Automation in Construction 162 (2024) 105392

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

From dense point clouds to semantic digital models: End-to-end AI-based automation procedure for Manhattan-world structures

Mansour Mehranfar *, Alexander Braun, André Borrmann

Chair of Computational Modeling and Simulation, School of Engineering and Design, Technical University of Munich, Arcisstr. 21, Munich 80333, Germany

ARTICLE INFO

Keywords:
Digital model
Point cloud
Artificial intelligence
Optimization process
Digital twinning

ABSTRACT

The paper presents a novel method for automatically creating semantic digital models for buildings in Manhattan environments from dense point cloud data. Unlike previous approaches, which rely solely on data-driven methods, our method integrates artificial intelligence with domain engineering knowledge to overcome challenges in indoor point cloud processing and geometry representation in complex layouts. A feature-based Decision Tree classifier extracts main building elements, utilized in a knowledge-based algorithm for 3D space parsing. On this basis, an optimization process generates parameterized floor plans, used to finally create volumetric digital models. The method was validated on datasets from the Technical University of Munich and Stanford University, achieving a mean accuracy of approximately 0.08 m for model placement and 0.06 m for estimating element parameters, which highlights its effectiveness for generating a building's semantic digital model. This approach underscores the potential of AI integration in digital twinning workflows for more automated solutions.

1. Introduction

1.1. Built environment and facility management

Nowadays, the term built environment has emerged as a resonant term in Architecture, Engineering and Construction (AEC). The built environment is a crucial aspect of modern society that encompasses a broad range the physical structures and infrastructures including buildings, roads, bridges, etc. The significance of buildings in the built environment is undeniable, given their central role in shaping the ways in which people interact with their surroundings. However, buildings must be properly maintained to remain functional, safe, and aesthetically pleasing. In this regard, Facility management (FM) serves as a crucial link between building maintenance and overall facility operations, providing a comprehensive approach that ensures the effective functioning and sustainability of buildings in the built environment.

The traditional approaches for building facility management and redesign activities heavily depend on conducting regular physical inspections and using the documented drawings to identify maintenance issues and make re-purposing decisions. However, due to the ongoing changes within the built environment, these approaches can be time-consuming, costly and prone to containing outdated information [1]. Recently, the use of virtual digital twin (DT) models emerged as an innovative approach enabling well-informed decision-making through the quick and coherent provision of data in a geometrically-semantically visualized context [2]. According to the practical explored studies, a digital replica of building assets can improve communication, collaboration, and decision-making for facility management and redesign purposes [3–5]. DT models can be used to simulate various scenarios and assess potential monitoring requirements without physical inspections [6]. In addition, DTs possess the capacity to effectively meet the ever-changing demands of decision-making and condition monitoring by facilitating real-time updates and fostering collaborative efforts.

Overall, the potential benefits of using DT models in facility management can be summarized as follow:

- Visualization: examining the building space from various angles enhances comprehension of the building's systems and components, enabling detection and addressing of potential problems beforehand [7].
- Maintenance and repair: a detailed record of a building's systems and components for simplifying planning and execution of maintenance to decrease emergency repair expenses [2].

* Corresponding author.
E-mail address: mansour.mehranfar@tum.de (M. Mehranfar).

https://doi.org/10.1016/j.autcon.2024.105392
Received 27 July 2023; Received in revised form 11 March 2024; Accepted 12 March 2024

- Space management: effective overseeing and controlling space utilization in a building, including occupancy rates, room bookings, and upcoming renovation plans [8].
- *Re*-design: optimize and enhance existing designs by leveraging real-time virtual simulations (various prototypes with different configurations, materials, and parameters) [7].
- Refurbishment: detailed assessment for efficient refurbishment strategies and resource allocation by leveraging simulations and providing a shared platform for collaboration and decision-making [5].
- Energy efficiency: simulating the building's energy performance to optimize energy usage and lowering operating expenses [9].

To support these applications, up-to-date digital representations of the buildings including both 3D geometry and semantics are crucial. However, currently, significant manual effort is required to create high-quality digital representations from raw surveying data. This becomes even more problematic when considering the need to repeat this process at regular intervals. Thus, there is an urgent need for end-to-end automation of the process to allow its cost-efficient and reliable execution. This paper contributes to providing the required technology.

1.2. Semantic digital models in BIM and DT realms

Today, Building Information Modeling (BIM) and Digital Twinning have emerged as transformative technologies within the domain of Architecture, Engineering, Construction, and Operations (AECO), heralding a new era for intelligent, data-driven infrastructure in built world. While BIM encompasses a thorough digital representation of the building's physical and functional characteristics, DTs further advance this concept by creating a real-time, virtual replica that faithfully mirrors the behavioral and performance aspects of the physical asset.

Specifically, BIM comprises a comprehensive digital representation of a building's geometric, spatial, and functional characteristics, enabling collaborative simulation and management of the projects throughout their lifecycle. DTs transcend static representation, incorporating real-time data streams from sensors, Internet of Things (IoT) devices, and operational systems to construct a dynamic, virtual replica that faithfully mirrors the behavioral, performance, and interactive aspects of the physical asset. At the core of these advancements lies the development and utilization of digital semantic models that bridge the gap between static physical-geometrical information and dynamic, real-time insights. These digital models leverage ontologies, semantic annotations, and standardized data schemas to impart data with meaning and context, thereby fostering interoperability and enabling seamless information exchange across disparate systems.

Over the past decade, the realization of semantic models has become achievable through the capturing of the built environment using multi-sensor remote sensing technologies. The raw data can be transformed into consumable information for human and machine interpretation using developed methodologies and techniques in the realms of computer vision and computational modeling. The aim of this research paper is to propose a novel 3D reconstruction algorithm for the automatic creation of the digital building model. The resulting digital models foster BIM and DT development, facilitating seamless interoperability and unlocking the potential for intelligent and data-driven infrastructure.

1.3. Problem statement

With the increasing complexity of buildings and the dynamics in their use, the need for up-to-date information about buildings and indoor environments for Operation and Management (O&M) has risen substantially [10]. DTs are a beneficial tool for digitally representing an asset's physical and functional properties and a shared data source for building information modeling (BIM) that provides a reliable basis for decision-making throughout the project life cycle. DTs using the efficient management of information and facilities have become an important tool in controlling costs and preventing waste in buildings' repair, maintenance, and energy resource consumption.

The utilization of a DT model enables the continuous updating of assets at frequent intervals, effectively incorporating the time dimension into the elements. This allows for real-time monitoring, catering to the requirements of efficient facility management for engineers and the maintenance team. Given the rapid and dynamic nature of urban assets, the development of an automated process to update the digital models can be replaced with traditional time-consuming and expensive manual methods. Developing a robust and dynamic pipeline requires using different geometric and semantic data and utilizing modern machine learning approaches to solve a multi-modal problem, including 3D geometry analysis and regression, type classification, and prediction of future asset changes [11].

Recent advancements in the remote sensing domain allow capturing accurate, multi-scale built environments with various technologies such as laser-scanning, image recording, and radar on a diversity of platforms including UAVs, cars, planes and satellites. In this regard, laser scanners have become a critical component in the creation of realistic and dynamic digital models for built environments, owing to their ability to capture highly accurate point clouds of scenes providing rich geometric and semantic detail [12–15]. Despite advanced development in engineering and technology, automatic or semi-automatic building digital model creation using point cloud data is still an open topic in the engineering and design society which needs novel approaches. Indoor environments can be large-scale and consist of complex spaces. Also, indoor scenes have been consistently associated with challenges such as noise, clutter, and obstructions [1]. Most developed methods for creating digital models rely on data-driven approaches, which are susceptible to poor data quality and encounter numerous challenges in reconstructing meaningful objects and simulating their relationships. Thus, the creation of a dynamic digital model based on the BIM and DT concepts needs an understanding of the sophisticated relationships between assets in the built environment [16,17]. To address these issues, this research aims to develop a novel approach for the creation of high-quality semantic digital models with coherent geometry which foster the development building DT with Manhattan-world structures. The proposed approach focuses on ensuring logical consistency between asset instances in the model, such as seamless connections between walls and slabs. This is achieved by incorporating semantic information and employing the parametric modeling technique to simulate their functional geometric relationships according to the Manhattan-world structural design.

1.4. Manhattan-world structures

Manhattan-world structures are characterized by a grid-based layout wherein structural elements (e.g. walls, slabs, and etc.) are interconnected in an orthogonal and aligned configuration. Adherence to the principles of orthogonality extends beyond aesthetic considerations, significantly impacting the building's performance and the efficient, dynamic utilization of its spaces. The design of Manhattan-world structures strategically maximizes usable space within the building, particularly in the layout of typical apartments.

In contemporary urban architecture, managers and urban planners increasingly value statistical analyses and objective observations pertaining to the utilization of the Manhattan-world structures design approach. The reasons for this embrace include enhanced efficiency, spatial optimization, simplifying construction processes, and facilitating the resource utilization.

Given the substantial number of educational and administrative buildings constructed with the Manhattan-world design, and the increasing trend of new buildings adopting this configuration, numerous efforts have been undertaken to develop 3D reconstruction methods for these structures using remote sensing data. While the geometric representation of Manhattan-world structures may seem simple, the automatic creation of a digital model for a building with a Manhattan-world configuration necessitates the precise formulation of the orthogonal grid structure within a highly parameterized system.

1.5. Contribution

Raw remote sensing data serve as the foundation for the creation of digital models. However, the interpretation of unstructured point clouds and images for accurately identifying and classifying building elements and their geometrical representation is still a challenging task. While common data-driven methods facilitate deriving contextual information and geometric representations using point cloud data, they often limited to automation and require numerous parameter settings and individual configurations for different datasets. These challenges are particularly intensified in the segmentation and classification of the main structural elements within point cloud. Conventional methods, such as plane fitting using the RANSAC algorithm or Region Growing, require a re-evaluation of the default parameters when confronted with different building point cloud. Recently, the capabilities of AI methods have promised tremendous potential for enhance the level of automation in indoor scene understanding task. However, achieving high performance in creating building digital models using AI techniques requires the collection of huge amount of training data and incorporation of domain knowledge contributing to enhanced reconstruction and simulation of the building system. These include inferring the interactive relationships between the elements of the building structure along with the assumptions and pre-determined characteristics (e.g. spatial organization, aesthetic considerations and etc.) in the design and construction of the building structure in the real world.

In this paper, we propose a novel framework for the automated creation of building digital models with rich semantic and coherent geometry from a dense laser-scanner point cloud dataset. The proposed method combines domain knowledge in design and construction with AI and optimization techniques capabilities to simulate the contextual relation between elements and improve their logical consistency. Thanks to applying the parametric modeling process, we are able to design a dynamic system with the ability to be manipulating the model by adjusting the value of parameters using inner predefined constraints which provide a basis for frequent geometric updates over the operation lifespan.

The resulting digital models offer a semantic volumetric representation of building structure corresponding to the schematic design of Level of Development (LOD 200) that supports decision-making processes in the operation and maintenance phases [18,19]. In the context of facility management, relying on extensively detailed digital models of buildings may not always be feasible or advantageous. In the BIM domain, the generalization concept has been developed to simplify the geometry of digital models while considering the LOD definition [20]. Geometry generalization involves simplifying the complex geometric representation of the model and abstracting it into simplified forms that still capture the essential structure of the building and its details. Meanwhile, the generalization concepts align with the LOD definition. The LOD framework provides a standardized scale for categorizing BIM models based on their level of detail, ranging from LOD100 (conceptual) to LOD500 (as-built) [21,22]. In this regard, a digital model in LOD(200) is a virtual volumetric model of a building system in which all graphical and non-graphical information including geometric property, and spatial relationships can be presented. By considering the LOD framework, the generalization methods ensure that the resulting models are appropriate for facility management purposes and allow them to effectively utilize digital models tailored to their specific operational needs, striking a balance between information richness and practicality.

In particular, this paper presents the following contributions:

- Designing a hybrid bottom-up, top-down approach for automatic building digital model creation.
- Utilizing an Artificial Intelligence (AI) classification technique for extracting main structural elements (e.g. wall, ceiling and etc.) within an indoor point cloud.
- Aligning the capabilities of AI in scene understanding with domain knowledge to overcome challenges in indoor 3D reconstruction.
- Developing a knowledge-based method for 3D space parsing in large-scale indoor point cloud.
- Reconstruction of the parametric building's floor plan using point cloud with low sensitivity to poor data quality.
- Creation of dynamic parametric building digital models (LOD200) for flexible facility management and space analysis.

The present paper is structured as follows: In Section 2, a literature review on the topic of building digital model creation using point cloud data is presented. Section 3 presents the developed methodology, which is described in detail from a theoretical standpoint. Section 4 showcases several case studies to demonstrate the feasibility of the proposed approach. Finally, Section 5 offers a discussion of the main findings of the study, along with future directions for research in the field.

2. Background

Over the past decade, there has been a significant increase in the utilization of advanced data acquisition techniques in the built environment, leading to a growing trend of automatic creation of digital representations of buildings with coherent geometry and rich semantic information. A significant number of studies have extensively investigated different facets of creating digital models of buildings under the term “Scan to BIM”. In this process, researchers utilize remote sensing data for creating digital models, which involves employing a range of sophisticated techniques encompassing data acquisition, processing, and modeling [10,23–26].

Xiong et al. [27] developed an automatic 3D reconstruction framework that used the voxelized point cloud to recognize the patches such as walls, ceilings, or floors based on boundary limits. Ochmann et al. [1] developed an automatic data-driven approach for reconstructing parametric 3D models of the indoor environment, including the floor, ceiling, and wall elements through volumetric 3D solid shapes. They employed a global optimization method to find and locate walls. However, their proposed method requires input data such as the availability of separate scans as an initial room segmentation. Later in [28], the authors proposed a novel method for reconstructing parametric, volumetric, and multi-story building models to improve their previous works. They define the modeling task as an integer linear optimization problem.

In [29] the authors used the combination of geometrical features of planar surfaces and their topological relation (e.g., distances and parallelism) to reconstruct the geometric and semantics of indoor volumetric models. In [30] the authors presented a combination method consisting of shape grammar and a data-driven approach that uses a reversible jump Markov Chain Monte Carlo (rjMCMC) algorithm to guide the automated application of grammar rules in the derivation of a 3D indoor model. In [31] the authors proposed an automated framework for generating building models utilizing a voxelized point cloud. The proposed method initially employs the planar cuts algorithm, followed by voxel-based morphological operators to segment the point cloud corresponding to different floors and rooms. Subsequently, for each segmented room, the boundary voxels are extracted, and volumetric wall instances are generated through morphological skeletonization.

In [32] the authors developed a multi-step automated approach for constructing 3D building models using point clouds. The method initiates with a preprocessing step in which point cloud data undergoes a series of operations, including the segmentation of 3D spaces, separation of walls, and subsequent clutter removal. Following this, an edge detection algorithm is employed to extract lines within a 2D image derived from projecting wall points onto a 2D plane (2D learning). Concurrently, a point cloud semantic segmentation algorithm is utilized to separate points into classes such as walls, doors, and windows (3D learning). Subsequently, the RANSAC plane fitting method is applied to separate wall instances, and the DBSCAN clustering algorithm is employed to cluster door and window instances. Finally, a bounding box is fitted to the points corresponding to each individual element.

Hu et al. [33] employed a robot-assisted mobile laser scanning approach for the 3D reconstruction and semantic segmentation of indoor point clouds. The proposed approach initially integrates the Simultaneous Localization and Mapping (SLAM) algorithm with robot motion control and path finding to achieve 3D mapping and acquire comprehensive point cloud data. Subsequently, the ResPointNet++ is employed for semantic segmentation of the reconstructed point cloud and group the architectural and structural objects.

In [34] the authors introduced a progressive model-driven approach for the 3D modeling of indoor spaces employing watertight predefined models. This approach initially segment spaces into rectangular and non-rectangular regions with an even number of sides. Subsequently, a point density occupancy map is used to enhance the level of detail in the intrusion and protrusion parts of Manhattan and non-Manhattan models.

To enhance the evaluation criteria for reconstructed digital models using Scan-to-BIM approaches, Jarzabek-Rychard and Maas [35] introduced an analytical method to estimate the level of geometric uncertainty in reconstructed digital models. The proposed approach assesses the dependency between point cloud data and the resulting 3D model. This involves calculating correlations between surface parameters of segmented instances in the point cloud and their corresponding reconstructed elements surfaces in the digital model. For each instance, distinct confidence intervals are computed based on specific tolerance thresholds, providing a quantified measure of the probability associated with the accuracy of the reconstructed building elements.

According to the [29,30], the process of creating a precise as-built digital model necessitates characterizing objects' shapes, relationships, and attributes. Therefore, the automatic creation process of digital models for building structures is typically divided into two major steps:

- Indoor point cloud processing for data semantic enrichment.
- Geometry provision for representation of parametric or non-parametric models.

Fig. 1 provides a comprehensive overview of the current methodologies utilized for point cloud feature enrichment, along with the prevalent techniques employed for presenting geometric digital models. The corresponding aspects of these approaches will be elaborated in the subsequent sections.

2.1. Indoor point cloud processing (semantic enrichment)

Indoor point clouds are a crucial tool for creating precise and accurate 3D models of indoor environments, which can be utilized for a variety of applications, including architectural design, construction planning, and virtual reality simulations. However, raw point cloud data is typically unstructured and challenging to interpret without proper labeling. The process of semantic enrichment, which involves assigning semantic labels to individual points or groups of points within the point cloud based on their geometrical/spectral attributes and features, is critical to transforming unstructured point cloud data into a

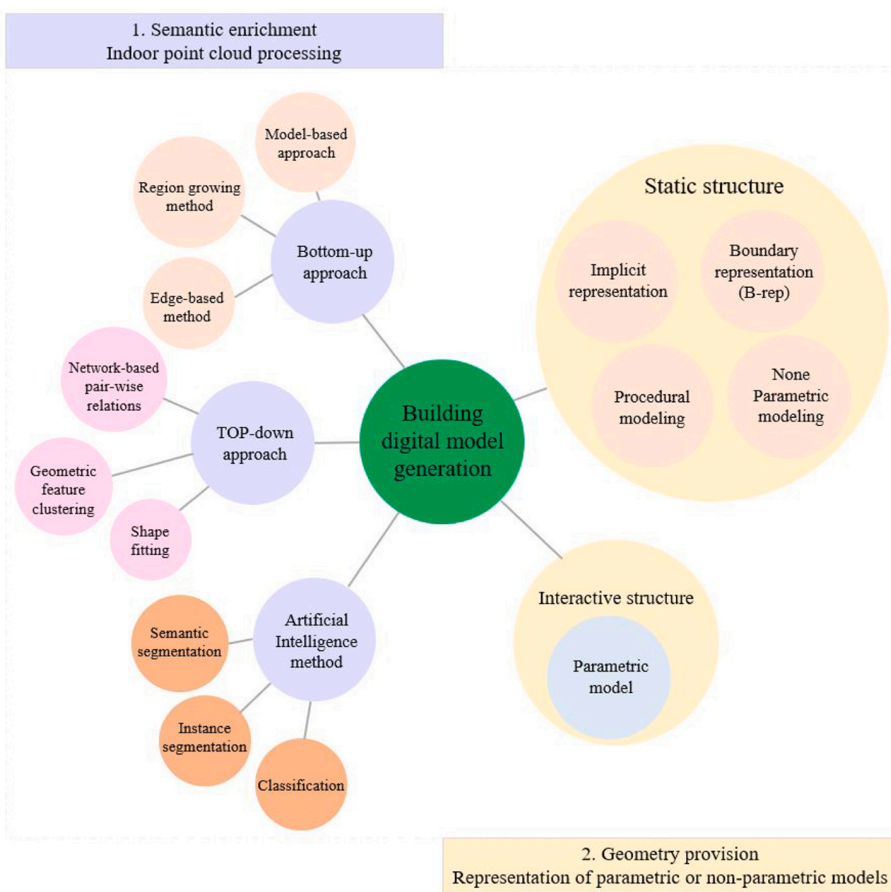


Fig. 1. Overview of the existing approach for semantic enrichment and geometry provision tasks in building digital model generation.

comprehensible and valuable resource. Accurate labeling is essential for distinguishing between structural elements within the environment. Various algorithms and techniques such as clustering, semantic segmentation, and classification can be employed to achieve accurate labeling through traditional methods or machine learning approaches. However, these techniques require careful parameter selection and tuning to ensure precise and accurate labeling.

