Exploiting Point Cloud Semantic Segmentation for Automatic Geometric Measurement in Modular Construction

Hongxu Chen[1], **Tingtian Li**[1], **Jie Hong**[1], **Xiao Li**[1]

[1]CI^3 Lab, Department of Civil Engineering, The University of Hong Kong, Hong Kong SAR, China

hongxu.chen@connect.hku.hk, tingtian@hku.hk, jiehong@hku.hk, shell.x.li@hku.hk,

Abstract -

Geometric measurement is critical for Modular Construction (MC) after fabrication because it helps prevent dimensional mismatches and safety risks that might otherwise be discovered only during on-site assembly. Terrestrial laser scanning (TLS) provides dense 3D measurements, but automated measurement in cluttered MC scenes remains challenging without robust semantic understanding and geometric reasoning. This paper presents an automated TLS point-cloud-based measurement framework for MC, consisting of data collection, deep-learning semantic segmentation, and geometry-driven post-processing. We evaluate PointNet, PointNet++, and Point Transformer V3 under transfer-learning and training-from-scratch settings. Using the segmented point clouds, we estimate room dimensions, assess wall verticality and orthogonality, estimate window dimensions, and evaluate wall flatness using RMS and point-to-plane distance metrics. Experiments on representative MC rooms demonstrate that the proposed pipeline achieves practical measurement accuracy and produces full-coverage deviation maps to locate nonconforming regions.

Keywords -

Modular Construction; Geometric Measurement; Semantic Segmentation in Point Clouds

1 Introduction

Modular Construction (MC) is an innovative and efficient building paradigm in which fully finished or semi-finished functional modules, including structural elements, interior fittings and utility systems, are prefabricated in factories, transported, and assembled on site [1]. However, integrating substandard MC modules can introduce quality and safety risks, potentially leading to demolition, repair, or replacement, which in turn delays construction schedules and increases costs. Therefore, geometric measurement is essential in MC projects.

Typical geometric tasks include room-dimension measurements, verification of the dimensions and placement of doors and windows, and assessment of wall surface flatness.

Conventional geometric measurement in the construction industry largely relies on manual measurements using tools such as measuring tapes and levels. These procedures are time-consuming and highly dependent on workers' experience and skill levels [2]. To mitigate these limitations, emerging sensing technologies, including computer vision [3, 4], laser scanning [5, 6], and infrared measurement technologies [7, 8] have been increasingly explored for construction geometric measurement.

Image-based approaches are easy and cost-effective to deploy and enable rapid data collection. However, due to inherent limitations in imaging and depth recovery, they often struggle to reach the precision required for fine-grained geometric inspection [9]. In contrast, terrestrial laser scanning (TLS) provides accurate 3D spatial information and has been widely used for measuring building elements such as doors and windows [10].

Early studies demonstrated the feasibility of TLS for automated dimensional inspection of prefabricated components. Kim et al. [11, 12] developed TLS-based pipelines to measure the dimensions, positions, and squareness of precast concrete elements by combining geometric feature extraction, error compensation, and BIM-based deviation analysis, achieving millimeter-level accuracy under controlled conditions. To extend TLS-based inspection to indoor environments, Tan et al. [13] proposed a geometry-driven approach to evaluate wall flatness, verticality, and door/window opening dimensions. Despite their accuracy, such methods typically rely on strong geometric assumptions and relatively clean scenes with limited occlusion, which restricts their applicability to complex and cluttered MC environments.

To improve robustness in unstructured scenes, recent studies have explored learning-based semantic understanding and component detection in point clouds. Tang et al. [14] integrated 3D deep learning with morphological processing to reconstruct semantically enriched BIM models from LiDAR and RGB-D data. Boudarbala et al. [15] proposed an automated workflow for door/window detection by jointly exploiting geometric and radiometric cues. Zhang [16] further addressed limited labeled data by training RandLA-Net on BIM-generated synthetic point clouds with data augmentation; real scans were aligned to the

BIM model via manual selection of feature points, after which the trained network was used to segment on-site point clouds and quantify dimensional deviations from the BIM design using cloud-to-cloud distance metrics. Despite improved automation and semantic understanding, these methods often rely on manual tuning or human intervention (e.g., achieving high-accuracy alignment between BIM-derived and as-built point clouds), and may provide limited accuracy for high-precision dimensional measurement.

Beyond dimensional measurement, TLS has also been applied to surface-related indicators. Wei et al. [17] quantified surface flatness of tunnel initial supports using deviation-based metrics on TLS data. Li et al. [18] developed a full-coverage TLS-based flatness inspection method for large-scale industrial floors by virtualizing the traditional 2-m ruler and feeler-gauge procedure on dense point clouds. However, existing TLS-based surface flatness methods are often tailored to relatively structured and simplified scenarios, making them difficult to extend directly to component-level surface inspection in cluttered, diverse, and heterogeneous environments.

