelling approaches, an annotated point cloud of the entire building can be obtained with the desired number of semantic classes. In our experiments, six classes were considered, yet the approach is not restricted to this configuration and can accommodate additional classes if required. Since the main objective of this work is to facilitate the preparation of datasets for training existing deep learning models for semantic segmentation, it is necessary to organize labelled point clouds into spatial units corresponding to individual rooms. Most existing segmentation datasets for indoor environments follow this principle, as it provides spatial context and manageable data subsets for model training.

The developed approach proved effective for generating such subsets. Nevertheless, several methodological aspects must be considered to ensure the robustness of the process. Room volumes are generated by extruding floor-plan surfaces defined by interior wall boundaries, with heights computed individually from each room’s floor-to-ceiling slabs, allowing variations within the same storey to be captured. This strategy allows the method to handle situations where multiple slabs or ceiling elements exist within the same room, ensuring that the dominant spatial volume of the room is correctly represented. Third, during the point filtering stage, a lower distance threshold of -10 m is applied in order to retain all points located inside the room volume, even when small geometric misalignments exist between the point cloud and the BIM model. An upper threshold of $+0.10$ m is introduced to include points located slightly outside the theoretical room boundary, particularly at interfaces such as slabs where larger geometric discrepancies may occur. Although this tolerance may occasionally include points belonging to adjacent rooms (for instance along shared walls), it's supposed that it does not negatively affect the deep model learning, as those points remain associated with their correct semantic class.

4.2.4 Assessment of the proposed methods: The effectiveness of the proposed labelling approaches was evaluated through qualitative and quantitative analyses:

Qualitative assessment - The qualitative inspection relies on visual inspection of the results obtained across several datasets described in Section 4.1, covering different building typologies, acquisition modes, and spatial scales, providing insight into the robustness of the methods under realistic conditions. Coherent segmentation was shown for both approaches, particularly for (system-family) elements such as walls, floors, ceilings, and stairs. This can be explained by the simple and regular geometry of these elements in both the BIM model and the point cloud.

Nevertheless, both methods show limitations when dealing with (loadable families) such as doors and windows, whose parametric representations in BIM models may not accurately reflect their actual geometry in the point cloud. Consequently, certain inaccuracies may occur during the labelling process. For example, only closed instances of these elements are reliably labelled, while partially or fully open doors or windows remain difficult to accurately segment, since BIM models typically represent them in a closed state. This issue is illustrated in Figure 5 for the SG-NN based approach case. In the BBX-based approach, this issue can be further amplified by the use of axis-aligned BBX, which may introduce additional geometric misalignment, as observed in Figure 4. Another difficulty concerns glass surfaces, especially for windows. Terrestrial laser scanning captures very few points on transparent materials, often limited to the window frames, which further complicates their identification and contributes to lower detection accuracy for this class. These qualitative findings are further substantiated through a quantitative evaluation presented in the following section.

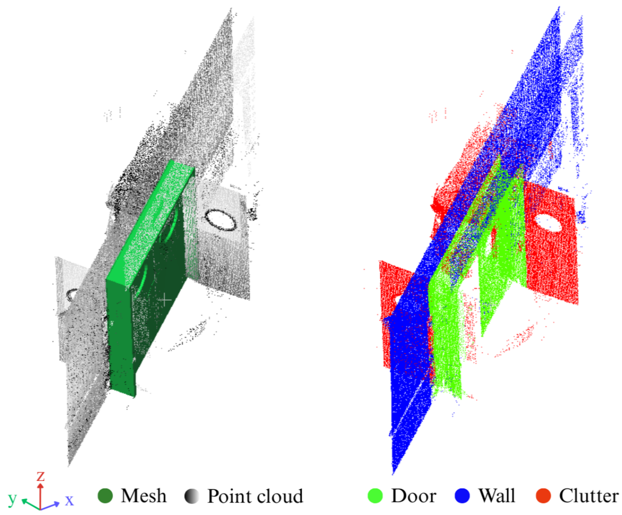

Figure 5. Example of labelling limitations for loadable BIM families (door case)