2.1.1. Bottom-up approach

The bottom-up approach involves the labeling process starting from several seed points and is gradually extended to all points until a higher-level surface, volume, or model is generated [36]. These higher levels are commonly represented by meshes [37], voxels [38], and planes [39]. This approach utilizes typical features such as normal vectors, curvatures, and RGB channel values to express the differences between the geometrical and spectral details of surfaces. The labeling process in this approach is achieved through popular methods such as Region Growing (RG), Model-based, and Edge-based techniques.

The process of Region Growing commences with a predetermined set of initial seeds and subsequently incorporates neighborhood points based on various criteria such as surface normal, curvature, co-planarity, or RGB values. This addition continues until certain prerequisite conditions are met, including the detection of a non-surface point or the seed point reaching a distance threshold. Region Growing algorithms can be applied to both structured and unstructured point clouds [40]. However, the algorithm tends to over-segment objects in the presence of non-trivial occlusions, thus posing a significant challenge when processing indoor building scenes. In [41] authors proposed a supervised Region Growing method for segmentation of unstructured point cloud using geometric features such as surface roughness and curvature. The proposed method utilizes a locally adaptive threshold based on a predefined parameter for the desired level of abstraction. This threshold is later used for considering the local context and aiding in seed finding during the region-growing steps.

Model-based methods use mathematical expressions for primitive shapes such as spheres, cylinders, cones, and planes for clustering points. These methods are repeated iteratively to find bunches of points that have the highest correlation with the pre-defined shapes. The most well-known model-based method is the random sample consensus (RANSAC), which is used to detect planar surfaces (e.g. walls, floors, ceilings, etc.,) in the point cloud of buildings [42,43]. Although model-based algorithms are quite effective even in the presence of noise and outliers due to mathematical foundations, they also have problems overlapping multiple references around boundaries and require prior knowledge about the data.

Edge-based segmentation methods characterize the objects based on the rapid changes in any geometrical or spectral feature of points. In [44,45] the authors proposed a method for representing surfaces by calculating the gradient of points and following the changes in the direction of the normal vectors in a neighborhood. In [46] the authors employed Hough Transform-based methods and used normal vector values to detect 3D roof planes in laser scanner point clouds. Similar to the model-based approach, Hough-transform is robust toward noisy and cluttered data. However, due to its sensitivity to parameter dimensions and high computational requirements, it cannot be applied to shapes with too many parameters.

2.1.2. Top-down approach

The top-down approach begins by dividing a point cloud into compositional sub-problems based on elements' existing similarities and dissimilarity features. In the top-down approach, perceptions start with the most general and move toward the more specific. In [47], the authors proposed a network based on pair-wise relationship rules such as parallel, equal height, above, under, and orthogonal for semantic mapping indoor building elements (e.g., walls, and floors). In [48], the authors used a top-down approach to reconstruct the geometry of architectural buildings from laser scanning point cloud data. They first cluster the planar surfaces of buildings based on the analyzing confidence rate and then create the polyhedron models by computing plane intersections and corners.

In [49], the authors proposed a hierarchical segmentation model which combined the top-down approach with a recurrent neural network (RNN) to decompose point clouds into different shape segments. The top-down and bottom-up methods represent two distinct approaches, each with its own set of advantages and disadvantages. Depending on the problem, either approach can offer a practical solution. However, one of the critical challenges when implementing these approaches is accurately defining the problem and effectively breaking it down into subproblems. Regardless of the method used, a high-level prior knowledge about the problem is typically required, and applying these methods to address any problem necessitates adherence to a specific set of rules and reasoning.

2.1.3. AI-based methods

After the significant expansion of machine learning and deep learning concepts over the past decade, large-scale data classification and segmentation have emerged as the most recent areas of research in computer vision and the construction industry. In [50] authors proposed a Density Based Spatial Clustering algorithm (DBSCAN) to segment planar objects within indoor point cloud. The proposed bottom-up method utilizes the geometric and spatial features of points extracted from plane fitting through K-means unsupervised learning. This approach is employed to evaluate the semantic contents and classify planes belonging to the major building planar elements. Finally, a bagged decision tree classifier is used to predict the object class belong to the walls, ceilings, and floors and etc.

Numerous network architectures have been specifically developed to effectively predict unknown labels in classification and segmentation problems while processing point clouds. Due to the unstructured nature of point clouds, voxelization deep learning methods were initially developed as a solution. These methods incorporate standard convolutions and leverage both local and global features of points to enable segmentation and classification of the point cloud based on a combination of these features [51,52]. In [53], the authors used 3D Convolutional Neural Networks (CNNs) to solve a binary classification problem. Later, in [54], the authors used raw point cloud directly as input data and learned the local features of points by Multi Layer Perception (MLP) and symmetric functions. Meanwhile, some research investigated utilizing Kernel Point Convolution to operate point clouds for segmentation tasks [55]. In the proposed method, the network learns the local geometry by considering kernel points and applying them to the input points close to them.

With the development of the graph concepts in the point cloud domain, graph-based convolutional methods were developed to reduce the computational cost and efficient use of neighborhood points' geometry [56]. The developed Graph-based networks use graph theory to find the spatial neighbors of each point and generalize CNN layers to adapt to the graph's structural data. To perform a semantic segmentation of an indoor point cloud, Armeni et al. first parsed the building environment into individual spaces using density-based histogram analysis [57]. They then implemented an instance segmentation model to parse spaces into their structural and building elements. In [58], the authors implemented an Image-based deep network for scene segmentation by projecting points to an intermediate 2D grid structure using different viewpoints. In [59], Dai et al. developed a method for 3D semantic scene segmentation of RGB-D scans in indoor environments which combines and joins RGB values and depth geometry data in end-to-end network architecture in their proposed method. Recently, in [60], the authors developed an efficient semantic segmentation network called point transformer, which uses the self-attention layer with a combination of simple linear layers and MLP. The point transformer layer is invariant to permutation and cardinality and is thus inherently suited to point cloud processing.

Despite the significant progress made in the field of AI, the developed networks for scene-understanding tasks face several challenges that can impede their performance. One major challenge is the requirement for numerous distinct datasets for training the network to achieve optimal results, particularly in the context of indoor environments with cluttered scenes, which can be both costly and difficult to obtain. Additionally, most AI architectures rely solely on local and global features in the training phase and decision-making process, while ignoring critical engineering construction knowledge such as general knowledge on building structures and the interaction of building elements. Incorporating such knowledge can serve as a key tool for overcoming obstacles and breaking down a general problem into smaller, more manageable tasks, ultimately leading to improved overall accuracy.

2.2. Geometry provision (representation of parametric or non-parametric models)

One of the fundamental prerequisites for achieving a DT is geometry provision, which involves representing the physical environment through three-dimensional models. The virtual model must exhibit topological consistency and geometric accuracy and should facilitate bidirectional interaction with a customized geometric abstraction of the real world. In recent years, various methods have been employed in the literature to represent 3D models, each with its own set of advantages and disadvantages. However, the creation and representation of the building digital model from point clouds has remained an unresolved issue, and most previous work is designed for specific types of buildings and is limited to reconstructing specific kinds of objects. In this regard, the state-of-the-art is reviewed across four distinct methodologies, as follows:

2.2.1. Implicit representation

Implicit shape representation is a method conducted by defining surfaces and primitive shapes with mathematical equations in Euclidean space. Implicit methods use implicit functions such as F (x, y, z) to represent curves and surfaces with arbitrary constructive topology. They are only applicable to specific symmetric shapes among all real-world 3D objects. In [61], the authors developed a dimension-independent algorithm based on the least-square method for fitting a shape (line, circle, conic, cubic, plane, sphere, quadric, etc.) to 2D and 3D space points. Kwon et al. [62] developed a real-time and accurate modeling algorithm to fit geometric primitives' shapes (such as planes, cuboids, and cylinders) to point clouds. In many studies, authors have recently tried to implement neural networks to learn a continuous implicit function representing shapes [63].

2.2.2. Boundary representation (B-rep)

The boundary representation method is a way to represent a 2D or 3D shape based on its vertices, edges, and loops and their topological relationship to form the object. In this method, vertices are described by their coordinates (x, y, z), and a parametric equation defines lines (straight or curved) or faces. In this regard, Xiong et al. in [27] developed an automatic 3D reconstruction framework that used the voxelized point cloud to recognize the patches such as walls, ceilings, or floors based on boundary limits. In the proposed approach, visibility reasoning is used to detect openings (windows and doorways) and holes in the surface. Finally, the building models generate by learning the unique features of points and their contextual relationships. In [64], the authors used histogram analysis to extract planar surfaces, such as walls, floors, and ceilings, from a laser scanner point cloud. Then, they implemented Hough transform to model candidate surfaces. They also used supervised learning by support vector machine (SVM) to detect opening parts. In [65], Stamati et al. developed an integrated approach based on morphology analysis techniques and the piecewise rational Bezier curves to extract features and boundaries from a 3D point cloud. In [42], the authors proposed an automatic approach for reconstructing and modeling architectural buildings. They first extract planes using the RANSAC algorithm and automatically create initial models by finding boundary polygons. They also extract local adjacency relations among parts of the polygons and use them to modify the fitted model through a snapping algorithm.

2.2.3. Procedural modeling

Procedural modeling is the method of 3D modeling geometries of solid shapes. The complete model of an object is represented by 3D primitive shapes (e.g., cuboids, cylinders, spheres, cones, etc.) or a combination of them using Boolean operations such as union, intersection, and difference. As one of the earliest investigations, Wang et al. [66] proposed a semiautomatic model-based approach for the 3D reconstruction of buildings using photogrammetric data. In their proposed pipelines, after selecting an appropriate model from the predefined library, a CSG (Constructive Solid Geometry) model is fitted to the data by implementing an optimization method. In [67], the authors proposed an automatic 3D reconstruction framework titled "Inverse CSG" to represent walls in indoor scenes with volumetric primitive shapes. Also, in [68], the authors presented an automatic method for representing 3D solid shapes of objects using point cloud data. In this method, a set of primitive shapes candidates are chosen, and then the target model is generated by the combination of a subset of candidates with corresponding Boolean operations using a binary optimization technique.

2.2.4. Parametric modeling

Parametric modeling involves the creation and representation of a dynamic geometric model system through dependencies and functions between parts/volumes. In contrast to solid modeling methods, the parametric modeling approach preserves a parametric format, allowing for the effortless modification of geometry by adjusting parameter values and incorporating metadata and semantic information. This adaptability proves particularly advantageous in the realms of BIM and digital twinning, where bidirectional links facilitate the updating of existing geometry based on input values [69]. It is considered a suitable method for digitally representing buildings and infrastructure, especially when frequent geometric updates are required over their service life.

In [70,71], the authors developed a parametric modeling approach for the creation of the digital model of bridges. The proposed method involves using a semantic segmentation network to extract various structural elements, such as the deck, abutments, railing, and piers. Subsequently, metaheuristic algorithms are used to adjust the parameter values of each structural element, resulting in the extraction of a model from the point cloud data and the creation of 3D volumetric shapes. In the case of building digital model creation, perpendicularity or parallelism of elements (e.g. walls, slabs etc.) or defining dimensional values such as length, distance, or angle could be a feasible example. In [1] the authors considered the parametric modeling of the building structure as a combination of model reconstruction and then set geometric rules for change the spaces layout and model modification.

Despite its effectiveness and extensibility, the parametric modeling process can be more labor-intensive comparing to the other modeling methods. Procedural parametric modeling requires a high level of prior knowledge about the geometry of the elements to establish a set of dependencies and topological integration between different parts and layers for adding geometric data to the target elements [69]. Moreover, an integrated parametric model require well-defined access points as crucial interfaces to accommodate received location-based queries and tailor the model to sub-views based on distinct use cases and LoDs [72].

2.3. Research gap

The rise in building construction projects has necessitated the need for smart asset management and sustainable decision-making in complex environments. In response, the concept of DTs has been extended to indoor environments to support building management and planning during the operation phase. Despite the growing interest in DTs, a significant research gap still remains in creating a digital representation model of indoor buildings from point cloud data.

Indoor point cloud data can be acquired through various methods such as laser scanning and photogrammetry. Despite their individual benefits and drawbacks, the quality of the point cloud data can be influenced by a range of factors, including lighting, occlusions, and reflections, which can pose challenges to the digital representation of the physical world. Thus, it is crucial to establish standardized protocols for data processing to ensure accuracy and consistency in the digital models. Most methods for processing indoor point cloud data are based on a bottom-up approach where the data is divided into sub-individual parts (such as an object or surface instances) and then 3D volumetric models are fitted. Although semantic segmentation/classification is intertwined with the model fitting process, this may not be easily achieved due to various obstacles such as noise, obstruction, and clutter in spaces. The data-driven approaches are particularly sensitive to data quality, especially occlusion, and their performance may decline in the presence of challenges such as clutter or complex building layouts. Therefore, leveraging domain knowledge through the parametric modeling technique is hypothesized to be a promising tool for creating digital models of real-world assets with coherent geometry and rich semantics in a comprehensive manner.

The parametric modeling approach enables the creation of accurate and detailed representations of assets, which can be used for various applications such as asset management, space planning, and renovation design. However, the design of appropriate parametric models and estimating the values accurately still have remained a challenge in representing the reliable digital model of indoor assets. In the realm of building digital modeling, extant literature expresses the parametric modeling process as a combination of model reconstruction and subsequently set geometric rules [1]. Also, in some research, parametric building modeling has been considered as the floor plan generation through a set of inputs and rules with the capacity to systematically generalize arbitrary constraints therein [73,74]. Nonetheless, an efficacious approach ought to incorporate the tenets of parametric modeling in the model reconstruction process to enable the estimation of parameters accurately and to guarantee consistency between the different components of digital models. This paper discusses the utilization of AI methods along with accumulated knowledge from construction and design domains to create a potent solution for the prevailing challenges in the field. The objective is to develop an automatic framework for the parametric modeling of a wide range of buildings under Manhattan world assumption, across different designs and layouts. Further insights into the proposed methodology will be expounded upon in the subsequent section.

3. Methodology

3.1. Overview

This section delineates the proposed workflow for creating highly parameterized digital models of indoor environments which rely on integrating the top-down and bottom-up approaches. The overall methodology of the proposed approach is visually depicted in Fig. 2. First, multi-step preprocessing techniques are employed to eliminate outliers, minimize noise as well as separate different levels in multi-story

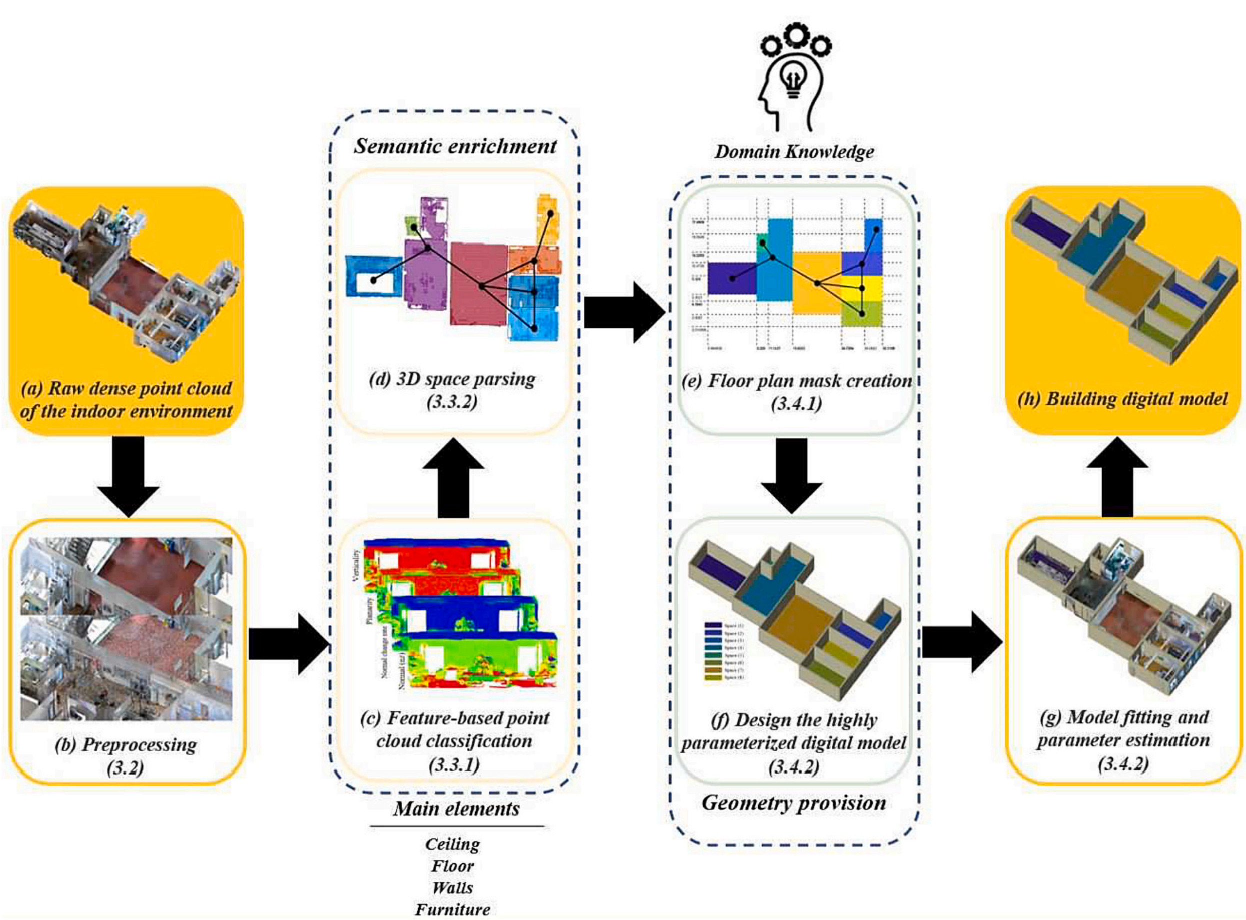


Fig. 2. Overview of the proposed end-to-end method for building digital model creation.

building scenarios to facilitate the modeling process and ensure the resultant 3D models effectively reflect the indoor environment. Next, a bottom-up approach is implemented through a Decision Tree classifier model to identify the main building's structural elements (e.g. walls, ceiling, and floors) in the indoor scene point cloud. Subsequently, the derived information is utilized in a knowledge-based 3D space parser algorithm that converts complex indoor environments into individual subspaces. The top-down approach involves designing highly parameterized building models that mirror the typical topology of buildings, presenting human knowledge in building design and engineering. This approach leverages a model-based technique, coupled with an optimization process, to create the highly parameterized floor plan maps which present the current spatial layout of the environment. Then, based on the pre-defined geometric and semantic rules and constraints, the extruded 3D models are fitted to the point cloud data using a high-dimensional optimization problem by employing simplex optimization methods. The modularized building models provide sufficient degrees of freedom to allow the modeling of a wide range of different real-world buildings (Fig. 3).