In summary, existing workflows for building inspection can be broadly categorized into purely geometry-driven methods and learning-based point-cloud segmentation methods. Geometry-driven approaches can be accurate under controlled conditions but rely on strong assumptions and are best suited to isolated components in structured scenes. Learning-based methods improve automation, yet in complex MC environments they may suffer from degraded segmentation and limited accuracy for fine-grained geometric measurement. Moreover, many pipelines still require substantial manual tuning and support only a limited set of component types.

To address these limitations, this paper proposes an automated geometric measurement framework based on TLS point clouds for MC applications. We adopt a more efficient point-cloud semantic segmentation model to improve component-level recognition accuracy in MC scenes, and develop more geometry-aware post-processing algorithms to enable reliable room-scale dimensional measurement. The main contributions of this paper are summarized as follows:

- A real-world TLS dataset of MC modules with diverse structural configurations is collected, providing practical data support for geometric measurement research in complex MC environments.

- An automated geometric measurement framework for MC is proposed based on TLS point clouds that integrates semantic segmentation with geometry-aware post-processing to enable robust component-level geometric measurement and analysis.

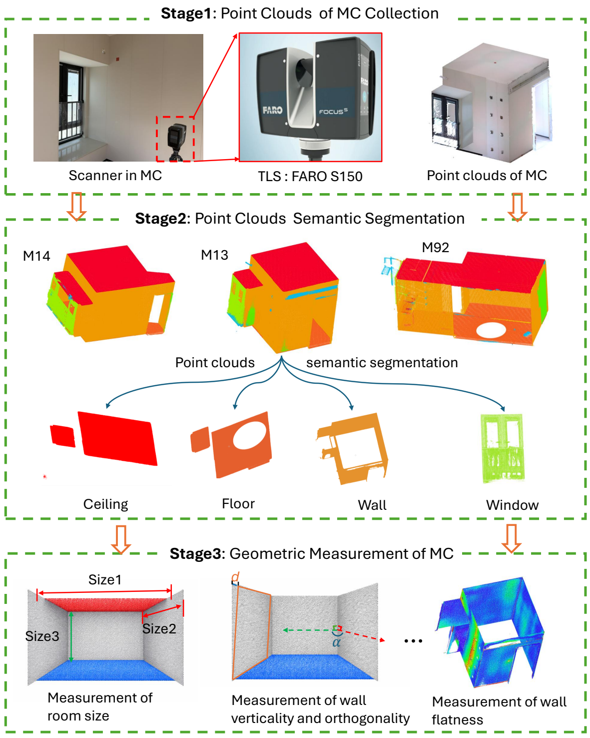


Figure 1. Proposed MC geometric measurement pipeline. TLS-acquired point clouds are first segmented by a semantic segmentation model, and geometric measurements are then performed for each predicted class.

2 Methodology

This paper proposes an automated geometric measurement workflow tailored to MC scenarios, as illustrated in Fig. 1. The framework consists of three stages. First, TLS point clouds of MC modules are acquired and organized. Second, a neural-network-based semantic segmentation model is used to identify key building components in the point clouds. Third, geometry-aware analysis is applied to the segmented components to compute dimensional and surface-flatness metrics, enabling fine-grained inspection in complex indoor environments.

2.1 Semantic Segmentation

To automatically identify MC components in raw scanner point clouds, we employ neural network-based semantic segmentation. Given the variety of MC elements and the complexity of indoor scenes, learning-based segmentation offers a practical alternative to purely geometry-driven pipelines. We evaluate three representative point cloud segmentation models on the proposed MC dataset: PointNet [19], PointNet++ [20], and the recently proposed

Point Transformer V3 [21]. These models cover different design paradigms, including point-wise global feature learning (PointNet), hierarchical local feature aggregation (PointNet++), and transformer-based representation learning for large-scale point clouds (Point Transformer V3).

Each point is represented using scanner-provided attributes, including 3D coordinates (xyz), RGB color, and surface normals (nx, ny, nz). These features are concatenated as the network input to capture both geometric cues and appearance-related information of MC components.

To evaluate learning-based semantic segmentation on MC point clouds, we consider two training strategies: (1) transfer learning from a large-scale indoor dataset followed by staged fine-tuning on MC data, and (2) training from scratch using only the collected MC dataset. All models are evaluated using the same train–test split.

Pretraining on S3DIS and two-stage fine-tuning. Given the limited amount of annotated MC point cloud data and the similarity between MC scenes and indoor environments, we initialize the networks with S3DIS-pretrained weights. The segmentation head is re-initialized to predict the MC classes. Fine-tuning on the MC dataset is performed in two stages to stabilize optimization:

- **Stage I (head-only training):** the backbone is frozen and only the segmentation head is trained on the MC dataset to adapt to the new class taxonomy.

- **Stage II (full fine-tuning):** the backbone is unfrozen, and the entire network is further fine-tuned on the MC dataset to better fit MC-specific geometry and component distributions.

Training from scratch on MC data. In addition to transfer learning, we train each model from scratch using only the MC dataset as an alternative training strategy. In this setting, all network parameters are randomly initialized and optimized end-to-end on MC data under the same train–test split and evaluation protocol.