Quantitative assessment - The quantitative evaluation, was performed on a smaller but controlled dataset corresponding to a single building floor containing 30 rooms manually segmented into 7 classes (walls, floors, ceilings, stairs, windows, doors and clutter). This limitation is due to the need for reliable ground-truth data obtained through manual segmentation of the point cloud, which is considered error-free and serves as the reference dataset for evaluating the two methods. Indeed, this manual segmentation process required approximately 15 minutes per room, amounting to more than 8 hours of work for the 30 rooms and covering about 40% of the total building area. This clearly highlights the time-consuming nature of manual point cloud annotation.

Based on the manually segmented point cloud used as ground truth, a metric-based evaluation of the labelling results produced by the two proposed approaches was conducted, following common practices in semantic segmentation assessment (Betsas et al., 2025). The evaluation relies on the standard confusion matrix components, including true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN), which quantify the correspondence between predicted labels and the reference data.

- TP: reference points that are correctly labelled.
- FN: reference points that are not detected by the labelling.
- FP: labelled points that do not correspond to the ground truth.
- TN: points correctly identified as not belonging to the considered class.

From these elements, several evaluation metrics can be derived, such as overall accuracy (OAcc) , precision, recall, and Intersection over Union (IoU).

$$OAcc = \frac{TP+TN}{TP+TN+FP+FN} \qquad Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN} \qquad IoU = \frac{TP}{TP+FP+FN}$$

The dataset exhibits a strong class imbalance. Walls, floors, and ceilings represent more than 70% of the points, while the clutter class accounts for about 15%, leaving less than 15% for smaller elements such as windows, doors, and stairs. As a result, metrics such as accuracy and precision may lead to biased interpretations. For this reason, the analysis focuses on Recall and IoU, which provide a more reliable evaluation of detection capability and spatial overlap between predicted and reference labels.

	SG-NN		**BBX**	
	Recall	**IoU**	**Recall**	**IoU**
Walls	97	95	95	91
Floors	98	96	87	86
Ceilings	99	97	98	96
Stairs	90	85	86	71
Windows	81	72	76	61
Doors	89	79	79	65
Clutter	92	91	89	87
Average	**92**	**88**	**87**	**78**

Table 1. Comparison of performance between the solid geometry and nearest-neighbour (SG-NN) based approach and the bounding box (BBX) approach.

The results reported in Table 1 show that the SG-NN based approach consistently outperforms the BBX approach. For the main structural classes (walls, floors, and ceilings), both approaches achieve very high performance, with Recall and IoU values above 90%. However, the SG-NN approach still provides slightly higher scores, particularly for floors where the Recall increases from 87% to 98%.

The difference becomes more significant for smaller and geometrically more complex elements such as stairs, windows, and doors. For these classes, the SG-NN approach provides clear improvements in both Recall and IoU, with the most notable gains observed for windows (IoU from 61% to 72%) and doors (IoU from 65% to 79%). These improvements indicate a better spatial correspondence between the predicted labels and the reference segmentation when using the exact object geometries rather than bounding-box approximations. However, the overall scores for these classes remain lower due to the factors discussed in the qualitative analysis, including the higher geometric complexity of these objects, potential misalignment between point clouds and loadable BIM families, the limitations of axis-aligned bounding boxes, and the scarcity of points on transparent glass surfaces.

Overall, the SG-NN method achieves higher average scores, with a mean Recall of 92% and a mIoU of 88%, compared to 87% and 78% for the BBX approach. This confirms the greater robustness of the solid-geometry-based method, particularly for architectural elements with more complex shapes. Although the average IoU does not reach the very high values observed for major structural classes, it remains a strong result given that it is affected by the lower performance of more challenging categories such as doors, windows, and stairs, whose geometric complexity and modelling discrepancies make accurate labelling more difficult.