3.2. Preprocessing

The process of manually digitizing and reconstructing digital models of objects needs a significant investment of time and costs. In this regard, the development of algorithms for the automated creation of digital models is a highly important area of research. Laser scanning is the technology that can capture data for the creation of 3D models, thereby motivating researchers to develop fully automatic algorithms in this field. However, it is crucial to pay special attention to the preprocessing of input data before executing the main computing processes. Among the essential factors, the subsampling of the point cloud for improving the processing time and filtering of the noises and obstruction are the utmost important steps that require particular attention in preprocessing methods.

Subsampling is a process of reducing the number of points in a point cloud while preserving the geometric structure and relevant information of the scene, which reduces the computational complexity of subsequent main operations. Subsampling can be achieved using various techniques, such as random sampling, and grid-based sampling. In random sampling, a subset of points is selected randomly from the original point cloud. This method is straightforward and efficient but may result in non-uniform point distribution and loss of important features.

Grid-based subsampling involves dividing the point cloud into a grid of equal-sized cells and selecting one point per cell. This method ensures a more uniform distribution of points but may result in the loss of fine details and features. Given that the proposed method employs model-driven approaches, the level of dependence on specific details is significantly reduced. Thus, a grid-based subsampling method is utilized to ensure the effectiveness of the proposed approach as well as to improve the time in the point cloud processing steps (Fig. 4).

Noise and outliers are undesirable points that do not conform to the overall data distribution, and can arise from different sources such as sensor noise, environmental conditions, and moving objects. To remove these unwanted points from the dataset, statistical or neighborhood analysis methods can be used. One of the effective methods for filtering out noise and outliers is the Connected Components Segmentation (CCS) approach, which involves setting a distance threshold and then identifying all connected points within the threshold as a separate segment [75] (Fig. 5). In this regard, noise and outliers often appear as small or isolated segments in indoor scenes, which can be easily identified and eliminated from the data.

In the preprocessing of the multi-story building point cloud, the separation of different building levels is a crucial step for the accurate modeling and representation of its structural elements. While several data-driven methods have been developed to separate floors and ceilings in a multi-story building point cloud, these methods rely on histogram analysis of points density along the Z axis and then filtering peaks as the level separators [76]. However, the developed methods are applicable only to a specific range of buildings. They face different challenges in the presence of complex and specialized structural designs, such as varying the floor's heights and the inclusion of stairs or ramps that connect different floors. Addressing the corresponding problems needs selecting parameters and performing the steps manually.

Consequently, there is a pressing need to develop more adaptable

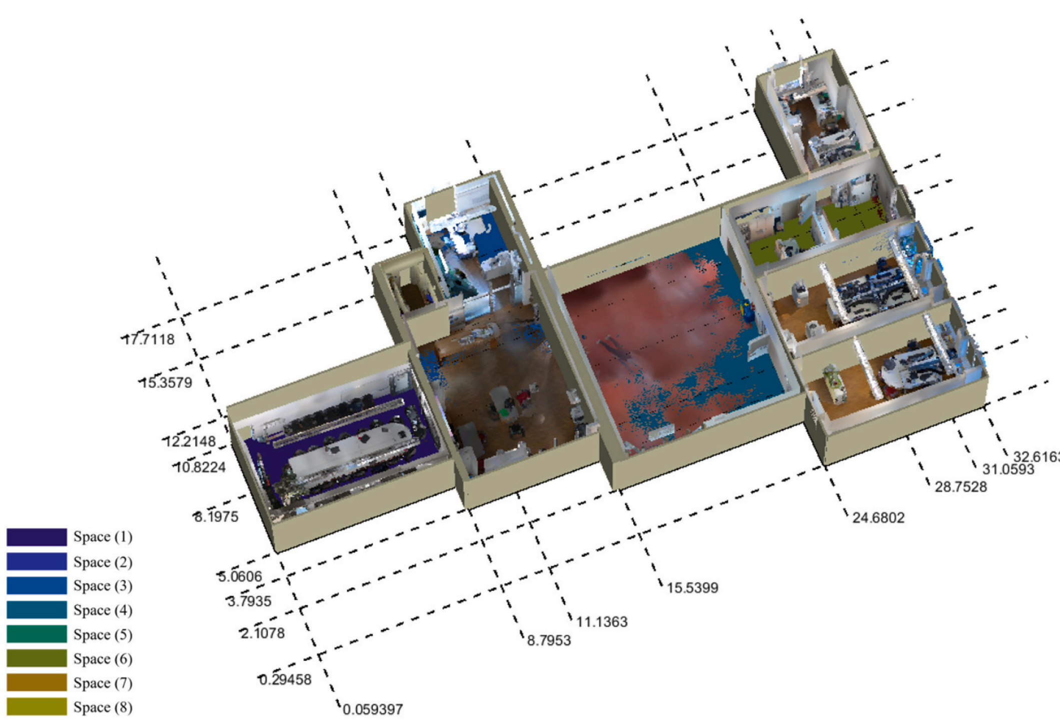


Fig. 3. Creation semantic digital building model through parametric design and model fitting.

Fig. 4. Grid-based point cloud subsampling.

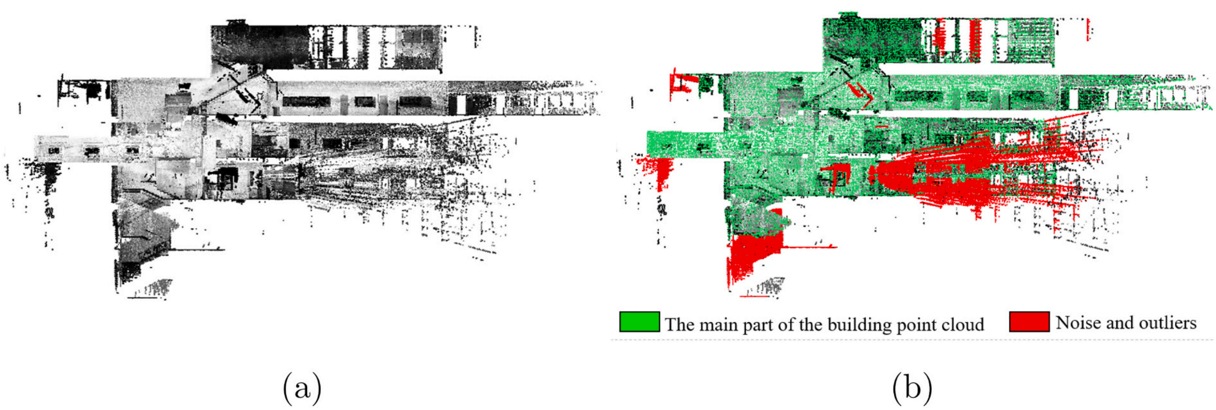


Fig. 5. Filtering of the noise and outliers using the CCS method: (a) raw point cloud data, (b) the result of filtering noise and outliers.

techniques to address these limitations and overcome the challenges posed by unique structural conditions. In this study, we employ the property of the point's normal vector to separate the various levels of the building. Specifically, we utilize this characteristic to differentiate between the ceiling and the floor of each level. Surface normals (n_x, n_y, n_z) are vectors that are perpendicular to the tangent plane at a specific point on the surface, and indicative of the surface's characteristics. In order to calculate the Normal vector values for a point on the surface in the point cloud space, the covariance matrix and eigenvector values of the nearest neighboring points are computed and analyzed [77]. This involves determining the covariance matrix c for a given point p by (1):

$$c = \frac{1}{k}\sum_{i=1}^{n}(p_i - \overline{p}).(p_i - \overline{p})^T \quad (1)$$

where k is the number of neighboring points, p_i and p also refer to the 3D coordinates of the points being considered. Additionally, the eigenvalues and eigenvectors are determined by (2) where λ and $\overrightarrow{v}$ are eigenvalues and eigenvectors respectively.

$$c.\overrightarrow{v}_j = \lambda_j.\overrightarrow{v}_j, \quad j \in \{0,1,2\} \quad (2)$$

For a simple room's point cloud, the Normal vector features of both the ceiling and floor points have typically the highest values along the Z axis. However, to ascertain the direction of the Normal vector of the ceiling and floor in point cloud space, a designated viewpoint is utilized, which is often situated at the geometric centre of the points. This enables the distinction between the ceiling and floor points based on their respective vector directions, which are opposite (Fig. 6).

In the context of a multi-story building scenario, the Verticality feature of points serves as a primary tool for the detection of horizontal planes (Fig. 7). The Verticality feature value is determined by computing the covariance matrix of neighboring points and normal vector value n_z by (3) [78]. For horizontal points of ceiling and floor elements with the highest n_z value, the corresponding Verticality feature value is equal to zero.

$$Verticality = 1 - n_z \quad (3)$$

Upon calculating the Verticality feature and extracting candidate points, the connected components segmentation (CCS) method with the parameters of a radius of about 25 cm and a minimum number of 5000 points per segment is utilized to decrease the noise of furniture elements and separate the major segments representing ceiling and floor points candidates on different levels. Calculating the Normal vector value and identifying its direction requires determining unique viewpoints on each

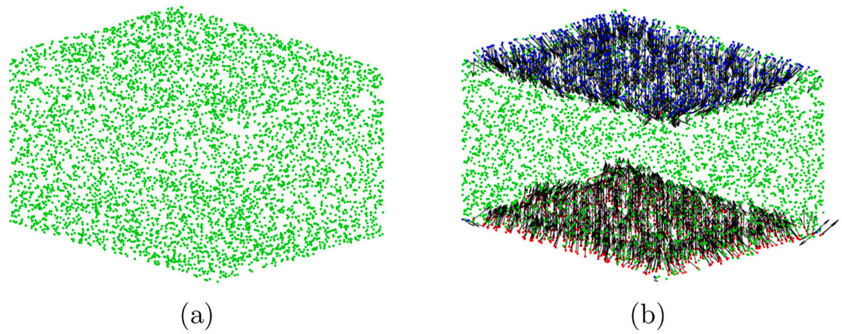


Fig. 6. Normal vector characteristic: (a) the point cloud of a simple room, (b) calculation of normal vector values and separating ceiling and floor points.

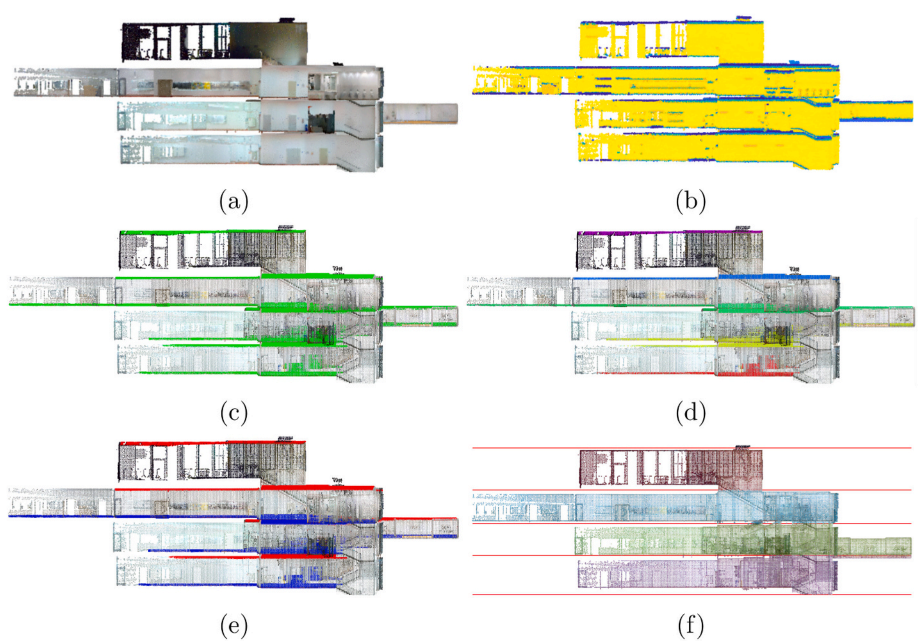


Fig. 7. Proposed method for level separation in multi-story building scenario: (a) original point cloud after the filtering process, (b) the verticality feature, (c) extracting major planar segments using the CCS method, (d) ceiling and floor point candidates in different level, (e) separating ceiling and floor points based on normal value and direction toward the viewpoints, (f) the point clouds of individual building levels.

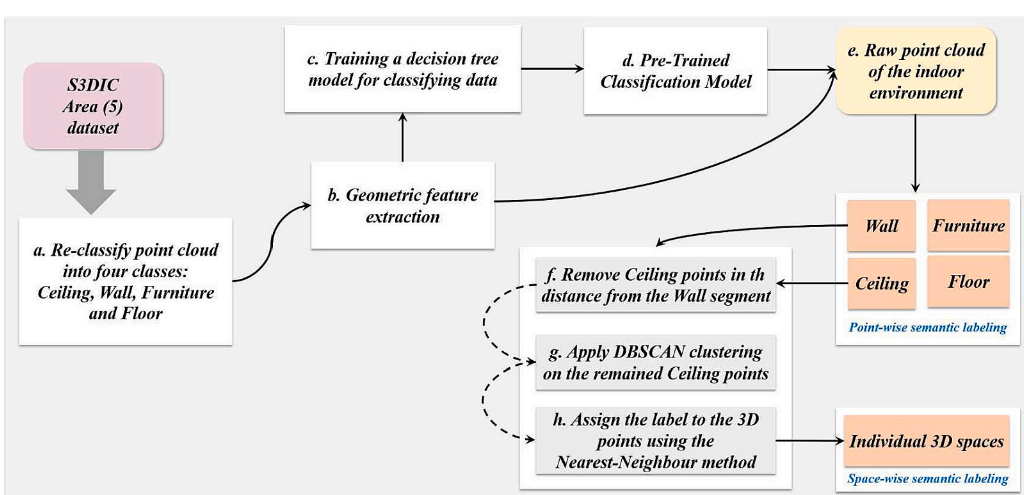


Fig. 8. Proposed workflow for semantic enrichment of indoor point cloud, a bottom-up approach for point-wise and space-wise semantic labeling.

building level. In this regard, the information of candidate segments is utilized and the middle points of two consecutive segments along the *Z*-axis are designated as viewpoints, and subsequently, the points are categorized according to the Normal vector direction.

Finally, the ceiling points information is employed to crop and partition the individual level's point cloud. The proposed method for separating different levels within a multi-story building point cloud is distinct from other extant approaches in that it is not sensitive to the geometric intricacies and structural complexity and obviates the necessity of executing manual operations.

3.3. Semantic enrichment

3.3.1. Indoor point cloud classification

Point clouds are a prevalent form of 3D data that can offer valuable geometric information about physical objects or scenes. However, the raw data collected from laser scanners can be complex, unstructured, and unordered. To support scene understanding and automated analysis in the geometry provision step, this data requires higher-level contextual information. Semantic enrichment is the process of adding semantic meaning to point clouds by labeling each point with contextual information to make it more consumable for human or machine interpretation (Fig. 8). In this research, the process of semantic enrichment is handled by employing a Decision Tree classifier model and leveraging the geometric features to partition the raw point cloud into smaller classes based on spatial relationships and object type.

Indoor environment scans have always been associated with the frequency of objects, complex layouts of spaces, clutter, and obstruction. The creation of a building's structure digital model requires accurate classification and extraction of the main structural elements. Due to the coplanarity as well as the coherent 3D distribution of the main structural element's points compared to the furniture objects, the use of appropriate geometrical features can lead to the accurate classification of point cloud data. To implement and train the classification model, first, Area 5 of the Stanford University 3D dataset (S3DIC) is selected and the point clouds are annotated into four classes including; wall, ceiling, floor, and furniture (Fig. 9) [57].

In the following, the geometric features such as Planarity, Verticality, Normal value along the Z axis (n_z), and surface variation are calculated for each point using a spherical neighborhood with a radius of 0.25 m (Fig. 10). In this regard, the Planarity and change in Normal rate features are utilized to differentiate the main structural element points from furniture objects. Furthermore, Verticality and n_z features are employed to distinguish the vertical walls of the building from the horizontal surfaces of the ceiling and floor. Similar to the surface normal vectors and Verticality features, the Planarity is calculated by analyzing the eigenvalues of the covariance matrix using (4). The surface variation is also computed by comparing the normal vectors of neighboring points and calculating the angular difference between them.

$$Planarity = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{4}$$

To classify multi-class data, the decision tree is built by recursively splitting the data into subsets based on the values of the features until the subsets be homogeneous or the tree reaches a maximum depth. At each node of the tree, a decision is made based on a feature, and the data is split into two or more subsets, which are then passed down to child nodes. This process continues until a leaf node is reached, which contains a prediction for the class label. The maximum depth parameter is a tuning parameter that affects the performance and interpretability of decision tree models and has a crucial role to balance between underfitting and overfitting to obtain a model that can accurately classify new data points.

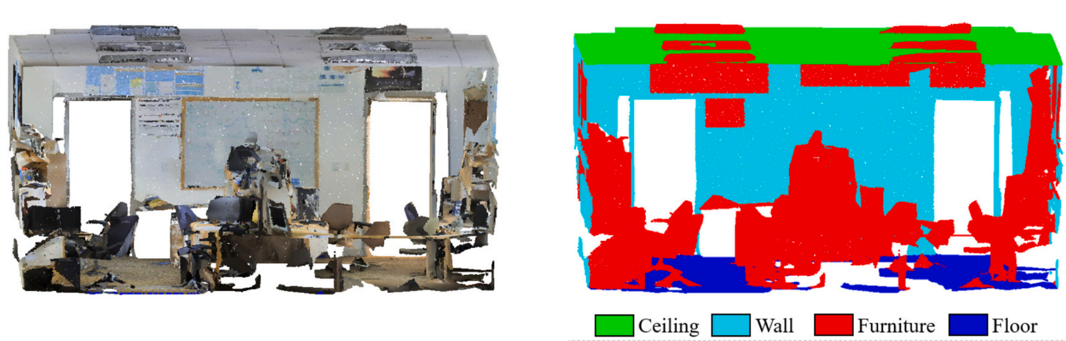
Fig. 11 shows the structure of a decision tree classifier with a maximum depth parameter equal to 3. Due to the complexity of indoor scenes and the use of different geometric features with similar values in the range of [−1,1], the depth parameter is experimentally considered equal to 10 to effectively explore the relationships between the features and search for relevant states.