Segmentation performance is evaluated using Intersection over Union (IoU), defined as the ratio of the intersection between the predicted segmentation and the ground-truth region to their union. A higher IoU indicates more accurate and complete segmentation of the MC components, thereby reducing subsequent measurement errors.

2.2 Geometric Measurement

This subsection presents geometric measurement modules operating on segmented MC point clouds, including (i) room size measurement, (ii) ceiling/floor elevation range evaluation, (iii) wall flatness evaluation, (iv) verticality/orthogonality assessment, and (v) window size measurement.

2.2.1 Room Size

Given a semantically segmented point cloud $\mathcal{P} = \{(\mathbf{x}_i, \ell_i)\}_{i=1}^{N}$, we estimate the room dimensions (length, width, and height) using the ceiling, floor, and wall regions, which provide stable planar references. The corresponding point subsets are $\mathcal{P}^{\text{ceil}} = \{\mathbf{x}_i \mid \ell_i = \ell_{\text{ceil}}\}$, $\mathcal{P}^{\text{floor}} = \{\mathbf{x}_i \mid \ell_i = \ell_{\text{floor}}\}$, and $\mathcal{P}^{\text{wall}} = \{\mathbf{x}_i \mid \ell_i = \ell_{\text{wall}}\}$.

For each structural subset, isolated points are first removed using a radius-based filter. Plane candidates are then extracted by RANSAC, and DBSCAN is subsequently applied to separate spatially disconnected but geometrically coplanar fragments. For each candidate plane, only the largest connected component is retained, and its parameters are finally refined by least-squares fitting (SVD) to obtain a unit normal $\mathbf{n}$ and centroid $\mathbf{c}$.

For the ceiling and floor subsets, the dominant plane is selected as the coherent patch with the largest number of inliers, yielding $(\mathbf{n}^{\text{ceil}}, \mathbf{c}^{\text{ceil}})$ and $(\mathbf{n}^{\text{floor}}, \mathbf{c}^{\text{floor}})$, respectively. For the wall subset, an iterative RANSAC procedure is used to extract multiple planar candidates. After connected-component filtering and SVD refinement, the minimum distance from each wall plane to the floor plane is computed. Only planes satisfying $h_{\min} \le t_{\text{touch}}$ are retained as valid wall candidates. These candidates are then ranked by the number of inliers, and the top four dominant walls are selected for subsequent analysis.

Finally, the room height is computed as the average perpendicular distance from the points in the dominant ceiling patch to the floor plane. The room length and width are estimated as the distances between two pairs of parallel walls among the four selected dominant walls.

2.2.2 Ceiling/Floor Elevation Range

Based on the dominant ceiling/floor plane estimated in Section 2.2.1, we compute a corner-level elevation range using four local patches. Let the plane be $\pi : \mathbf{n}^\top \mathbf{x} + d = 0$ with $\|\mathbf{n}\|_2 = 1$, and let $\mathbf{x}_{\text{corner}}^{(i)}$ $(i = 1, \ldots, 4)$ denote the four 3D corners of the fitted room rectangle. For each corner, we define a square patch (with side length s) whose center is given by

$$\mathbf{x}_{\text{center}}^{(i)} = \mathbf{x}_{\text{corner}}^{(i)} + o\,\widehat{\mathbf{t}}^{(i)}, \qquad \widehat{\mathbf{t}}^{(i)} = \frac{(\mathbf{I} - \mathbf{n}\mathbf{n}^\top)(\mathbf{c} - \mathbf{x}_{\text{corner}}^{(i)})}{\|(\mathbf{I} - \mathbf{n}\mathbf{n}^\top)(\mathbf{c} - \mathbf{x}_{\text{corner}}^{(i)})\|_2} \tag{1}$$

where $\mathbf{c}$ is the rectangle center and o is an inward offset along the plane.

Using an orthonormal in-plane basis $(\mathbf{u}, \mathbf{v})$, we collect points $\mathcal{X}^{(i)}$ within each $s \times s$ patch centered at $\mathbf{x}_{\text{center}}^{(i)}$ to obtain a robust local estimate. The average point of each patch is computed and its signed distance to the plane is evaluated as $\delta(\mathbf{x}) = \mathbf{n}^\top \mathbf{x} + d$. Finally, the elevation range is defined as the difference between the maximum and minimum average distances among the four corner patches.

2.2.3 Wall Flatness

We quantify wall surface flatness using a patch-level root-mean-square (RMS) deviation as a global indicator, complemented by a dense per-point distance-to-plane map $\{D_i\}$ to capture localized deviations.