The overall processing chain of the proposed framework, including the labelling stage based on solid geometry and the subsequent geometric grouping, is illustrated in Figure 6. Depending on the application requirements, the two labelling strategies can be used in a complementary manner: the BBX approach provides faster processing and is suitable for large-scale datasets, whereas the SG-NN approach offers higher geometric precision when more accurate annotations are required, which is generally the case when preparing datasets for deep learning based semantic segmentation.

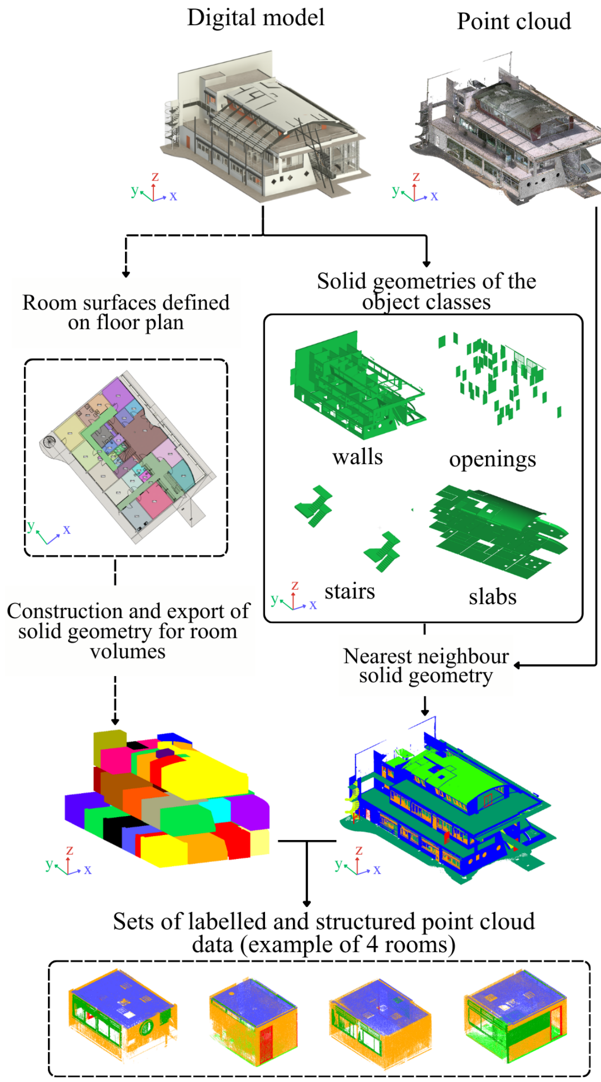

Figure 6. Proposed processing chain for labelling (by solid geometry approach) and room-based geometric grouping of building point cloud datasets

5. Conclusion

This study presented an automated framework for labelling real building-scale point clouds using their corresponding BIM models, addressing one of the major bottlenecks in preparing training datasets for existing deep learning models of semantic segmentation. Two approaches were developed: a computationally efficient method based on bounding boxes (BBX) and a more precise method relying on solid geometries and a nearest-neighbour principle (SG-NN).

Experimental results show that both approaches can produce reliable semantic annotations. The BBX approach offers high computational efficiency and is well suited for large-scale datasets, although its reliance on axis-aligned BBX may introduce geometric approximations. In contrast, the SG-NN approach provides more accurate labelling by using the exact geometry of BIM objects, achieving higher Recall and IoU scores, particularly for geometrically complex elements such as doors and windows. In addition, a room-based geometric grouping strategy was proposed to organize labelled point clouds into spatially coherent subsets. This step is essential for preparing datasets compatible with existing deep learning architectures for indoor semantic segmentation.

Overall, the framework is fully generalizable and applicable to any point cloud associated with a digital model, enabling large-scale, automated dataset generation without manual intervention, i.e. without inconsistencies of human interpretation. Although the BBX approach remains attractive for rapid large-scale processing, the SG-NN method consistently achieves higher accuracy and more reliable spatial correspondence, making it better suited for high-quality dataset generation.