To calculate the amount of uncertainty in data and determine the optimal split at each node of the tree, the criterion entropy is used. At each node, the algorithm calculates the entropy of the data based on the distribution of the target classes. It then splits the data into subsets based on the values of the features and calculates the entropy of each subset. The entropy criterion is less prone to bias when compared to other criteria (e.g. Information gain, Gini, Reduction in Variance, etc.), as it considers the distribution of the target classes in the data. For a dataset with N classes the entropy is calculated by (5):

$$E(s) = -\sum_{i=1}^{c} p_i log_2(p_i) \tag{5}$$

Where s is the current state, and p_i is the probability of an event i of state s.

The S3DIS dataset area 5 is divided into two sets, namely training and validation, with an 80/20 ratio, respectively. The common data partitioning ratios include 90/10, 80/20, and 70/30 subsets. The selection of these ratios is based on fulfilling the requirements of adequate training data, effective generalization assessment, computational efficiency, and the mitigation of data imbalance. The training set is used to train the Decision Tree classifier model aim to robust parameter learning, fostering the model's capacity to capture intricate patterns, while the validation set is utilized to evaluate the performance of the model on new, unseen data, mitigate overfitting and ensureing the comprehensive evaluation without compromising computational efficiency. To evaluate the performance of the model, annotated reference data and the classified points are compared using the standard metrics of



(a) (b)

Fig. 9. Preparing point cloud dataset for training the Decision Tree Classifier model: (a) the point cloud of the Stanford University 3D dataset (S3DIS) Area 5 [57], (b) the annotated point cloud data.

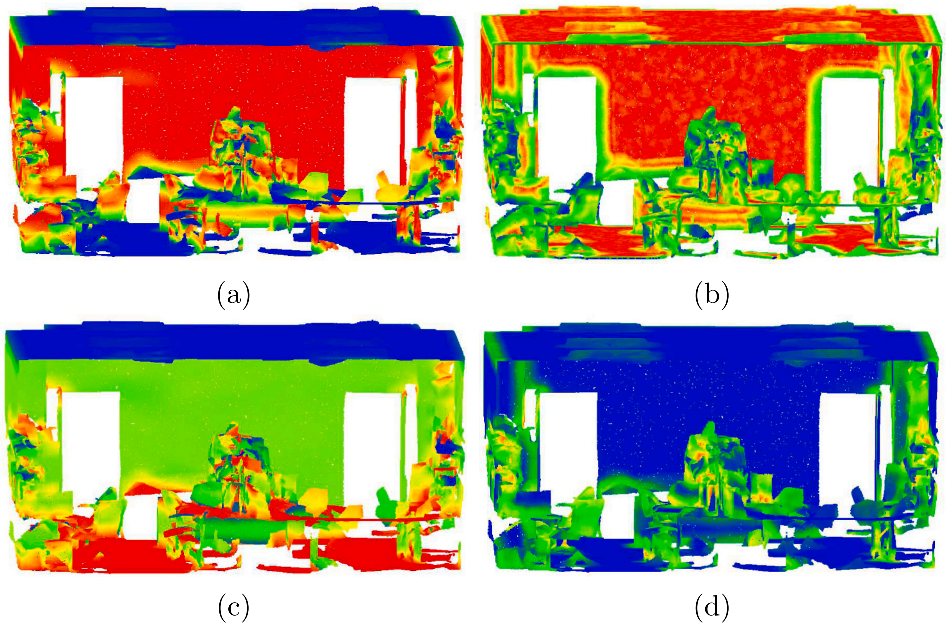


Fig. 10. Point cloud feature extraction: (a) Verticality, (b) Planarity, (c) Normal value along the Z axis (n_z), (d) Normal change rate.

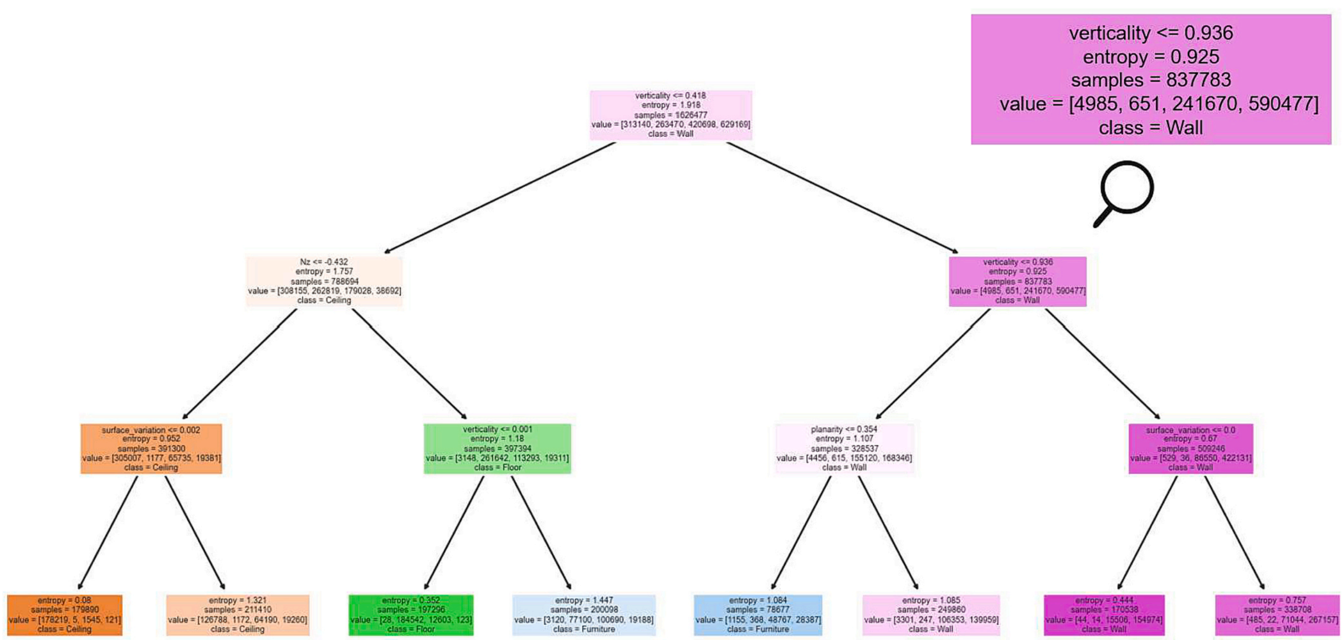


Fig. 11. The structure of the decision trees model with a maximum depth of three for the classification task.

Table 1
Accuracy evaluation of classification model.

	Precision	Recall	F-Score
Ceiling	0.90	0.92	0.91
Floor	0.83	0.93	0.88
Furniture	0.66	0.49	0.56
Wall	0.79	0.87	0.83
Overall accuracy = 0.79			

Precision, Recall, F-Score, and the results are reported in Table 1 based on Eqs. (6)–(8) where TP, TN, FP, and FN are True Positive, True Negative, False Positive, and False Negative, respectively.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F\text{-}score = 2.\frac{Precision.Recall}{Precision + Recall} \tag{8}$$

3.3.2. 3D space parsing

The skeleton of a building is delineated by pertinent data regarding its structural and architectural characteristics such as the area of interior spaces, the number of walls, openings, etc. When crafting an elaborated digital replica of a building, it becomes imperative to accord primacy to

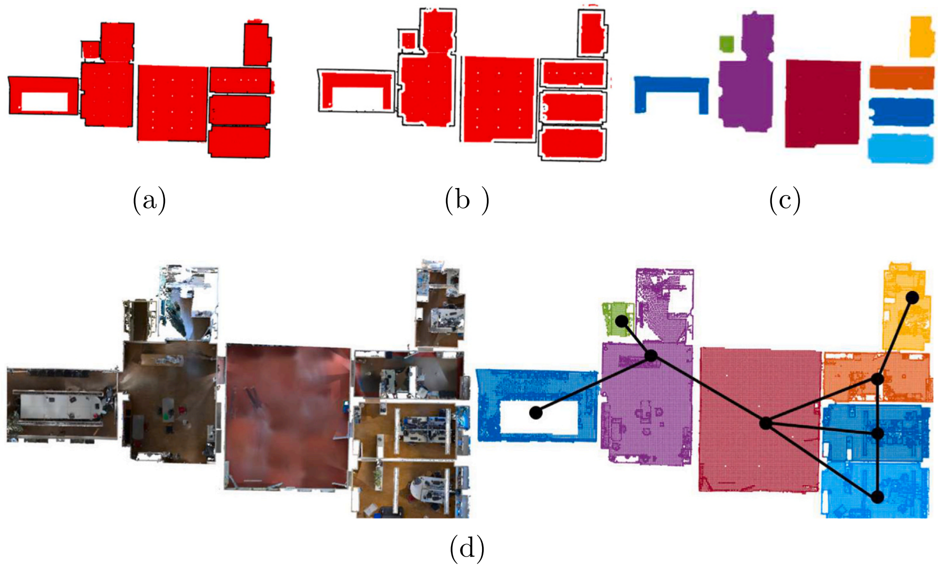


Fig. 12. Overview of the proposed algorithm for 3D space parsing: (a) ceiling and walls points on X-Y plane, (b) removing ceiling points at distance *d* from walls, (c) clustering remained ceiling points, (d) he result of 3D space parsing after applying the nearest neighbor method, along with the adjacency graph of 3D individual spaces (The ceiling points were removed to improve the visual representation).

the constituent rooms and spaces. The individual spaces play a pivotal role in facilitating the creation of an accurate and intricate virtual representation of the entire building [79]. According to the literature, an enclosed interior space is characterized by fundamental structural components such as the floor, ceiling, and walls, which exhibit topological relationships with other adjacent spaces [1].

Inferring the prevailing topological relationships between spaces is necessary for accurate geometric modeling of complex spaces and their analysis in internal navigation applications [29]. The major existing approaches for partitioning 3D spaces rely on prior knowledge of space layout or the precise location of laser scanners, rendering them suitable only for small-scale environments [80]. Moreover, unsupervised segmentation methods that employ features such as point density to delineate spaces typically produce results with over-segmentation defects and other associated drawbacks [57].

Considering all the aspects and limitations of existing methods, we propose a knowledge-based approach to overcome constraints. The proposed methodology combines both top-down and bottom-up techniques to enable effective partitioning of spaces within complex 3D environments. The proposed approach involves analyzing the intersection of wall and ceiling points to determine the central part of an enclosed space that is encompassed by walls (Fig. 12).

In this regard, the algorithm initially removes the ceiling points that are in *d* distance from the wall points, resulting in a ceiling segment that is fragmented and dispersed at a significant distance from the walls. The appropriate value for *d* is based on the average thickness of interior walls

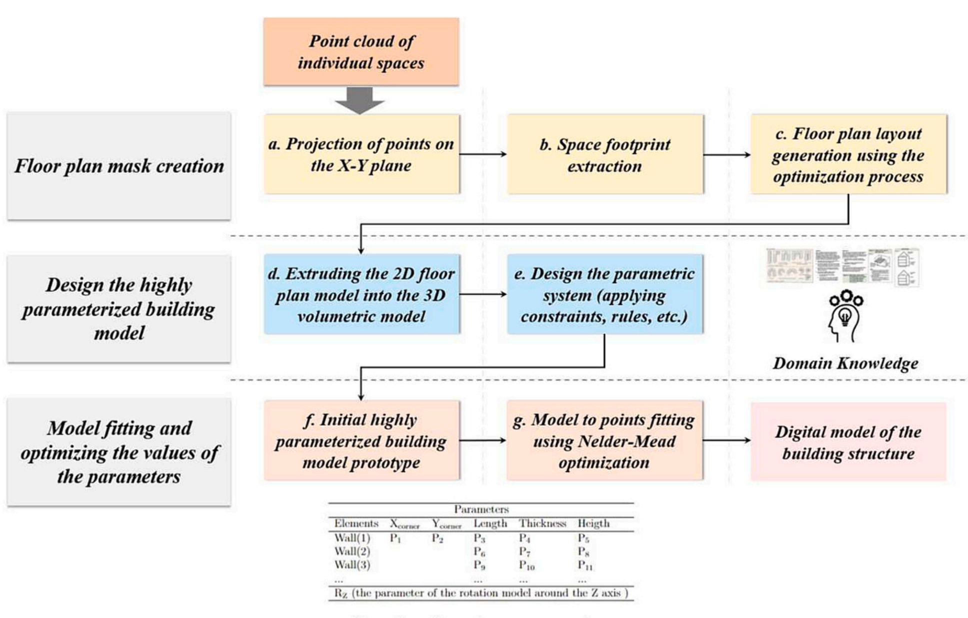


Fig. 13. The proposed workflow for geometry provision; the creation of the highly parametrised digital building model.

that are typically found in the specific type of buildings being examined. In the subsequent step, the remaining ceiling points undergo a density-based clustering procedure using the DBSCAN algorithm, which aims to generate distinct clusters [81]. DBSCAN is a well-suited technique for the clustering of sparsely distributed points as it identifies clusters as groups of points with high density within a particular range or neighborhood radius, which are separated by regions of lower density. Lastly, a hierarchical nearest neighbor approach is utilized to label each 3D point in the point cloud space, encompassing the ceiling, walls, floors, and furniture, with the appropriate cluster assignment obtained from the preceding clustering procedure. The pseudo-code of the proposed method for 3D space parsing is shown in Algorithm 1.

In indoor environments, individual spaces are typically separated by common walls or connected by openings such as doors and windows. To support various applications such as BIM, robotics, and path planning, an adjacency graph is constructed to represent the adjacency relationships among all spaces, regardless of whether they are physically connected. The adjacency graph, denoted as *G (V, E)*, consists of vertices (or nodes) that correspond to individual spaces, and edges that denote adjacency between two distinct spaces. To determine the adjacency relationships, the distance between the point clouds of each space instance is computed using a neighborhood tolerance *t*, which enables the identification of adjacent spaces in the adjacency graph.

Algorithm 1 3D space parsing of indoor point cloud

Input:
One point $\boldsymbol{p_S} \in \boldsymbol{S}$, laser-scanned point cloud set $\boldsymbol{S}$;
One point $\boldsymbol{p_i} \in \boldsymbol{C}$, Ceiling point cloud set $\boldsymbol{C}$;
Wall point cloud set $\boldsymbol{W}$;
The Ceiling to Wall points distance threshold $\boldsymbol{D}$;
The minimum number of points per space's cluster $\boldsymbol{M}$;

Function to calculate the distance of a query point $\boldsymbol{p_i}$ in $\boldsymbol{C}$ from all points of $\boldsymbol{W}$ in 3D space $PointToPointcloudDist()$;
Function for density-based clustering of remaining Ceiling point cloud set $\boldsymbol{C'}$ $DBSCAN()$;
Function to count the number of points in clusters $PointsInCluster()$;
Function to assign the cluster id of the remaining Ceiling point cloud set $\boldsymbol{C'}$ to the Laser-scanned point cloud set $\boldsymbol{S}$ using 2D nearest neighbour method $AssignLabelToPoints()$;

Initialize:
Matrix used to save remaining Ceiling point cloud set $\boldsymbol{C'}$;
Matrix used to save all labels of $DBSCAN()$ result $\boldsymbol{l}$;
Matrix used to save labels for all points in $\boldsymbol{S}$ point cloud set $\boldsymbol{L}$;

Algorithm:
```
for p_i ∈ C do
    if PointToPointcloudDist (p_i , W) < D
        remove p_i from C
    end if
end for
l = DBSCAN(C')
for all l' ∈ l do
    if PointsInCluster (l) <= M
        remove l' from l
    end if
end for
for p_S ∈ S do
    L(p_S)=AssignLabelToPoints(p_S,l)
end for
```

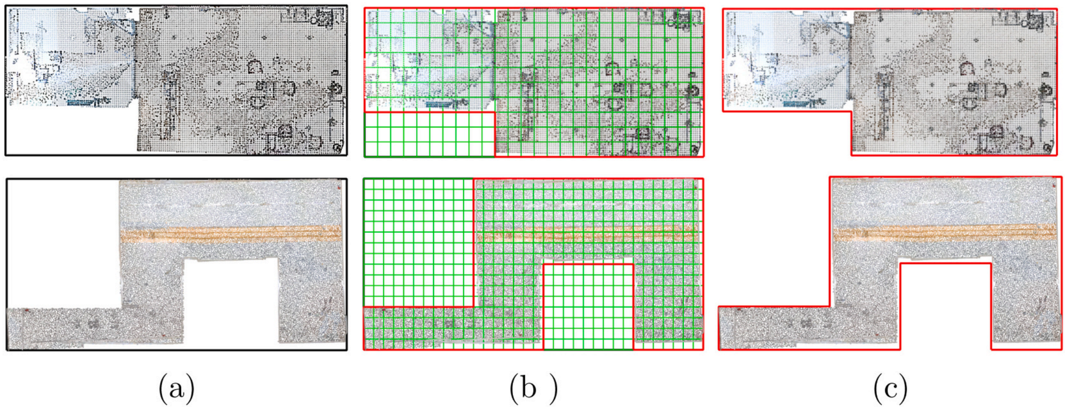


Fig. 14. The creation of the 2D geometric model of spaces with the parameter of *L* equal to 0.5: (a) fitting the 2D bounding box to the points in the X-Y plan, (b) install rectangular grid, identification, and removal of empty cells from the initial 2D bounding box model, (c) 2D geometric footprint model of spaces.

3.4. Geometry provision

Parametric modeling is the process of generating an interactive digital model that operates on a set of predetermined values, algorithms, and rules known as 'parameters'. The model is designed to have a flexible structure where different elements are interconnected through parametric rules, enabling any modification to the model to be automatically managed by internal logic arguments. When it comes to creating meaningful building objects and their relationships, traditional data-driven methods can encounter significant challenges [1,30].

However, a top-down approach that involves fitting a highly parameterized building model to observed data has emerged as a promising alternative which due to establishing parameters and rules that govern the relationships between building components, can create a more accurate and meaningful representation of the building. In the top-down approach, the building models are designed based on a typology of typical office and residential buildings, allowing for greater flexibility and adaptability to changes. This brings the potential to significantly advance the way of modeling building assets enabling a more effective management approach. Fig. 13 illustrates the proposed workflow for geometry provision.

3.4.1. Floor plan mask creation

In this research, the creation of the building digital model is handled through a top-down approach and the task of digital geometric representation is considered designing a highly parameterized floor plan problem. The difference among the floor plans is primarily ascribed to the dimensions and the layout of spaces in the 2D plane, in which the position of walls and slabs in building space play important roles. In this regard, the information extracted from the space parsing step is utilized to create an initial floor plan mask.

In the context of the Manhattan world, the spaces exhibit rectangular shapes that align with the coordinate axes. This phenomenon arises from the fact that the shortest path between two points within a Manhattan world is measured along the orthogonal axes, rather than through a straight line. This leads to a characteristic appearance of spaces that are "squared off" or "blocky" in nature. The aforementioned concept is harnessed via a model-based approach for designing a parametric floorplan mask (Fig. 14).

Initially, a 2D bounding box is fitted to the points of individual spaces in the X-Y plane to serve as a generalized model of the space's footprint, featuring low geometric intricacy. To enhance geometric detail in more complex spaces such as those with L-shaped configurations etc., the method of utilizing the bounding box grid is employed. This technique involves the installation of a rectangular grid with the dimensions of *L* on the points, followed by the identification and removal of empty cells from the initial 2D bounding box model through polygon subtracting operations. By adjusting the parameter *L*, the level of detail in the final model can be tailored to the user's desired accuracy, in which smaller L values lead to increased geometric detail but higher processing times.