Given a contiguous planar patch extracted in Section 2.2.1, denoted as $\mathcal{S} = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$, we fit a plane via SVD to obtain the centroid $\mathbf{c}$ and unit normal $\mathbf{n}$ ($\|\mathbf{n}\|_2 = 1$). The absolute deviation of each point from the fitted plane is defined as

$$D_i = \left|\mathbf{n}^\top (\mathbf{x}_i - \mathbf{c})\right|, \qquad \mathbf{x}_i \in \mathcal{S} \tag{2}$$

and the patch-level RMS deviation is computed as

$$\text{RMS} = \sqrt{\frac{1}{M}\sum_{i=1}^{M} D_i^2} \tag{3}$$

The set of values $\{D_i\}$ forms a dense deviation map that can be used for visualization and analysis of local surface irregularities.

2.2.4 Wall Verticality and Orthogonality

We evaluate wall geometry quality from two aspects: (i) wall-to-floor verticality and (ii) orthogonality between adjacent walls. The dominant floor plane $(\mathbf{n}_f, \mathbf{c}_f)$ is estimated from $\mathcal{P}^{\text{floor}}$ (Section 2.2.1), where $\|\mathbf{n}_f\|_2 = 1$. Four dominant wall patches $\{\mathcal{W}_k\}$ are extracted using the same decomposition procedure as described in Section 2.2.1, and each patch is refitted via SVD to obtain a unit normal $\mathbf{n}_k$ and centroid $\mathbf{c}_k$.

Wall-to-floor verticality. For each floor-connected wall patch $\mathcal{W}_k$, we define the height of a point $\mathbf{x} \in \mathcal{W}_k$ along the floor normal as $h(\mathbf{x}) = (\mathbf{x} - \mathbf{c}_f)^\top \mathbf{n}_f$. The verticality deviation is quantified by the angular deviation between the wall normal and the floor normal:

$$\theta_k = \arcsin\left(\left|\mathbf{n}_k^\top \mathbf{n}_f\right|\right) \tag{4}$$

To provide an interpretable geometric measure, we further approximate the lateral deviation along the wall height as

$$d_k \approx \theta_k \,(h_{\max} - h_{\min}), \tag{5}$$

where $h_{\min} = \min_{\mathbf{x}\in\mathcal{W}_k} h(\mathbf{x})$ and $h_{\max} = \max_{\mathbf{x}\in\mathcal{W}_k} h(\mathbf{x})$

Wall-to-wall orthogonality. Adjacency between two wall patches $\mathcal{W}_i$ and $\mathcal{W}_j$ is determined by the minimum Euclidean distance between their inlier point sets; a pair is considered adjacent if $d_{ij}^{\min} \leq t_{\text{conn}}$. For each adjacent pair, the orthogonality error is defined based on their unit normals as

$$e_{ij} = \left|\arccos\left(\left|\mathbf{n}_i^\top \mathbf{n}_j\right|\right) - \frac{\pi}{2}\right| \tag{6}$$

2.2.5 Window Size

We first extract the window points $\mathcal{P}^{\text{win}} = \{\mathbf{x}_i \mid \ell_i = \ell_{\text{window}}\}$ and apply DBSCAN to remove noise, obtaining a denoised point set $\mathcal{P}^{\text{win}}_{\text{db}}$. A planar model is then fitted to $\mathcal{P}^{\text{win}}_{\text{db}}$ using RANSAC, and the largest inlier set is retained as $\widetilde{\mathcal{P}}^{\text{win}}$. Next, the normal of the window plane is estimated from $\widetilde{\mathcal{P}}^{\text{win}}$ via PCA. Among the wall planes extracted in Section 2.2.1, the one whose normal is most parallel to the window-plane normal is selected as the supporting wall plane. The points in $\widetilde{\mathcal{P}}^{\text{win}}$ are orthogonally projected onto this wall plane and then mapped to a 2D local coordinate system, yielding the 2D point set C^{2D}.

To reduce the influence of boundary artifacts, we iteratively trim the point set along both axes by removing boundary points whose distances to their nearest interior neighbors exceed a threshold τ.

Finally, we compute the 2D convex hull of the trimmed points and fit a minimum-area bounding rectangle using OpenCV `minAreaRect`. The window width and height are defined as the side lengths of this rectangle.

3 Experiments and Results

3.1 Data Collection and Preparation

The data used in this study were collected at an MC manufacturing base in Zhuhai, China. Terrestrial laser scanning was conducted using a FARO Focus S150 to acquire as-built point clouds of MC modules with different structural types, including both steel-structured and concrete-structured modules. According to the device specifications, the ranging accuracy is better than 0.7 mm within 10 m. Under the adopted scanning configuration, the point density is approximately 87,507 pts/m^2 at a distance of about 1 m from the scanner, corresponding to an average point spacing of roughly 3 mm on surfaces perpendicular to the laser beam direction. In total, 44 point cloud files were collected and organized for subsequent semantic segmentation and geometric measurement experiments.

For semantic segmentation, all point cloud files were manually annotated. The labeling scheme follows the S3DIS class definitions and is extended to better reflect the structural characteristics and functional components specific to MC modules. As a result, the annotation taxonomy was expanded to a total of 22 semantic classes: ceiling, floor, wall, beam, column, window, door, table, chair, sofa, bookcase, board, clutter, electronic hole, electronic box, electronic pipe, water pipe, basin, toilet, ceramic tile, rebar, and lamp. Here, the electronic-related classes refer to openings, boxes, and conduits associated with electrical installations commonly found in MC modules.