6. Acknowledgements

The authors gratefully acknowledge Futurmap for providing the datasets used in this study and for funding this study. We also thank all contributors who supported the development and execution of this work.

References

Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., Savarese, S., 2016. 3d semantic parsing of large-scale indoor spaces. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1534–1543.

Bai, X., Luo, Y., Jiang, L., Gupta, A., Kaveti, P., Singh, H., Ostadabbas, S., 2023. Bridging the Domain Gap between Synthetic and Real-World Data for Autonomous Driving. *arXiv preprint arXiv:2306.02631*.

Betsas, T., Georgopoulos, A., Doulamis, A., Grussenmeyer, P., 2025. Deep learning on 3d semantic segmentation: A detailed review. *Remote Sensing*, 17(2), 298.

Birkeland, A. A., Udnæs, M., 2024. Semi-automated dataset creation for semantic and instance segmentation of industrial point clouds. *Computers in Industry*, 155, 104064.

Boudarbala, S., Hajji, R., Gourguechon, C., Macher, H., Landes, T., 2024. Automatic detection of doors and windows using point clouds in the context of as-built BIM. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 37–44.

Campagnolo, D., Camuffo, E., Michieli, U., Borin, P., Milani, S., Giordano, A., 2023. Fully automated scan-to-bim via point cloud instance segmentation. *2023 IEEE International Conference on Image Processing (ICIP)*, IEEE, 291–295.

Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T., Nießner, M., 2017. Scannet: Richly-annotated 3d reconstructions of indoor scenes. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5828–5839.

De Geyter, S., Bassier, M., Vergauwen, M., Shults, R., Dlesk, A., 2022. Automated training data creation for semantic segmentation of 3D point clouds. *Measurement, Visualisation and Processing in BIM for Design and Construction Management II*, 46(W1), 59–67.

Hu, D., Gan, V. J., Zhai, R., 2024. Automated BIM-to-scan point cloud semantic segmentation using a domain adaptation network with hybrid attention and whitening (DawNet). *Automation in Construction*, 164, 105473.

Huang, H., Tang, S., Wang, W., Li, X., Guo, R., 2022. From bim to pointcloud: Automatic generation of labeled indoor pointcloud. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 73–78.

Ma, J. W., Czerniawski, T., Leite, F., 2020. Semantic segmentation of point clouds of building interiors with deep learning: Augmenting training datasets with synthetic BIM-based point clouds. *Automation in construction*, 113, 103144.

Mérizette, M., Audebert, N., Kervella, P., Verdun, J., 2025. 3DSES: an indoor Lidar point cloud segmentation dataset with real and pseudo-labels from a 3D model. *arXiv preprint arXiv:2501.17534*.

Roman, O., Bassier, M., De Geyter, S., De Winter, H., Farella, E. M., Remondino, F. et al., 2024. BIM Module for Deep Learning-driven parametric IFC reconstruction. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 403–410.

Tang, S., Li, X., Zheng, X., Wu, B., Wang, W., Zhang, Y., 2022. BIM generation from 3D point clouds by combining 3D deep learning and improved morphological approach. *Automation in Construction*, 141, 104422.

Wu, B., Zhou, X., Zhao, S., Yue, X., Keutzer, K., 2019. Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud. *2019 international conference on robotics and automation (ICRA)*, IEEE, 4376–4382.

Xiao, A., Huang, J., Guan, D., Zhan, F., Lu, S., 2021. Transfer Learning from Synthetic to Real LiDAR Point Cloud for Semantic Segmentation. *arXiv preprint arXiv:2107.05399*.

Yoo, H., Kim, Y., Ryu, J.-H., Lee, S., Lee, J. H., 2024. Semi-Automated Building Dataset Creation for 3D Semantic Segmentation of Point Clouds. *Electronics*, 14(1), 108.

Zhang, H. X., Zou, Z., 2023. Quality assurance for building components through point cloud segmentation leveraging synthetic data. *Automation in Construction*, 155, 105045.