The creation of building floor plans and the generation of its 3D structural models from the point cloud dataset have a formidable

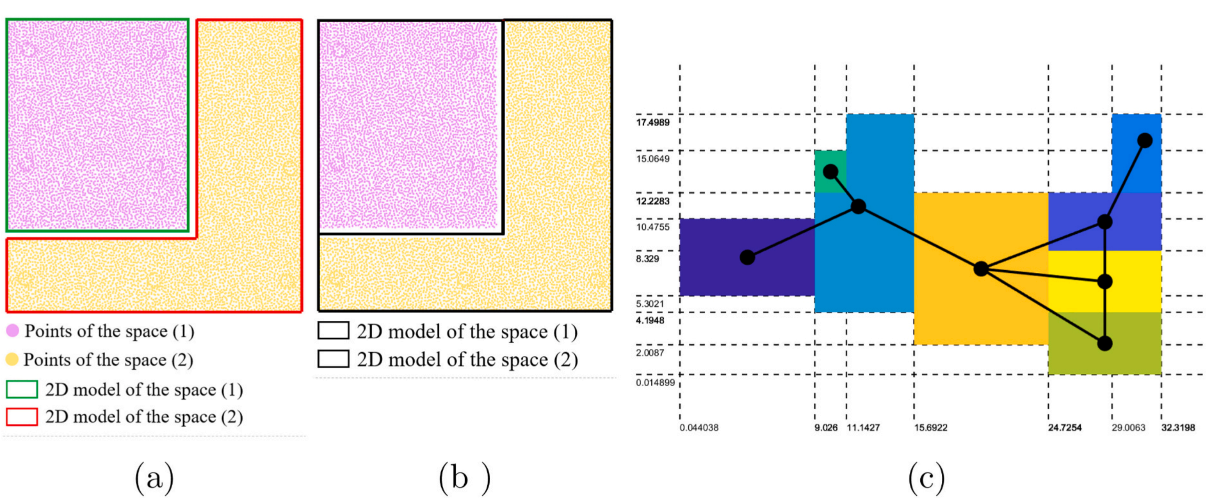


Fig. 15. Generation of the 2D parametric floor plan mask of the building layout using the simplex Nelder–Mead optimization algorithm: (a) 2D model of individual spaces, (b) finding and placement of shared walls between spaces using optimization problem, (c) parametric floor plan mask model of the building layout.

challenge in identifying the position of common walls shared between spaces. Existing methods rely on a variety of assumptions and multiple distance and angle thresholds to locate parallel candidate planes that may belong to common walls. They are with significant challenges in accurately identifying these common walls due to the complexity of indoor scenes and the heterogeneity of building layouts and designs.

To address the challenges, we define the process of determining the position of common walls shared between spaces and the creation of the parametric floor plan as an optimization problem (Fig. 15). Each 2D model of the individual spaces has two parameters X_c and Y_c values, which are utilized for the precise placement of these models within the original point cloud. The optimal placement of 2D models in the point cloud should aim to maximize the number of points from the corresponding space that are contained within the boundaries of the models.

In this regard, the simplex Nelder–Mead optimization algorithm [82] is utilized to identify the best values of the X_c and Y_c parameters for each space's 2D model by minimizing the objective function (1). During the optimization process, an important internal operation involves the selection of adjacent edges in the 2D models of adjacent spaces as shared walls. This is achieved by utilizing the adjacency graph of spaces and replacing the selected edges with their respective middle line. This process is contingent upon the verification of the objective function (2) and is applied only if it does not lead to an increase in the value of the objective function (1). By replacing the adjacent edges with their respective middle line, the algorithm seeks to improve the overall quality of the solution while preserving the geometric and topological properties of the original layout. This approach provides a viable means of optimizing the layout of spaces and improving their functionality in various architectural and design contexts.

$$Obj_1 = min \sum_{k=1}^{n} \left(P_{space(k)} - P_{model(k)} \right) \tag{9}$$

$$Obj_2 = \textbf{\textit{select}}\ e_i\ from\ \{e_1, \ldots, e_n\} \in S_1 \quad and \quad e_j\ from\ \{e_1, \ldots, e_m\} \in S_2 \tag{10}$$

Where k is the number of individual spaces, $P_{space(k)}$ is the number of points belonging to the space k, $P_{model(k)}$ is the number of points belonging to the space k within the boundaries of the corresponding model, e_i, e_j are the selected adjacent edges from the 2D models of two adjacent spaces, and $\{e_1,\ldots,e_n\}$, $\{e_1,\ldots,e_m\}$ are all edges of the 2D models of two adjacent spaces S_1, S_2 respectively.

Fig. 17. The process of fitting the highly parameterized digital model to the point cloud dataset (the ceiling points were removed to improve the visual representation).

3.4.2. Parametric design and model fitting

Upon generating the initial floor plan mask, a sophisticated system of geometrical and mathematical rules and constraints is employed to establish internal relationships between the walls and slabs. These rules and constraints are rooted in the principles of building design and construction, and aim to optimize the functionality and aesthetics of the structure. For instance, buildings intended for office or educational purposes are often designed in adherence to the Manhattan world, a layout that prioritizes practicality and ease of navigation. As a result, in the case of such buildings, a restriction of perpendicularity and parallelity between walls and slabs is imposed. The output of parametrization is a digital model in which the rules and restrictions are designed and applied so that any change in the internal parameters of a wall and slabs, such as the length and height, etc., will affect other related walls and elements (Fig. 16).

The level of geometric accuracy needed for a digital model in facility management can vary depending on the specific use case and requirements of the facility. A high level of geometric accuracy is desirable to ensure the model accurately represents the physical environment. Despite reliable topology, the initial parametric model may have inadequate geometric accuracy and imprecise parameter values. To address the problem, the Nelder-Mead optimization algorithm is employed once more to refine the rough model and bring it in closer alignment with the point cloud data and the real world (Fig. 17). To fit the volumetric model to the main structural element point cloud, the objective function of the

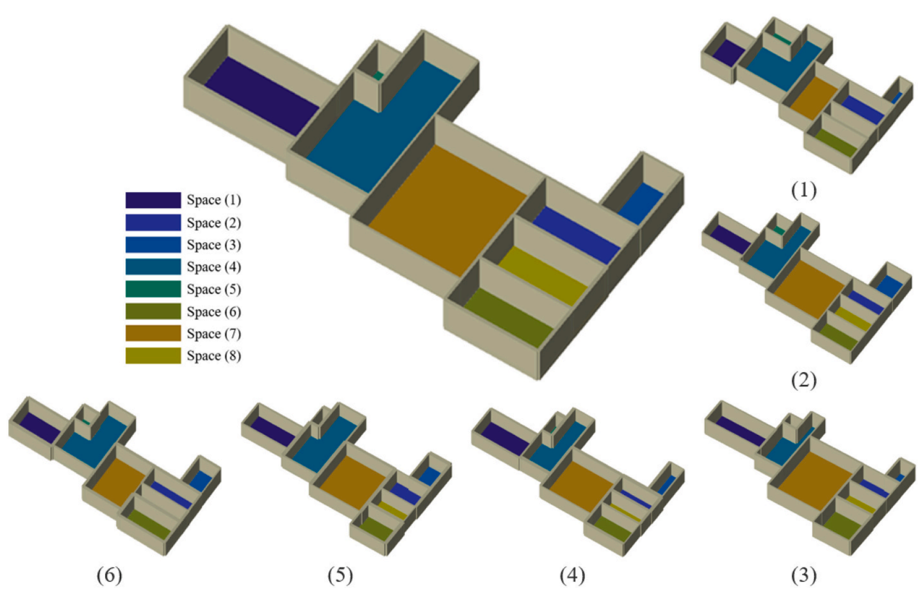


Fig. 16. Designing the highly parameterized digital model of the building's structure; the process of changing the parameters.

Table 2
Encoding Unknown parameters of highly parameterized digital model of the building's structure.

	Parameters				
Elements	X_{corner}	Y_{corner}	Length	Thickness	Heigth
Wall(1)	P_1	P_2	P_3	P_4	P_5
Wall(2)			P_6	P_7	P_8
Wall(3)			P_9	P_{10}	P_{11}
…			…	…	…

R_Z (the parameter of the rotation model around the Z axis).

Points-To-Model distance is considered. The process entails computing the distance between points and the planes of the model to reconstruct precise geometric models. In this regard, a lower Points-To-Model distance value indicates better adaptation of the digital model to the point cloud data, resulting in the estimation of parameters more accurately toward their actual values.

One of the major technical challenges in the creation of digital models of indoor environments is accurately estimating the thickness of walls. This is particularly challenging due to the complexity of building geometries and the need to account for variations in wall thickness throughout the structure. The proposed method incorporates the thickness values of walls into the optimization problem as unknown parameters, treating the walls as boxes with varying dimensions of the thickness (t), length (l), and height (h) during the model fitting process which enables more flexible and accurate estimation of wall parameters.

The design and optimization of parametric models for buildings involve a significant number of degrees of freedom which account for various geometric properties of the walls, including their length, thickness, height, and 2D location (Table 2). To estimate these unknown

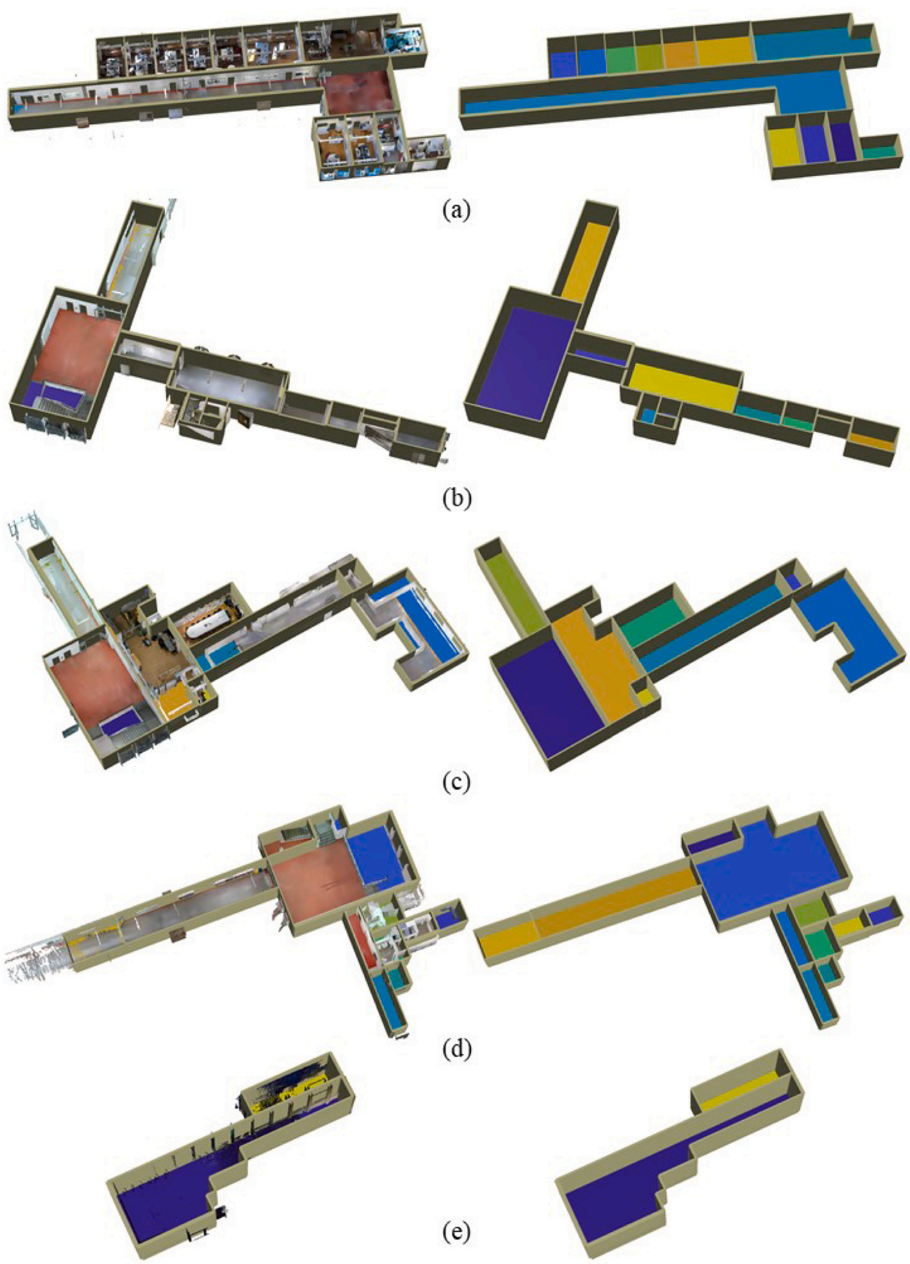


Fig. 18. Overview of the datasets and corresponding reconstructed digital models; (a) TUM CMS chair, (b) TUM - Floor (2), (c) TUM - Floor (3), (d) TUM - Floor (4), (e) TUM - Floor (5).

parameters accurately, they are included in the optimization problem as both explicit and implicit mathematical equations, subject to geometric constraints. The specific equations and values utilized in the optimization problem can vary depending on the case study and the particular building being modeled. In the next section, the efficacy and dependability of the proposed method are scrutinized through tests conducted on different case studies which offer insights into the performance of the proposed method and its potential applications in the domains of operation and management.

4. Implementation and result

4.1. Case study

This research study presents a comprehensive evaluation of the proposed method for the creation of the building digital model by utilizing distinct building datasets with varying architectural designs and layout appearances. The proposed approach is implemented in Python and MATLAB on a research computer (11th Gen Intel(R) Core(TM) i7-1165G7, with 16.0 GB memory). The evaluation metrics encompass various aspects of the proposed method, including its accuracy, efficiency, and scalability, thereby offering a comprehensive analysis of its potential for practical implementation.

The investigation involves the analysis of five distinct point cloud datasets, including two from the Technical University of Munich (TUM) in Germany and three from Stanford University in the United States (S3DIC) (Figs. 18–19). These datasets for Germany's sites have been captured by NavVis VLX laser scanner (www.navvis.com) and the US datasets were generated by Matterport scanner (www.matterport.com). The buildings represented in the datasets are primarily utilized for educational and research purposes, featuring various areas such as rooms, offices, hallways, and stairwells. A detailed summary of the input data's pertinent properties, including dimensional information, the number of points etc. are provided in Table 3.

Table 3
Data used in this research.

Datasets	Length (m)	Width (m)	Number of points	Clutter
TUM CMS chair	25.51	59.73	534.136	Low
TUM - Floor (2)	34.42	47.70	683.586	Moderate
TUM - Floor (3)	44.67	48.62	675.294	Moderate
TUM - Floor (4)	51.78	27.80	551.092	Moderate
TUM - Floor (5)	24.31	10.22	88.625	Moderate
S3DIC Area (1)	24.63	48.26	1.120.758	Low
S3DIC Area (3)	29.23	25.71	484.874	Low
S3DIC Area (6)	23.16	45.34	1.059.076	Low

4.2. Preprocessing

To facilitate the execution of the main steps of the proposed algorithm, the datasets undergo preprocessing procedures prior to their utilization using the steps described in Section 3.2. In this regard, a grid-based subsampling approach with a 5 cm length is utilized to simplify dense point clouds. Additionally, the CCS method with a distance threshold of 0.25 m is employed to eliminate noise and outlier data in the point cloud space and separate the main building part from the raw point cloud. The TUM building (1) datasets is a multi-story building consisting of four floors. In order to separate the points belonging to each level, the proposed data-driven method is used (Fig. 6). First, the verticality and normal vector features are calculated for each point with a spherical neighborhood with a radius of 0.25 m. Next, in order to separate the major horizontal segments including the ceiling and floor

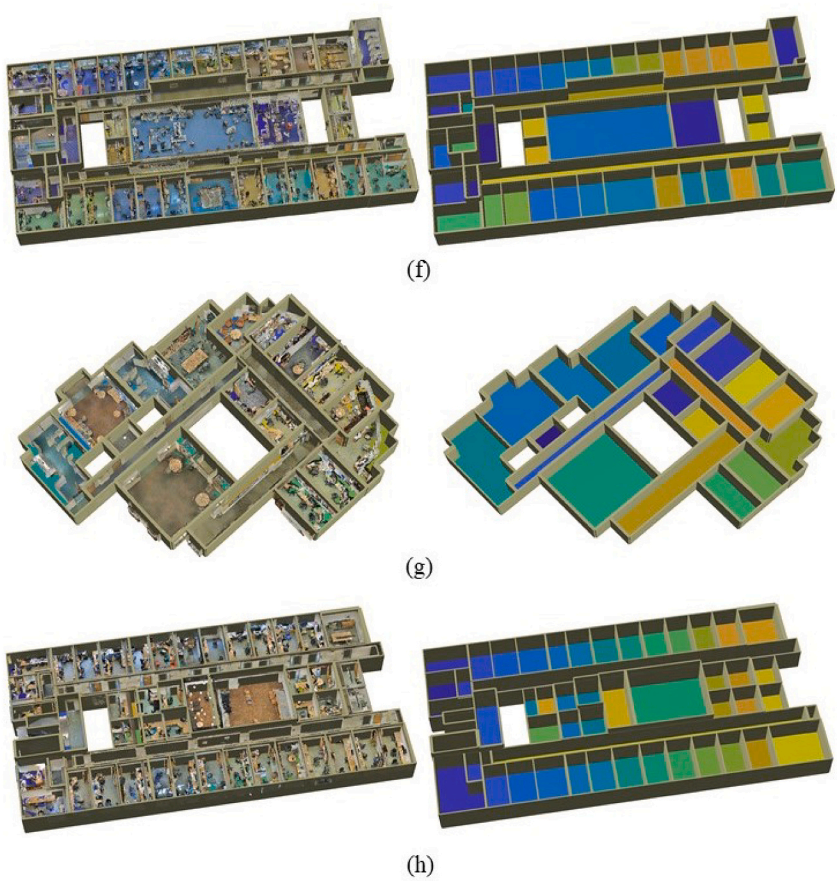


Fig. 19. Overview of the datasets and corresponding reconstructed digital models; (f) S3DIC Area 1, (g) S3DIC Area 3, (h) S3DIC Area 6.

Table 4
Accuracy evaluation of the point cloud classification (Precision, and Recall metrics).

	Precision				Recall			
Datasets	Ceiling	Floor	Furniture	Wall	Ceiling	Floor	Furniture	Wall
TUM CMS chair	0.98	0.95	0.39	0.88	1.00	0.98	0.34	0.88
TUM - Floor (2)	0.88	0.95	0.29	0.95	0.99	0.74	0.38	0.95
TUM - Floor (3)	0.78	0.81	0.26	0.75	0.95	0.85	0.48	0.42
TUM - Floor (4)	0.80	0.89	0.16	0.87	0.99	0.79	0.24	0.77
TUM - Floor (5)	0.89	0.89	0.34	0.98	0.98	0.94	0.58	0.88
S3DIC Area (1)	0.86	0.80	0.59	0.84	0.93	0.91	0.51	0.83
S3DIC Area (3)	0.89	0.85	0.64	0.76	0.94	0.92	0.46	0.86
S3DIC Area (6)	0.84	0.78	0.62	0.82	0.92	0.91	0.51	0.82

Table 5
Accuracy evaluation of the point cloud classification (F-Score, and overall accuracy metrics).