From the 44 collected point cloud files, three representative scenes were selected for testing, and the remaining

Table 1. Semantic segmentation results on the MC TLS test set.

Model	Training data	$\mathrm{mIoU}_{\mathrm{rep}}$	Ceiling	Floor	Wall	Window	Door
PointNet	S3DIS+MC	0.69	0.93	0.93	0.86	0.59	0.12
PointNet++	S3DIS+MC	0.73	0.93	0.94	0.90	0.69	0.20
PT V3	S3DIS+MC	0.72	0.90	0.93	0.90	0.49	0.40
PointNet	MC	0.73	0.93	0.90	0.86	0.57	0.38
PointNet++	MC	0.57	0.83	0.89	0.84	0.29	0
PT V3	MC	**0.89**	**0.94**	**0.95**	**0.93**	**0.89**	**0.74**

files were used for training. This split was adopted because several scans contain relatively simple, repetitive layouts; as a result, a conventional 80/20 split would yield a test set that is not sufficiently representative. The test set was therefore curated to better reflect typical MC environments and to provide a more meaningful evaluation of generalization. The selected test scenes are shown in Fig. 1 (Stage 2), labeled from left to right as M14, M13, and M92. Although the test set is small, each test scene contains multiple components and cluttered structures commonly observed in MC modules, enabling component-level geometric measurement under realistic conditions.

3.2 Semantic Segmentation in Point Clouds

We evaluate semantic segmentation on the proposed MC TLS dataset using three representative architectures: PointNet, PointNet++, and Point Transformer V3. The results are summarized in Table 1. For concise presentation, Table 1 only reports the classes whose IoU exceeds 0.5 in our experiments. Classes whose IoU remains below 0.5 across all architectures are omitted because their predictions are insufficiently distinguishable from other categories, which can lead to unreliable inputs and reduced accuracy in subsequent geometric measurement. Accordingly, we report $\mathrm{mIoU}_{\mathrm{rep}}$, defined as the mean IoU over the reported classes.

As shown in Table 1, the highest $\mathrm{mIoU}_{\mathrm{rep}}$ is achieved by Point Transformer V3 trained from scratch (0.89), with consistently high IoU on the dominant structural classes (Ceiling/Floor/Wall) and strong performance on opening-related components such as *Window* and *Door* (0.89/0.74). Under S3DIS initialization, PointNet++ benefits from pretraining, improving $\mathrm{mIoU}_{\mathrm{rep}}$ from 0.57 (scratch) to 0.73 (fine-tuned), suggesting that pretrained hierarchical local features can be helpful when MC supervision is limited.

In contrast, Point Transformer V3 shows degraded performance after S3DIS pretraining, with mIoU dropping from 0.89 (scratch) to 0.72 (fine-tuned). A plausible explanation is the domain gap between S3DIS and our MC scenes. Since Transformer-based models rely heavily on global context, the pretrained representations may be biased toward S3DIS-specific scene priors, causing negative transfer on MC data. Therefore, the subsequent geometric measurement experiments use predictions from Point Transformer V3 trained from scratch.

3.3 Geometric Measurement

Based on the semantic segmentation results of MC point clouds, we develop geometry-aware algorithms to measure room size, ceiling/floor elevation range, wall flatness, wall verticality and orthogonality, and window size. The quantitative results are summarized in Table 2, which compares our measurements with manual measurements from the point clouds, and reports the corresponding errors. Detailed analyses of each measurement item are provided in the following sections.

3.3.1 Room Size

For M13, the measured room length, width, and height are 2346 mm, 2569 mm, and 2769 mm, respectively, compared with the corresponding manual measurements of 2344 mm, 2568 mm, and 2769 mm, yielding absolute errors of 2 mm, 1 mm, and 0 mm. For M14, the estimated length, width, and height are 3116 mm, 3057 mm, and 2767 mm, respectively, corresponding to absolute errors of 2 mm, 2 mm, and 1 mm.

These results demonstrate that the proposed method can reliably recover global room dimensions with millimeter-level precision. Compared with Truong-Hong et al. [22], who reported a 2.9% error in ceiling area measurement, our method achieves a substantially smaller ceiling-area error of 0.1% (assuming the ceiling area is approximated by length × width).

3.3.2 Ceiling/Floor Elevation Range

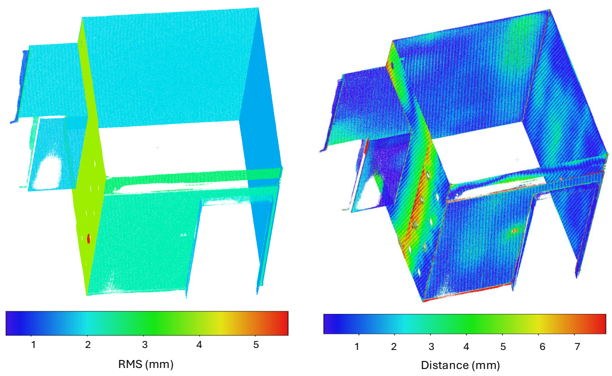
In this study, the inward offset is set to $o = 300$ mm and the patch side length to $s = 10$ mm. As shown in Table 2, the automatically measured ceiling elevation ranges are 2.8 mm for M13 and 2.3 mm for M14, corresponding to absolute errors of 0.3 mm and 0.9 mm relative to manually obtained reference values. For the floor, the measured elevation ranges yield errors of 0.2 mm (M13) and 0.0 mm (M14), demonstrating similarly high measurement accuracy for floor surfaces.

3.3.3 Wall Flatness

Figure 2a presents the RMS flatness of the extracted wall patches in room M13. All major wall patches exhibit an RMS ≤ 4 mm, indicating good overall planarity at the patch level. Among the four walls, the left wall shows the largest RMS value (4.0 mm), primarily due to irregular openings and voids that introduce larger local deviations and consequently inflate the RMS metric. In contrast, the right wall demonstrates the best flatness performance, with

Table 2. Geometric measurement results of major MC components from M13 and M14 point clouds.

Items	M13			M14		
	Measurement	Manual	Error	Measurement	Manual	Error
Length (mm)	2346	2344	2	3116	3114	2
Width (mm)	2569	2568	1	3057	3055	2
Height (mm)	2769	2769	0	2767	2766	1
Ceiling elevation range (mm)	2.8	2.5	0.3	2.3	3.2	0.9
Floor elevation range (mm)	1.3	1.1	0.2	0.9	0.9	0.0
Wall flatness (RMS, mm)	1.6, 1.9 2.3, 4.0	1.7, 1.9 1.8, 3.9	0.1, 0.0 0.5, 0.1	2.6, 2.3 2.1, 3.2	2.4, 1.8 2.2, 4.1	0.2, 0.5 0.1, 0.9
Wall verticality (mm)	0.6, 1.5 1.7, 1.1	0.1, 3.2 0.5, 1.5	0.5, 1.7 1.2, 0.4	5.2, 4.2 1.3, 0.1	5.6, 4.1 1.0, 0.1	0.4, 0.1 0.3, 0.0
Wall orthogonality (°)	0.2, 0.3 0.4, 0.3	0.2, 0.2 0.4, 0.3	0.0, 0.1 0.0, 0.0	0.1, 0.0 0.0	0.1, 0.0 0.0	0.0, 0.0 0.0
Window size (mm)	1730, 1037	1743, 1028	13, 9	1766, 2476	1750, 2469	16, 7



(a) RMS map. (b) Distance-to-plane map.

Figure 2. Wall flatness visualization for the M13 room.

an RMS of only 1.6 mm. When compared to manual measurements, the maximum and minimum absolute errors are 0.5 mm and 0.0 mm, respectively. Furthermore, the proposed method achieves substantially improved measurement accuracy over the work of Li et al. [23], who reported an average flatness measurement error of 7.5 mm for the M4 room in their dataset.

Figure 2b further shows that most wall regions have absolute point-to-plane distances within approximately 1 mm, indicating that the major wall surfaces remain close to an ideal plane. Notably, deviations exceeding 3 mm are mainly concentrated around the opening regions and electrical box areas. Using a practical acceptance criterion of 3 mm for point-to-plane deviation, room M13 is assessed as broadly compliant. Similarly, for room M14, the RMS values range between 2 and 4 mm. Compared to manual measurements, the maximum and minimum absolute errors for this room are 0.9 mm and 0.1 mm, respectively.

Overall, the TLS data and the proposed method enable both quantitative evaluation and intuitive visualization of wall flatness.

3.3.4 Wall Verticality and Orthogonality

In this experiment, we set $t_{\text{touch}} = t_{\text{conn}} = 0.15$ m. For the M13 room, the estimated verticality deviations of the four dominant wall planes are 0.6 mm, 1.5 mm, 1.7 mm, and 1.1 mm, respectively. Compared with the corresponding manual measurements, the absolute errors are 0.5 mm, 1.7 mm, 1.2 mm, and 0.4 mm, yielding a mean absolute error of 0.95 mm. This result is lower than the average verticality measurement error of approximately 1.33 mm reported by Tan et al. [13], demonstrating the higher accuracy of the proposed method. Under the commonly adopted tolerance requirement that the wall verticality deviation should be within 4 mm for walls lower than 3 m, the M13 room satisfies the verticality criterion.

We also assess the orthogonality of adjacent wall planes in room M13. The computed deviations for the four wall pairs are 0.2°, 0.3°, 0.4°, and 0.3°, with corresponding absolute errors of 0.0°, 0.1°, 0.0°, and 0.0° relative to manual measurements. As all values are well below the standard tolerance of 1°, room M13 satisfies the orthogonality requirement with a large margin.