	F-Score				Acc.
Datasets	Ceiling	Floor	Furniture	Wall	Overall
TUM CMS chair	0.99	0.97	0.36	0.88	0.88
TUM - Floor (2)	0.93	0.83	0.33	0.95	0.88
TUM - Floor (3)	0.86	0.83	0.33	0.54	0.64
TUM - Floor (4)	0.89	0.84	0.20	0.81	0.76
TUM - Floor (5)	0.93	0.91	0.43	0.94	0.89
S3DIC Area (1)	0.90	0.85	0.55	0.83	0.79
S3DIC Area (3)	0.91	0.88	0.53	0.81	0.78
S3DIC Area (6)	0.88	0.84	0.56	0.82	0.78

points, the CCS method with a distance threshold value of 0.25 m is used. Subsequently, after classifying the candidate points into classes of ceiling and floor using normal vector directions the derived information is used to separate the individual level's points.

4.3. Experimental results of point cloud semantic enrichment

The evaluation of the point cloud semantic enrichment approaches is involved two main steps of points classification using the decision tree classification model and 3D space parsing. For this purpose, the required features including Verticality, Planarity, Normal Z, and Normal change rate are calculated for each point in a 0.25 m neighborhood. Next, the pre-trained Decision Tree classifier model is applied to the datasets, and the main structural elements of the building are extracted. To evaluate the performance of the classification step, reference data are manually generated for each building point cloud and the comparison results between the reference data and the classified points are reported in Tables 4–5.

For each dataset, the standard quality metrics of precision, recall, and F-score are calculated using Eqs. (8)–(10). According to the results, the average accuracy of classification for all buildings datasets is about 80%. As previously mentioned, the objective of the classification step is to identify and extract the main structural elements of the building, such as ceilings, and walls, and utilize them in the proposed knowledge-based method for 3D space parsing. According to the result provided in Table 5, the accuracy of classifying walls in building environments is lower compared to ceilings and floors, reasoned by factors such as scene complexity and geometric feature similarity between wall points and furniture elements (Fig. 20). Addressing this issue requires the preparation of huge and diverse datasets while collecting such data for indoor environments is challenging and expensive. Therefore, leveraging existing knowledge of building design and element-space interaction is important for feature extraction and modeling tasks.

To assess the performance of the method used for point cloud classification, a quantitative comparison is conducted between the Decision Tree method and other supervised learning methods, including GaussianNB (Gaussian Naive Bayes) [83] and SVM (Support Vector Machine) [84]. Each model is trained with annotated S3DIC area 5 data and is tested on each building datasets.

GaussianNB utilizes the Gaussian distribution to model feature dependencies and calculates the probability of an instance belonging to a specific class based on its feature values and the information gained from the training data. SVM classifies data by finding the optimal hyperplanes that maximally separate different classes in the feature space. Due to the nonlinear and complex relationship between normalized geometric features, the sigmoid kernel is used for training the SVM classifier model.

According to Table 6, the Decision Tree model has achieved higher accuracy compared to other classification methods. While GaussianNB

Table 6
Quantitative comparison of the overall accuracy of various supervised learning methods in the classification of indoor point clouds.

Supervised learning methods	Decision Tree	GaussianNB	SVM
Overall accuracy	0.80	0.76	0.57

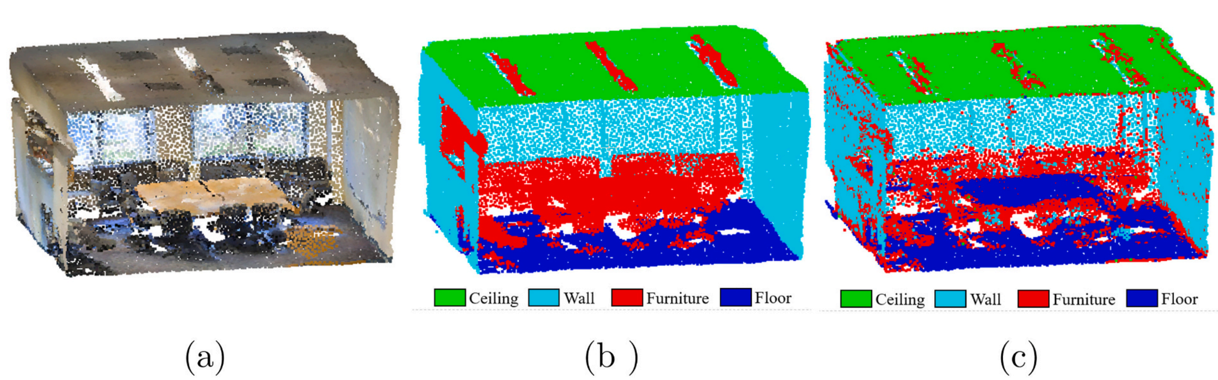


Fig. 20. The result of points classification: conference room of the Stanford University Area (1) dataset: (a) original point cloud, (b) ground truth for the points classification, (c) the result of points classification using the Decision Tree classification model.

demonstrates strong performance of 76%, with only a 4% difference in overall accuracy compared to the Decision Tree model, SVM has reached the lowest accuracy at about 57%. Although the Decision Tree and GaussianNB handle data complexity differently, due to the nature of the SVM learning process, it is more sensitive to data imbalance. A substantial proportion of the training data corresponds to structural elements of walls, ceilings, and floors, posing challenges for the SVM model in capturing different nonlinear relationships and distinguishing between the classes of furniture and structural elements. Furthermore, despite the use of various distinct geometric features, there are still similarities between the geometric feature values of wall and ceiling elements and other furniture objects such as bookshelves and desks. These similarities can lead to more confusion in the model's learning process.

To disjoint 3D spaces, the ceiling points situated 35 cm away from the wall element are eliminated. The distance threshold of 35 cm is selected based on the standard thickness of interior and exterior walls in educational building design using materials such as concrete, stone slabs, etc. Subsequently, the DBSCAN clustering technique is applied along with the nearest neighbor algorithm to group points in the 3D space into distinct clusters. This approach is employed to identify areas of dense point concentration in the 3D space and to create boundaries between these clusters, thus forming separate and disjointed 3D spaces. For determining the adjacency relationships between spaces in accordance with the indoor layout, the distances between all segmented spaces are calculated.

A neighborhood distance of 1 m is employed as a tolerance threshold, whereby any spaces within this distance are deemed to be adjacent. To assess the effectiveness of the proposed algorithm for 3D space parsing, a quantitative evaluation is conducted which involves comparing the statistical parameters of the algorithm's results with the information provided by the facilities department (Table 7). The evaluation metrics include the unsupervised clustering similarity metric Rand Index (RI) [85] and the number of individual spaces identified by the algorithm. In this regard, the overall accuracy for 3D space parsing is 93.5% which indicates the utility and performance of the proposed algorithm for partitioning 3D spaces in different indoor environments with diverse layouts and design.

In the realm of 3D space parsing for large-scale and complex indoor environments, a key obstacle is the accurate separation of hallways from each other and mitigating the issue of over-segmentation. The proposed knowledge-based algorithm effectively addresses this challenge by seamlessly grouping and distinguishing individual hallways and corridors as a unified space, without requiring any post-processing tasks (Fig. 21). The principal objective of creating digital models utilizing BIM concepts is to furnish both geometric and semantic data about an object in a simultaneous manner. In this regard, the segmented 3D spaces and corresponding wall instances can be utilized across various applications, including scan-to-BIM, 3D modeling, and navigation.

4.4. Experimental results of geometry provision

To design the highly parameterized building digital models, first, the 2D floor plans that reflect the building layouts are generated. Individual space points are initially projected on the X-Y plane, and their corresponding 2D models are extracted using the model-based approach outlined in Section 3.4.1. To compare the geometric detail in result models, the parameter of grid dimension in the corresponding model-based approach is considered with different values of 0.25, 0.5 and 1 m. The extracted 2D models are then imported into the optimization process to determine their optimal placement within the point cloud space, thereby generating comprehensive floor plans for the datasets.

Table 7
Accuracy evaluation of 3D space parsing.

Datasets	Number of spaces: Ground truth	Number of spaces: Our method	Rand Index%
TUM CMS chair	15	12	0.93
TUM - Floor (2)	9	11	0.85
TUM - Floor (3)	9	9	0.89
TUM - Floor (4)	9	11	0.94
TUM - Floor (5)	2	2	0.97
S3DIC Area (1)	44	45	0.97
S3DIC Area (3)	23	19	0.95
S3DIC Area (6)	48	51	0.98

In this regard, the geometric centre of the point cloud pertaining to each individual space is considered as the initial value for unknown parameters in the optimization process. This serves as a starting point for the iterative optimization algorithm, which aims to identify the most suitable spatial arrangement of the 2D models within the point cloud. Due to iteratively refining the placement of spaces based on predefined objective functions and constraints, the optimization process effectively generates cohesive and well-organized floor plans that accurately represent the spatial relationships and layout of the datasets.

Subsequently, the initial floor plan models are extended into 3D volumetric-parametric models and subjected to geometrical and mathematical rules and constraints, including the requirement for perpendicularity between elements such as walls, slabs, and ceilings. Finally, the highly parameterized building models are fitted to the point cloud through the optimization process and the best value for the parameters of the building model are extracted. Table 8 shows the values of parameters of optimization processes for creating highly parameterized building models. One of the notable challenges in the accurate creation of the digital model of an indoor environment lies in estimating the thickness of exterior walls.

During the scanning process, these walls appear as single planar surfaces, and the optimization process typically yields thickness values of around 1–2 cm. To address this issue, a modification is implemented to the attributes of these walls and their thickness is considered with the minimum value observed in the thickness of the shared walls within the respective models.

In order to evaluate the accuracy of the proposed method for creating digital building models, three evaluation metrics of the precision and consistency across different aspects, including geometry, semantics, and topology extents are examined. These metrics are calculated by comparing the reconstructed digital models with reference BIM models (Fig. 22).

To assess the accuracy of the geometric representation and semantic properties within the digital models, a comparative analysis is conducted between the parameters of reconstructed models and those of reference models. Initially, the corresponding walls and ceiling elements in both the reconstructed and reference models are identified by utilizing the coordinates of their corresponding endpoints. For each individual wall instance in the reference model, three rectangular buffers with the dimensional thresholds 5, 10 and 20 cm are considered and the corresponding closest wall instance in the reconstructed model is selected. For each model, the accuracy of matching between elements in the reference model and reconstructed model is measured by Eq. (11).

$$Matching\text{-}score = \frac{\sum_{k=1}^{m}(Rec_b)}{\sum_{k=1}^{n}(Ref)} \tag{11}$$

where the $\sum_{k=1}^{n}(Ref)$ is the number of elements in the digital reference model and $\sum_{k=1}^{m}(Rec_b)$ is the corresponding elements in the reconstructed digital model extracted using buffer with the threshold of b.

Subsequently, for each paired element, the disparities in parameters such as length, height, thickness, and endpoint displacement are quantified. This aids in assessing that in which extent the proposed methodology accurately presents the spatial information of the environment, while also demonstrating the consistency in semantic properties assigned to various building components. Tables 9–11 present the statistical evaluation results using the standard measure of mean errors.

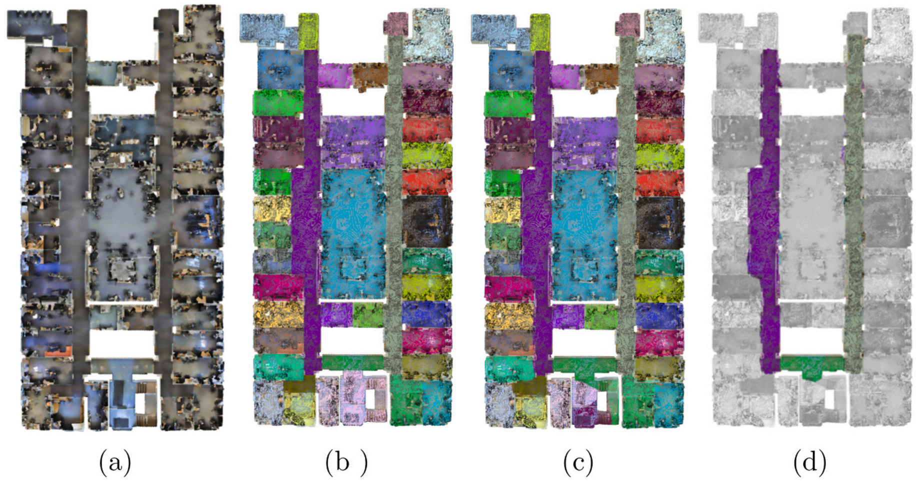


Fig. 21. The result of 3D space parsing on Stanford University datasets Area (1) (bottom view): (a) original point cloud, (b) ground truth for 3D space parsing, (c) the result of the proposed method for 3D space parsing, (d) separation of individual hallways.

Table 8
The initial values of parameters for optimization processes.

Parameters			
Problem	Tolerance_X	Tolerance_obj	Iterations
Floor plan generation	0.0001	0.0001	50
Volumetric digital model fitting	0.0001	0.0001	50

In this regard, the average accuracy of about 0.08 m in differences between the coordinate of elements endpoints and 0.06 m for estimating the parameters of elements highlights the efficiency of the proposed method in the creation of volumetric-parametric digital models for buildings with diverse designs and layouts from the single floor buildings to the multi-story scenario (Fig. 23). According to the results, the highest error is related to S3DIC Area (3) building because the data contain some rooms in non-manhattan shape and make the comparison between the reference model and the reconstructed digital models less precise.

5. Discussion

In the realm of built environment digital twinning, the principles rest upon the precise reflection of real-world assets and the establishment of a synergy between physical and virtual realities. In this regard, parametric-volumetric digital models play a pivotal role in the development of DTs, offering dynamic representations, capturing intrinsic behaviors and relationships, and facilitating comprehensive simulations. Recently, remote sensing data, including point clouds, has emerged as a primary tool for unveiling a nuanced understanding of physical environments and creating digital models. However, developing automated algorithms to convert raw remote sensing data into consumable digital models for human and machine interpretation still remains a challenging endeavor. This paper presents a novel automatic algorithm for the creation highly parameterised digital building model with rich semantics and coherent geometry. The main idea of the proposed workflow is to combine the capabilities of bottom-up, and top-down approaches to overcome the limitations of common traditional methods in point cloud processing and model representation. The hybrid approach aligns the capabilities of AI methods in scene understanding with domain knowledge to improve automation and efficiency in the face of all real-world building designs.

5.1. Comparison with other methods

To assess the effectiveness of the proposed method in creating digital models comparing with other existing algorithms, a quantitative comparison is conducted between our proposed parametric modeling method and the 3D reconstruction method proposed by [32]. For implementation, the points belonging to the walls are first extracted

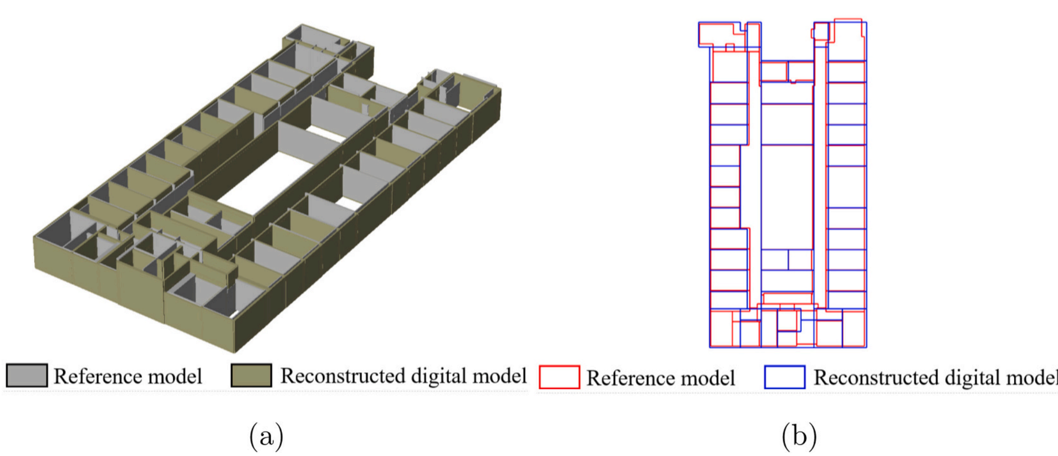


Fig. 22. The comparison between the reconstructed digital model and the reference BIM model of Stanford University datasets Area (1): (a) the evaluation of volumetric parameters, (b) the evaluation of wall coordinates.

Table 9
Accuracy evaluation of digital model reconstruction (grid dimension = 0.25m).

$Buffer_{th}$	*matching*	δ displacement	δ length	δ thickness	δ height
TUM CMS chair					
5 cm	78%	0.04 m	0.05 m	0.04 m	0.03 m
10 cm	83%	0.06 m	0.07 m	0.04 m	0.04 m
20 cm	95%	0.09 m	0.08 m	0.05 m	0.04 m
TUM - Floor (2)					
5 cm	83%	0.05 m	0.06 m	0.05 m	0.07 m
10 cm	89%	0.07 m	0.08 m	0.06 m	0.06 m
20 cm	96%	0.08 m	0.08 m	0.05 m	0.06 m
TUM - Floor (3)					
5 cm	76%	0.05 m	0.06 m	0.04 m	0.04 m
10 cm	83%	0.10 m	0.06 m	0.04 m	0.05 m
20 cm	88%	0.14 m	0.08 m	0.06 m	0.05 m
TUM - Floor (4)					
5 cm	69%	0.05 m	0.04 m	0.05 m	0.03 m
10 cm	75%	0.10 m	0.06 m	0.06 m	0.03 m
20 cm	89%	0.14 m	0.07 m	0.05 m	0.03 m
TUM - Floor (5)					
5 cm	100%	0.03 m	0.02 m	0.02 m	0.01 m
10 cm	100%	0.03 m	0.02 m	0.02 m	0.01 m
20 cm	100%	0.03 m	0.02 m	0.02 m	0.01 m
S3DIC Area (1)					
5 cm	73%	0.05 m	0.05 m	0.05 m	0.05 m
10 cm	80%	0.09 m	0.07 m	0.05 m	0.06 m
20 cm	92%	0.14 m	0.11 m	0.07 m	0.06 m
S3DIC Area (3)					
5 cm	69%	0.05 m	0.06 m	0.04 m	0.03 m
10 cm	78%	0.10 m	0.07 m	0.04 m	0.03 m
20 cm	88%	0.14 m	0.10 m	0.06 m	0.03 m
S3DIC Area (6)					
5 cm	74%	0.05 m	0.06 m	0.05 m	0.04 m
10 cm	87%	0.08 m	0.08 m	0.07 m	0.04 m
20 cm	93%	0.12 m	0.11 m	0.08 m	0.04 m

Table 10
Accuracy evaluation of digital model reconstruction (grid dimension = 0.5m).