For the M14 room, the estimated verticality deviations are 5.2 mm, 4.2 mm, 1.3 mm, and 0.1 mm, with corresponding absolute errors of 0.4 mm, 0.1 mm, 0.3 mm, and 0.0 mm, respectively. The mean absolute error is 0.2 mm. Although the estimated values are in close agreement with the manual measurements, two walls exceed the 4 mm tolerance limit, indicating that M14 does not fully satisfy the verticality criterion. For wall orthogonality in the M14 room, the computed deviations are 0.1°, 0.0°, and 0.0°, which are consistent with the manual measurements and all well below the tolerance threshold of 1°. Therefore, the M14 room satisfies the orthogonality requirement.

3.3.5 Window Size

Considering the 3 mm point density, we set the threshold $\tau = 3$ mm. For the M13 room, the measured window height and width are 1730 mm and 1037 mm, respectively. Compared with the manual measurements, the absolute errors are 13 mm in height and 9 mm in width. For the M14 room, the measured window height and width are 1766 mm and 2476 mm, respectively, while the corresponding manual measurements are 1750 mm and 2469 mm. The resulting absolute errors are 16 mm in height and 7 mm in width. Despite these deviations, the results still show good agreement with the manual measurements, indicating that the proposed method can provide reliable window-size estimation in different room layouts.

Compared with Tang et al. [14], who reported door/window location deviations exceeding 120 mm in BIM reconstruction, our proposed method achieves substantially higher geometric accuracy in these case studies.

3.3.6 Error Analysis

This subsection analyzes the measurement errors between the proposed automatic geometric measurement method and manual measurements. Since the scanner used in this study is highly accurate, the analysis focuses primarily on other error sources.

The results show that the measurement errors are small for room dimensions and the geometric attributes of ceilings, floors, and walls. This is mainly because these components achieve high semantic segmentation performance, with IoU values all above 0.93, indicating only limited point misclassification. Moreover, these targets are relatively large, and the proposed post-processing method is robust to minor segmentation errors, thereby maintaining high measurement accuracy. The remaining errors may be partly attributed to manual measurement, which is usually performed on selected points or local regions and may therefore be less comprehensive and more sensitive to local irregularities. In contrast, the proposed method evaluates the geometry of the entire component, providing more comprehensive and stable results.

The window measurement errors are relatively larger than those of other components, mainly because of lower segmentation accuracy for windows. It should be noted that this measurement refers to the installed window within the wall opening rather than the outer structural window frame. Consequently, during semantic segmentation, the boundary between the window and the surrounding wall may be ambiguous, causing some wall points to be misclassified as window points, as observed in the length measurement of M13 and the height measurement of M14. In the current window-size measurement procedure, a rectangle is fitted to the segmented window points, making the result sensitive to such boundary misclassification and potentially leading to overestimation or underestimation. Although denoising strategies are applied to mitigate this effect, their influence cannot yet be fully eliminated.

4 Conclusion

This paper presents an automated TLS point-cloud-based geometric measurement framework tailored to MC scenarios. Using FARO-scanned MC point cloud data, we first perform learning-based semantic segmentation with PointNet, PointNet++, and Point Transformer V3 in both transfer-learning and training-from-scratch settings, and then integrate geometry-driven post-processing for quantitative measurement. The proposed framework supports a range of measurement tasks, including room-level dimensional measurement, evaluation of wall verticality and orthogonality, analysis of ceiling/floor elevation range, window size measurement, and wall surface flatness assessment. Experiments demonstrate practical accuracy for room-scale measurements and effective localization of non-conforming regions in flatness maps, enabling tolerance-based decision-making.

Limitations mainly stem from the current dataset scale and the difficulty of segmenting extremely sparse classes. Future work will expand the dataset, improve robustness to long-tailed class distributions and occlusions, and extend the framework to additional MC geometric-measurement items, such as opening squareness, door/window verticality, and geometric measurements of embedded components (e.g., their dimensions and locations), with more comprehensive validation against industrial standards.

5 Acknowledgments

This study is supported by RISUD, PolyU (P0052942), HKU Seed Fund for Collaborative Research (2407102267) and Innovation and Technology Fund (PRP/048/24LI).

References

[1] Yuanyang Qi, Xiao Li, and Ruiqi Jiang. Vision-based segmentation, measurement, and pose estimation in modular integrated construction. In *Proc. 42nd Int. Symp. Autom. Robot. Constr. (ISARC)*, pages 1097–1104, Montreal, Canada, July 2025. Int. Assoc. Autom. Robot. Constr. (IAARC).

[2] Zhiliang Ma, Yu Liu, and Jiayi Li. Review on automated quality inspection of precast concrete components. 150:104828.

[3] Guofeng Ma, Ming Wu, Zhijiang Wu, and Wenjing Yang. Single-shot multibox detector- and building information modeling-based quality inspection model for construction projects. *J. Build. Eng.*, 38:102216, 2021.

[4] Yangze Liang and Zhao Xu. Intelligent inspection of appearance quality for precast concrete components based on improved yolo model and multi-source data.

Eng. Constr. Archit. Manag., 32(3):1691–1714, 10 2023.

[5] Min-Koo Kim, Jack C.P. Cheng, Hoon Sohn, and Chih-Chen Chang. A framework for dimensional and surface quality assessment of precast concrete elements using bim and 3d laser scanning. *Autom. Constr.*, 49:225–238, 2015. 30th ISARC Special Issue.

[6] Hanbin Luo, Jianjiang Zhan, Henry J. Liu, Wei Zhou, and Weili Fang. A novel uncertainty-aware point cloud approach for geometric quality monitoring in construction. *Adv. Eng. Inf.*, 68:103642, 2025.

[7] Eberechi Ichi and Sattar Dorafshan. Effectiveness of infrared thermography for delamination detection in reinforced concrete bridge decks. *Autom. Constr.*, 142:104523, 2022.

[8] Siyu Zheng, Jiaxin Zhang, Rui Zu, and Yunqin Li. Vision transformer-enhanced thermal anomaly detection in building facades through fusion of thermal and visible imagery. *J. Asian Archit. Build. Eng.*, 24 (4):2854–2868, 2025.

[9] Ali Algadhi, Panos Psimoulis, Athina Grizi, and Luis Neves. Experimental assessment of the performance of terrestrial laser scanners in monitoring the geometric deformations in retaining walls. 15(6):1885–1900.

[10] Sander Varbla, Raido Puust, and Artu Ellmann. Quality evaluation of sizeable surveying-industry-produced terrestrial laser scanning point clouds that facilitate building information modeling—a case study of seven point clouds. *Buildings*, 14(11), 2024.

[11] Min-Koo Kim, Hoon Sohn, and Chih-Chen Chang. Automated dimensional quality assessment of precast concrete panels using terrestrial laser scanning. *Autom. Constr.*, 45:163–177, 2014.

[12] Min-Koo Kim, Qian Wang, Joon-Woo Park, Jack C.P. Cheng, Hoon Sohn, and Chih-Chen Chang. Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and bim. *Autom. Constr.*, 72:102–114, 2016.

[13] Yi Tan, Xin Liu, Shuaishuai Jin, Qian Wang, Daochu Wang, and Xiaofeng Xie. A terrestrial laser scanning-based method for indoor geometric quality measurement. *Remote Sens.*, 16(1), 2024.

[14] Shengjun Tang, Xiaoming Li, Xianwei Zheng, Bo Wu, Weixi Wang, and Yunjie Zhang. Bim generation from 3d point clouds by combining 3d deep learning and improved morphological approach. *Autom. Constr.*, 141:104422, 2022.

[15] S. Boudarbala, R. Hajji, C. Gourguechon, H. Macher, and T. Landes. Automatic detection of doors and windows using point clouds in the context of as-built bim. *ISPRS Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, XLVIII-2/W8-2024:37–44, 2024.

[16] Hao Xuan Zhang and Zhengbo Zou. Quality assurance for building components through point cloud segmentation leveraging synthetic data. *Autom. Constr.*, 155:105045, 2023.

[17] Xiao Wei, Jijun Wang, Chunzhong Xiao, Qasim Zaheer, Weidong Wang, Xianhua Liu, Jin Wang, and Shi Qiu. Quantification and evaluation of roughness of initial support using terrestrial laser scanning. *J. Comput. Civ. Eng.*, 39(1):04024052, 2025.

[18] Xianzhe Li, Haiming Li, Ting Jiao, Changzheng Chi, and Mingfeng Huang. Flatness detection of super-flat emery floor based on terrestrial laser scanning. *Vibroeng. Proc.*, 54:186–192, April 2024.

[19] R. Qi Charles, Hao Su, Mo Kaichun, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pages 77–85, 2017.

[20] Charles R. Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: deep hierarchical feature learning on point sets in a metric space. In *Proc. 31st Int. Conf. Neural Inf. Process. Syst. (NIPS)*, NIPS'17, page 5105–5114, Red Hook, NY, USA, 2017. Curran Associates Inc.

[21] Xiaoyang Wu, Li Jiang, Peng-Shuai Wang, Zhijian Liu, Xihui Liu, Yu Qiao, Wanli Ouyang, Tong He, and Hengshuang Zhao. Point transformer v3: Simpler, faster, stronger. In *2024 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pages 4840–4851, 2024.

[22] L. Truong-Hong and Roderik Lindenbergh. Extracting structural components of concrete buildings from laser scanning point clouds from construction sites. *Advanced Engineering Informatics*, 51: 101490, 2022.

[23] Dongsheng Li, Jiepeng Liu, Shenlin Hu, Guozhong Cheng, Yang Li, Yuxing Cao, Biqin Dong, and Y. Frank Chen. A deep learning-based indoor acceptance system for assessment on flatness and verticality quality of concrete surfaces. *Journal of Building Engineering*, 51:104284, 2022.