$Buffer_{th}$	*matching*	δ displacement	δ length	δ thickness	δ height
TUM CMS chair					
5 cm	54%	0.05 m	0.06 m	0.06 m	0.03 m
10 cm	67%	0.07 m	0.06 m	0.08 m	0.04 m
20 cm	80%	0.11 m	0.09 m	0.08 m	0.04 m
TUM - Floor (2)					
5 cm	62%	0.04 m	0.06 m	0.04 m	0.03 m
10 cm	74%	0.09 m	0.08 m	0.05 m	0.03 m
20 cm	81%	0.15 m	0.09 m	0.05 m	0.03 m
TUM - Floor (3)					
5 cm	57%	0.05 m	0.05 m	0.06 m	0.07 m
10 cm	66%	0.10 m	0.07 m	0.07 m	0.07 m
20 cm	74%	0.12 m	0.09 m	0.09 m	0.08 m
TUM - Floor (4)					
5 cm	59%	0.04 m	0.05 m	0.04 m	0.03 m
10 cm	67%	0.07 m	0.06 m	0.07 m	0.03 m
20 cm	71%	0.12 m	0.09 m	0.08 m	0.04 m
TUM - Floor (5)					
5 cm	100%	0.03 m	0.03 m	0.02 m	0.01 m
10 cm	100%	0.04 m	0.03 m	0.02 m	0.01 m
20 cm	100%	0.04 m	0.03 m	0.02 m	0.01 m
S3DIC Area (1)					
5 cm	67%	0.05 m	0.06 m	0.04 m	0.03 m
10 cm	75%	0.09 m	0.07 m	0.05 m	0.02 m
20 cm	89%	0.15 m	0.09 m	0.05 m	0.03 m
S3DIC Area (3)					
5 cm	49%	0.05 m	0.07 m	0.06 m	0.03 m
10 cm	64%	0.10 m	0.09 m	0.08 m	0.04 m
20 cm	72%	0.16 m	0.14 m	0.09 m	0.04 m
S3DIC Area (6)					
5 cm	57%	0.05 m	0.04 m	0.05 m	0.04 m
10 cm	68%	0.09 m	0.07 m	0.04 m	0.05 m
20 cm	77%	0.14 m	0.10 m	0.04 m	0.05 m

Table 11
Accuracy evaluation of digital model reconstruction (grid dimension = 1m).

$Buffer_{th}$	*matching*	δ displacement	δ length	δ thickness	δ height
TUM CMS chair					
5 cm	31%	0.05 m	0.08 m	0.06 m	0.03 m
10 cm	47%	0.10 m	0.11 m	0.06 m	0.05 m
20 cm	69%	0.18 m	0.15 m	0.08 m	0.05 m
TUM - Floor (2)					
5 cm	36%	0.04 m	0.10 m	0.05 m	0.02 m
10 cm	45%	0.09 m	0.14 m	0.07 m	0.03 m
20 cm	74%	0.17 m	0.21 m	0.07 m	0.03 m
TUM - Floor (3)					
5 cm	41%	0.05 m	0.09 m	0.06 m	0.07 m
10 cm	55%	0.10 m	0.14 m	0.06 m	0.07 m
20 cm	69%	0.19 m	0.18 m	0.08 m	0.08 m
TUM - Floor (4)					
5 cm	48%	0.04 m	0.08 m	0.04 m	0.04 m
10 cm	57%	0.09 m	0.13 m	0.05 m	0.05 m
20 cm	66%	0.16 m	0.19 m	0.05 m	0.05 m
TUM - Floor (5)					
5 cm	100%	0.05 m	0.04 m	0.03 m	0.01 m
10 cm	100%	0.04 m	0.05 m	0.03 m	0.01 m
20 cm	100%	0.04 m	0.05 m	0.03 m	0.01 m
S3DIC Area (1)					
5 cm	38%	0.03 m	0.14 m	0.05 m	0.03 m
10 cm	40%	0.08 m	0.14 m	0.07 m	0.02 m
20 cm	76%	0.16 m	0.14 m	0.08 m	0.03 m
S3DIC Area (3)					
5 cm	24%	0.06 m	0.26 m	0.10 m	0.04 m
10 cm	41%	0.08 m	0.23 m	0.07 m	0.03 m
20 cm	66%	0.13 m	0.21 m	0.10 m	0.03 m
S3DIC Area (6)					
5 cm	36%	0.04 m	0.25 m	0.02 m	0.04 m
10 cm	60%	0.07 m	0.23 m	0.04 m	0.03 m
20 cm	71%	0.15 m	0.23 m	0.06 m	0.03 m

using the classification model. Subsequently, the RANSAC plane fitting algorithm is employed along with the DBSCAN clustering method to separate points of individual wall instances based on their orientation. Finally, the Manhattan 3D box estimation method is utilized to generate the 3D model of the walls. For comparison, the resulting digital models are once again compared with the reference digital model with respect to the criteria outlined in Section 4.4.

As can be seen in Table 12, our proposed method achieved precise accuracy compared to the 3D reconstruction method proposed by [32], in term of estimating the parameters of the wall instances. One of the challenges of proposed method in [32] is the utilization of data-driven 3D box estimation to extract the parameters of the walls which is highly sensitive to the quality of the input data, as well as noise and clutter points. It also has some limitations in considering various distance threshold values for clustering the points of walls, which can impact the estimation of thickness for shared wall instances. In contrast, our proposed parametric modeling approach benefits from the optimization process and can overcome common challenges in point cloud processing.

Table 13 provides comparison between the key features of proposed method and the most recent approaches [28–30,32,34,79,86] in creating semantic digital models of buildings. Each column of the table present a key feature and the potential of the proposed method in creation semantic digital models.

As can be seen in Table 13, most of the developed methods can create volumetric models of walls and spaces. However, among these methods, only one is capable of accurately representing volumetric models and employing topological relationships between elements and spaces during the reconstruction process. Furthermore, the majority of the proposed modeling methods rely on non-parametric data-driven approaches. As mentioned earlier, these methods encounter significant challenges when dealing with common challenges in point cloud data, as well as in inferring and simulating topological relationships between the structural elements. These challenges pose obstacles to the automatic

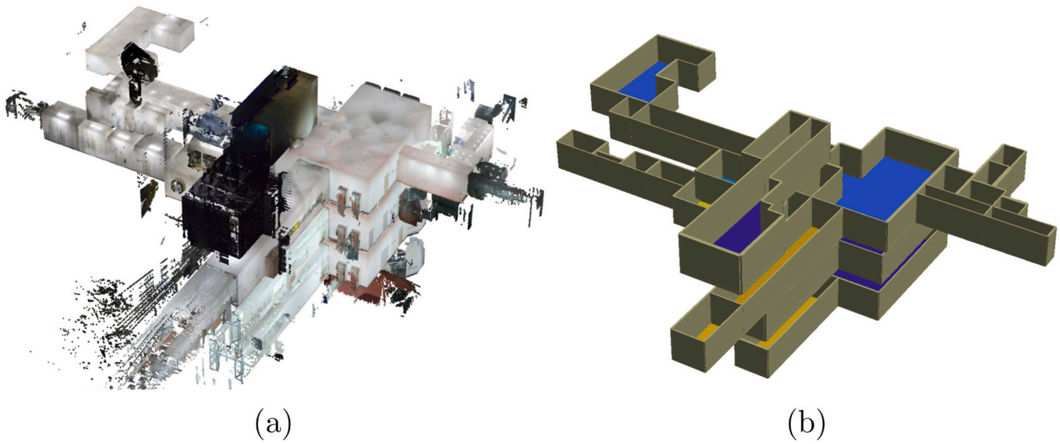


Fig. 23. Creation of the digital model of multi-level building scenario (TUM building (1) dataset): (a) original point cloud, (b) the highly parameterized digital model.

Table 12
Quantitative comparison of the results between the model reconstruction approach proposed by [32] and our proposed parametric modeling approach (grid dimension = 0.25m).

Method	δ displacement	δ length	δ thickness	δ height
[32]	0.06 m	0.11 m	0.08 m	0.05 m
Ours	0.08 m	0.06 m	0.05 m	0.04 m

Table 13
Comparison of key features of the proposed method with five state-of-the-art methods.

Method	Volumetric walls	Volumetric spaces	Topological relation	Parametric modeling
Ochmann et al. [28]	Yes	Yes	No	No
Nikoohemat et al. [29]	Yes	Yes	No	No
Tran and Khoshelham [30]	Yes	Yes	Yes	No
Cai and Fan [86]	No	Yes	No	No
Wu et al. [32]	Yes	Yes	No	No
Abdollahi et al. [34]	No	Yes	No	No
Pan et al. [79]	Yes	Yes	No	No
Ours	Yes	Yes	Yes	Yes

Table 14
Overview of the essential parameters and assumptions employed during the creation of a digital model using the proposed method.

Parameter	Value
Preprocessing:	
1. Grid-based point cloud subsampling distance	0.05 m
2. Distance threshold for noise and outliers removal	0.25 m
3. Neighborhood radius for geometric feature extraction	0.25 m
Semantic enrichment:	
4. Neighborhood radius for geometric feature extraction	0.25 m
5. Depth of Decision Tree model	10
6. Ceiling to Wall points distance threshold (space parsing)	0.35 m
7. Distance tolerance for the DBSCAN clustering	0.35 m
8. Neighborhood distance for space adjacency matrix	1 m
Assumptions:	
- The individual spaces are separated by common walls.	
- The minimum length and width of spaces are equal to 0.35 m.	
Geometry provision:	
9. Grid dimension for model-based space footprint extraction	0.25 m 0.5 m 1 m
Assumptions:	
- The building's space layout represents the Manhattan world.	
- Walls are parallel, and the slabs are perpendicular to the walls.	
- Ceiling and walls can have different height levels.	

reconstruction process of the semantic digital model.

In this regard, our proposed hybrid bottom-up, top-down method is capable of inferring topological relationships between elements and 3D spaces and employ them in a parametric modeling approach results in the creation of highly parameterized digital model with coherent geometry. In this context, each wall instance within the reconstructed digital models possesses its distinct height, while the shared walls between adjacent spaces exhibit varying face heights (Fig. 24). This characteristic enhances the overall fidelity of the reconstructed digital models, aligning them more closely with the physical reality of real-world structures, where buildings often feature distinct spaces with varying height differentials in their ceiling elements. Also, thanks to employing of the optimization method in the 3D reconstruction process, our proposed method is able to effectively overcome the challenges in

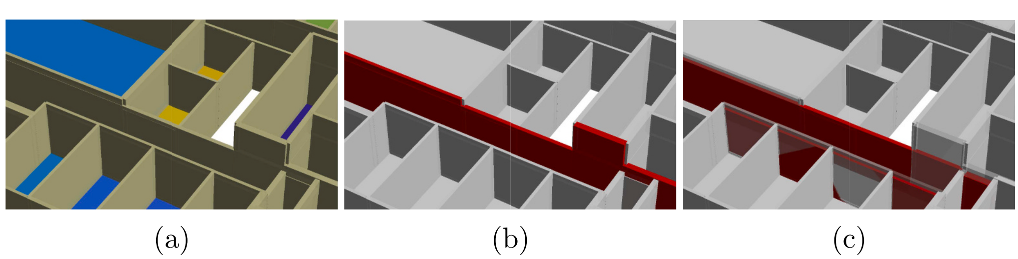


Fig. 24. Accurate representation of wall height in the reconstructed digital model: (a) the reconstructed digital model, (b) the difference of the height within the wall instances, (c) the difference of the height in different faces of the wall instances.

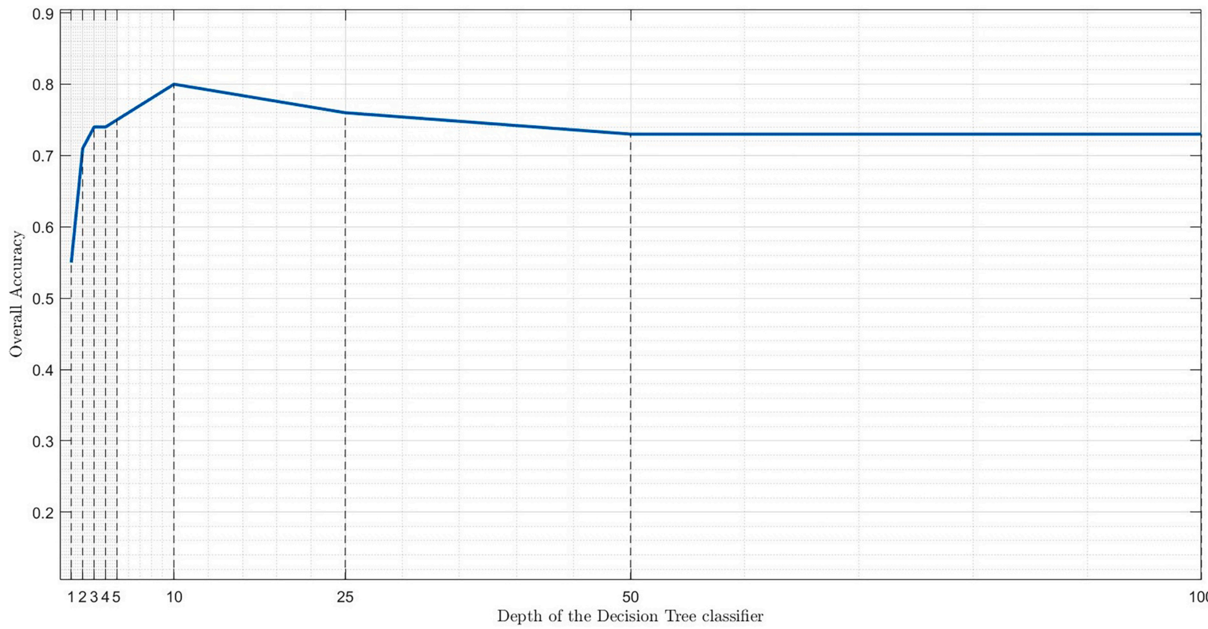


Fig. 25. Evaluating sensitivity of the proposed algorithm to the assigned parameter in semantic enrichment steps: the impact of the'Depth of Decision Tree model' parameter on the overall accuracy of test dataset classification.

point cloud data.

5.2. Sensitivity to the parameters

Table 14 presents an exhaustive enumeration of the diverse values assigned to parameters and assumptions employed in the creation of the digital models through the utilization of the proposed methodology. In order to assess the sensitivity of the developed algorithm to the assigned parameter values, a comprehensive evaluation is conducted by systematically testing various parameter values. This evaluation aim to investigate the impact of these parameters on the corresponding results at each step.

Within the framework of semantic enrichment steps, a quantitative assessment is conducted to investigate the impact of the depth parameter on the learning process of the decision tree classification model. This examination further explores the subsequent effects of the parameter on the overall accuracy in the classification of the test datasets. As illustrated in the Fig. 25, increasing the depth parameter value from one to ten results in a notable improvement in the overall classification accuracy, rising from 55% to 80%. However, the depth parameter value beyond ten leads to a gradual decline in overall accuracy, stabilizing at approximately 74%. These imply that, in accordance with the properties of the main structural elements within the indoor point cloud and the distinctive characteristics of each geometric feature calculated for element classification, the decision tree model achieves optimal performance at a depth of ten. This optimal depth allows for the proficient

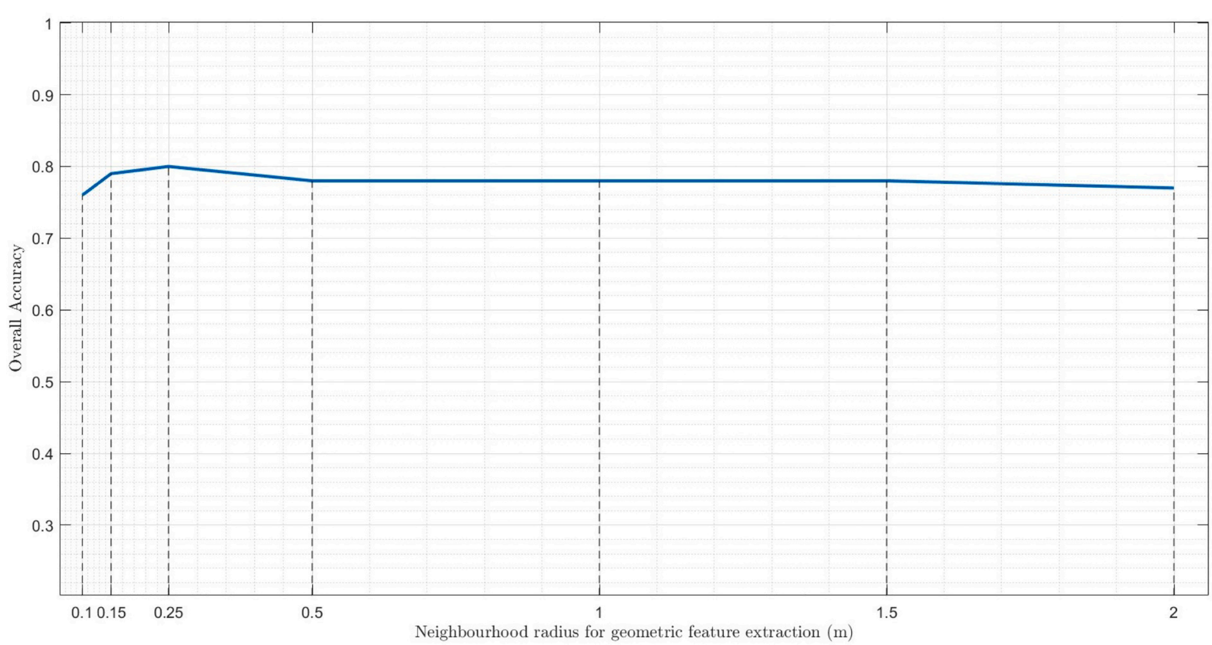


Fig. 26. Evaluating sensitivity of the proposed algorithm to the assigned parameter in semantic enrichment steps: the impact of the different Neighborhood radius for geometric feature extraction used in points classification using Decision Tree model.

categorization of selected geometric features, thereby facilitating the accurate classification of different classes based on their distinctive characteristics.

To investigate the affect of the different neighborhood radius values for geometric feature extraction on the overall accuracy of classification using the decision tree model, various radius values ranging from 10 cm to 200 cm are examined. According to the Fig. 26, employing a neighborhood radius of 25 cm yields the highest overall accuracy of about 80% in the point cloud classification task. However, the overall accuracy obtained with other selected radius values is consistently high, averaging at 78%. The most important factor in this context is the imbalance of data among classes. Specifically, a huge proportion of the point cloud data corresponds to planar surfaces of the structural elements, including the ceiling, floor, and walls of the building. These structural elements points can be accurately distinguished from points belonging to furniture class by adjusting different neighborhood radius. The marginal superiority in overall accuracy achieved by the neighborhood radius of 25 cm is also attributed to its enhanced precision in classifying points corresponding to small furniture elements which have different Normal change rate values comparing to the structural elements points.

As mentioned in Section 3.3.2, the proposed algorithm for 3D space parsing step encompass crucial parameters of the ceiling to wall points distance threshold, along with the maximum distance tolerance for the DBSCAN clustering. In this regard, various values for both pertinent parameters are systematically examined to assess their impact on the overall Rand Index similarity value between the result of the proposed algorithm and reference data. According to the Fig. 27, a distance value of 35 cm produces the highest accuracy of about 93.5% in partitioning individual 3D spaces. In this regard, the initial selection of 35 cm as the default value for algorithm implementation was grounded in the average thickness of internal walls observed across all data. Departing from this value, by opting a distance parameter less than or greater than 35 cm, leads to decreased accuracy due to challenges such as over-segmentation or merging of individual spaces.

A required parameter for floor plan mask creation and subsequent parametric digital model creation is the grid dimension used for model-based space footprint extraction. The consideration of different values for the grid dimension parameter impact the quality and overall accuracy of the reconstructed digital model. To assess the influence of this parameter on the ultimate accuracy of the reconstructed digital model, a statistical analysis is conducted based on the values outlined in Tables 8–10. In this context, the grid dimensions are adjusted to 25, 50, and 100 cm, and subsequent digital models are reconstructed following the procedures detailed in Section 3.4.

According to the Fig. 28, the grid size of 25 cm produced the highest accuracy, resulting in a displacement accuracy of about 7 cm and an element parameter estimation accuracy of 5 cm for the reconstructed digital model. Meanwhile, in alternative scenarios with grid dimensions of 50 cm and 100 cm, the attained accuracy closely approximate the optimal accuracy achieved with a 25 cm grid dimension. This underscores the robustness of the proposed algorithm, indicating its minimal dependence on the initial parameter value. Furthermore, it highlights the exceptional performance of the employed optimization algorithm. Notably, despite a substantial disparity in grid dimension values between 25 cm and 100 cm, their overall accuracies differ by only 3 cm.

5.3. Limitations

Despite thorough considerations, the proposed method is subject to certain limitations that hinder its ability to generate digital models with coherent geometry.

The developed algorithm employs an AI-based decision tree classification network to separate wall and ceiling elements within the indoor point cloud. The effectiveness of this solution relies heavily on the performance of the trained model, which necessitates a huge amount of annotated data and computational power for optimal training.

Moreover, the algorithm developed exhibits certain limitations when applied to buildings constructed with glass materials. The data collection process faces challenges in capturing accurate information from glass surfaces, making it difficult for the algorithm to recognize and process these data effectively.

Due the consideration of ceiling to wall points distance threshold value equal to 0.35 in 3D space parsing step, the proposed method poses difficulties in accurately separating super-narrow hallways or emergency exit ways.

Additionally, the utilization of optimization processes during the geometry provision steps results in increased processing time for large-

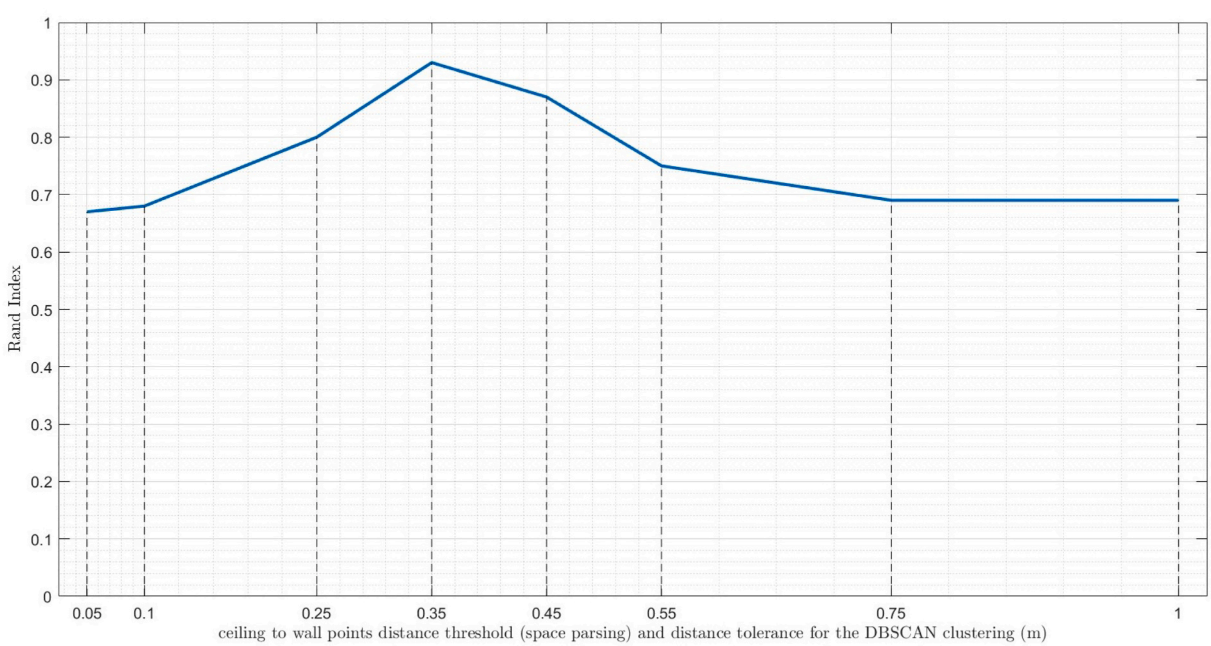


Fig. 27. Assessment the impact of the different values for ceiling to wall points distance threshold and distance tolerance for the DBSCAN clustering parameters on the overall Rand Index similarity value in 3D space parsing task.

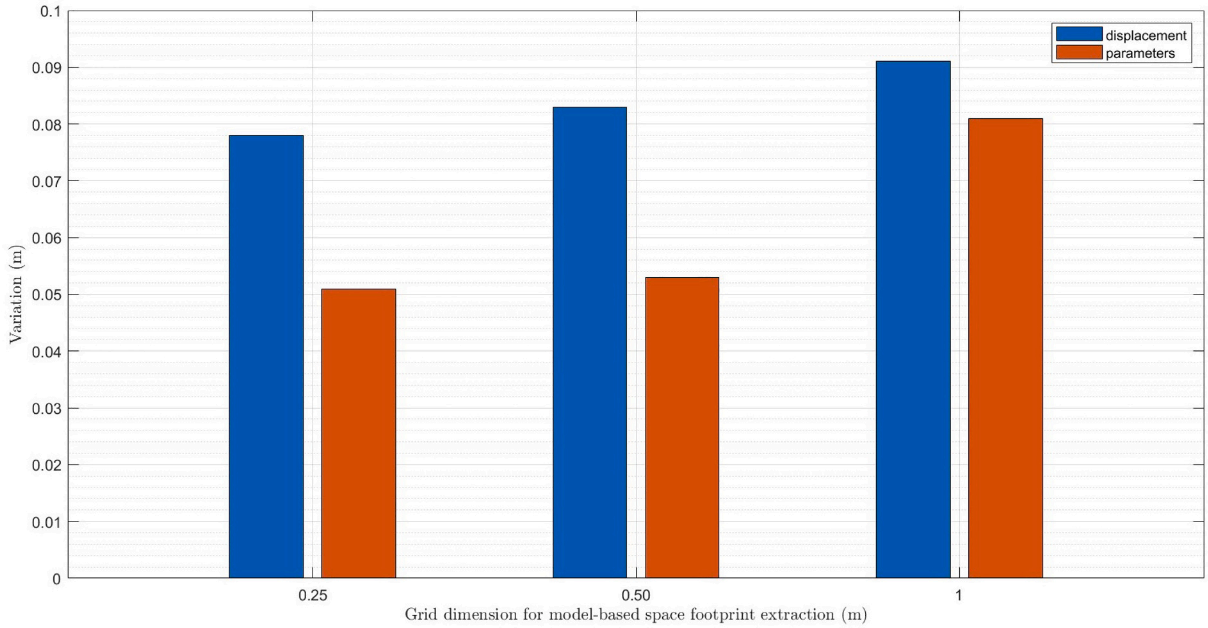


Fig. 28. Evaluating the impact of different values for the grid dimension parameter used in model-based space footprint extraction on the overall accuracy of reconstructed digital models.

scale multi-story buildings, demanding substantial processing time and computing power.

6. Conclusions

The creation of digital building models using point clouds is an advanced and cutting-edge area of research in computer vision and the AEC industry which is always associated with significant challenges in automation and required accuracy. The proposed method is a novel framework for automatically generating digital building models which rely on integrating capabilities of AI techniques in scene understanding with the domain knowledge in building design and construction.

Unlike existing data-driven approaches, our model-based proposed method produces high-quality digital models that accurately represent the semantics of the components and simulate proper relationships between them. Thanks to applying the parametric modeling process, we are able to consider the semantic connections between components, thereby enabling us to overcome prevalent obstacles and challenges in complex building point clouds such as noise and clutter within intricate indoor scenes. Due to the average accuracy of about 0.06 m in estimating models parameters, our proposed approach provides significant progress in the field of “Scan-to-BIM,” ultimately delivering high-quality digital models with high geometric accuracy and rich semantic information, providing various analysis possibilities for facility management and experts in the field.

The basis of the proposed algorithm for the creation of parametric digital models relies on extracting primary semantic information from points and partitioning individual spaces. However, this approach may present challenges when dealing with buildings constructed using glass and mirror materials. In such cases, the laser reflections and scattering from vitreous surfaces result in scattered and disjointed points, impeding the efficiency of the algorithm in the classification of points. Furthermore, a significant issue arises in the identification and differentiation between the fundamental structural walls and the movable partition walls within the building. These two types of structures share common geometric features, making it difficult to separate and distinguish them during the points classification process. Consequently, this leads to the division of more than individual spaces within the building and, subsequently, the generation of multiple excess walls in the model.

Further investigations can be carried out on improving the accuracy of the proposed method particularly in point cloud semantic enrichment through AI techniques and extending the domain knowledge for parameterization of non-Manhattan structures. This can be conducted by formulating connections of oblique wall and slab instances. Additionally, there is potential to enrich the level of development and semantic information of the result digital models by creating digital representations of additional structural-architectural elements (e.g., doors, windows, columns, etc.), enabling a more comprehensive representation of the real-world environment.

CRediT authorship contribution statement

Mansour Mehranfar: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alexander Braun:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **André Borrmann:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Mansour Mehranfar reports financial support was provided by TUM Georg Nemetschek Institute.

Data availability

The authors do not have permission to share data.

Acknowledgments

The research presented has been performed in the frame of the AI4TWINNING project (“Artificial Intelligence for the automated creation of multi-scale digital twins of the built world") funded by the TUM Georg Nemetschek Institute Artificial Intelligence for the Built World (GNI), which is thankfully acknowledged.

References

[1] S. Ochmann, R. Vock, R. Wessel, R. Klein, Automatic reconstruction of parametric building models from indoor point clouds, Comput. Graph. 54 (2016) 94–103, https://doi.org/10.1016/j.cag.2015.07.008.
[2] V. Stojanovic, M. Trapp, R. Richter, B. Hagedorn, J. Döllner, Towards the generation of digital twins for facility management based on 3D point clouds, in: 34th Annual Association of Researchers in Construction Management Conference (ARCOM 2018), Belfast, UK, 2018, pp. 270–279, last access: March 09, 2024. URL, https://www.arcom.ac.uk/-docs/proceedings/b65e593d342a8de045cf05698677e600.pdf.
[3] M. Kassem, G. Kelly, N. Dawood, M. Serginson, S. Lockley, BIM in facilities management applications: a case study of a large university complex, Built Environ. Proj. Asset Manag. 5 (3) (2015) 261–277, https://doi.org/10.1108/BEPAM-02-2014-0011.
[4] N. Bortolini, M. Forcada, M. Macarulla, BIM for the integration of building maintenance management: A case study of a university campus, in: eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2016: Proceedings of the 11th European Conference on Product and Process Modelling (ECPPM), CRC Press, Limassol, Cyprus, 2016, pp. 427–434, https://doi.org/10.1201/9781315386904.
[5] O.C. Madubuike, C.J. Anumba, Digital twin–based health care facilities management, J. Comput. Civ. Eng. 37 (2) (2022) 04022057, https://doi.org/10.1061/jccee5.cpeng-4842.
[6] H.D. Duong, J.J. Lin, Reality model-based facility management framework for existing building, Front. Built Environ. 8 (2022) 815672, https://doi.org/10.3389/fbuil.2022.815672.
[7] M. Mahmoodian, F. Shahrivar, S. Setunge, S. Mazaheri, Development of digital twin for intelligent maintenance of civil infrastructure, Sustainability 14 (14) (2022) 8664, https://doi.org/10.3390/su14148664.
[8] J. Underwood, U. Isikdag, Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies, 1st Edition, IGI Global, 2009 (ISBN: 9781605669281).
[9] O. Vera-Piazzini, M. Scarpa, F. Peron, Building energy simulation and monitoring: a review of graphical data representation, Energies 16 (1) (2022) 390, https://doi.org/10.3390/en16010390.
[10] R. Lu, I. Brilakis, Digital twinning of existing reinforced concrete bridges from labelled point clusters, Autom. Constr. 105 (2019) 102837, https://doi.org/10.1016/j.autcon.2019.102837.
[11] M. Mohri, A. Rostamizadeh, A. Talwalkar, Foundations of Machine Learning, 2nd Edition, MIT Press, 2018 (ISBN: 9780262039406).
[12] M. Austin, P. Delgoshaei, M. Coelho, M. Heidarinejad, Architecting smart city digital twins: combined semantic model and machine learning approach, J. Manag. Eng. 36 (4) (2020) 04020026, https://doi.org/10.1061/(ASCE)ME.1943-5479.0000774.
[13] M. Chen, A. Feng, R. McAlinden, L. Soibelman, Photogrammetric point cloud segmentation and object information extraction for creating virtual environments and simulations, J. Manag. Eng. 36 (2) (2020) 04019046, https://doi.org/10.1061/(ASCE)ME.1943-5479.0000737.
[14] A. Francisco, N. Mohammadi, J.E. Taylor, Smart city digital twin–enabled energy management: toward real-time urban building energy benchmarking, J. Manag. Eng. 36 (2) (2020) 04019045, https://doi.org/10.1061/(ASCE)ME.1943-5479.0000741.
[15] G. Schrotter, C. Hürzeler, The digital twin of the City of Zurich for urban planning, PFG-J. Photogram. Remote Sens. Geoinform. Sci. 88 (1) (2020) 99–112, https://doi.org/10.1007/s41064-020-00092-2.
[16] Q. Wang, M.-K. Kim, Applications of 3D point cloud data in the construction industry: a fifteen-year review from 2004 to 2018, Adv. Eng. Inform. 39 (2019) 306–319, https://doi.org/10.1016/j.aei.2019.02.007.
[17] F. Xue, W. Lu, K. Chen, A. Zetkulic, From semantic segmentation to semantic registration: derivative-free optimization-based approach for automatic generation of semantically rich as-built building information models from 3D point clouds, J. Comput. Civ. Eng. 33 (4) (2019) 04019024, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000839.
[18] J. Abualdenien, A. Borrmann, Levels of detail, development, definition, and information need: a critical literature review, J. Inform. Technol. Construct. 27 (2022) 363–392, https://doi.org/10.36680/j.itcon.2022.018.
[19] R. Sacks, C. Eastman, G. Lee, P. Teicholz, BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers, 3rd edition, John Wiley & Sons, 2018 (ISBN: 9781119287537).
[20] A. Dlesk, K. Vach, R. Shults, P. Doubrava, Generalization of BIM model for purposes of facility management, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B4–2022, 2022, pp. 309–314, https://doi.org/10.5194/isprs-archives-XLIII-B4-2022-309-2022.
[21] J. Abualdenien, A. Borrmann, Vagueness visualization in building models across different design stages, Adv. Eng. Inform. 45 (2020) 101107, https://doi.org/10.1016/j.aei.2020.101107.
[22] BIMForum, Level of Development (LOD) Specification (Part 1). http://bimforum.org/lod/, 2023 last access: March 09, 2024.
[23] F. Bosché, M. Ahmed, Y. Turkan, C.T. Haas, R. Haas, The value of integrating scan-to-BIM and scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: the case of cylindrical MEP components, Autom. Constr. 49 (Part B) (2015) 201–213, https://doi.org/10.1016/j.autcon.2014.05.014.
[24] R. Laing, M. Leon, J. Isaacs, D. Georgiev, Scan to BIM: the development of a clear workflow for the incorporation of point clouds within a BIM environment, WIT Trans. Built Environ. 149 (2015) 279–289, https://doi.org/10.2495/BIM150241.
[25] J.C. Sarmiento, A.M. Bercero, J.C. Fernandez, P. Garcia-Almirall, SCAN to BIM beyond a final BIM: why, when and how, in: proceedings of, IOP Conf. Ser. Mater. Sci. Eng. 603 (2019) 042090, https://doi.org/10.1088/1757-899X/603/4/042090.
[26] E. Agapaki, I. Brilakis, Geometric Digital Twinning of Industrial Facilities: Retrieval of Industrial Shapes, arXiv:2202.04834, (Accessed 10 February 2022), 2024, https://doi.org/10.48550/arXiv.2202.04834.
[27] X. Xiong, A. Adan, B. Akinci, D. Huber, Automatic creation of semantically rich 3D building models from laser scanner data, Autom. Constr. 31 (2013) 325–337, https://doi.org/10.1016/j.autcon.2012.10.006.
[28] S. Ochmann, R. Vock, R. Klein, Automatic reconstruction of fully volumetric 3D building models from oriented point clouds, ISPRS J. Photogramm. Remote Sens. 151 (2019) 251–262, https://doi.org/10.1016/j.isprsjprs.2019.03.017.
[29] S. Nikoohemat, A.A. Diakité, S. Zlatanova, G. Vosselman, Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management, Autom. Constr. 113 (2020) 103109, https://doi.org/10.1016/j.autcon.2020.103109.
[30] H. Tran, K. Khoshelham, Procedural reconstruction of 3D indoor models from lidar data using reversible jump Markov Chain Monte Carlo, Remote Sens. 12 (5) (2020) 838, https://doi.org/10.3390/rs12050838.
[31] J. Martens, J. Blankenbach, VOX2BIM + − a fast and robust approach for automated indoor point cloud segmentation and building model generation, PFG – J. Photogram. Remote Sens. Geoinform. Sci. 91 (4) (2023) 273–294, https://doi.org/10.1007/s41064-023-00243-1.
[32] Y. Wu, L. Maosu, F. Xue, Towards fully automatic scan-to-BIM: A prototype method integrating deep neural networks and architectonic grammar, in: Proceedings of European Conference on Computing in Construction (EC3) and the 40th International CIB W78 Conference, Vol. 4, European Council on Computing in Construction, Heraklion, Greece, 2023, https://doi.org/10.35490/EC3.2023.257.
[33] D. Hu, V. Gan, C. Yin, Robot-assisted mobile scanning for automated 3D reconstruction and point cloud semantic segmentation of building interiors, Autom. Constr. 152 (2023) 104949, https://doi.org/10.1016/j.autcon.2023.104949.
[34] A. Abdollahi, H. Arefi, S. Malihi, M. Maboudi, Progressive model-driven approach for 3D modeling of indoor spaces, Sensors 23 (13) (2023) 5934, https://doi.org/10.3390/s23135934.
[35] M. Jarzabek-Rychard, H.-G. Maas, Modeling of 3D geometry uncertainty in scan-to-BIM automatic indoor reconstruction, Autom. Constr. 154 (12) (2023) 105002, https://doi.org/10.1016/j.autcon.2023.105002.
[36] E. Borenstein, E. Sharon, S. Ullman, Combining top-down and bottom-up segmentation, in: Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), IEEE, Washington, DC, USA, 2004, p. 46, https://doi.org/10.1109/CVPR.2004.314.
[37] Z.C. Marton, R.B. Rusu, M. Beetz, On fast surface reconstruction methods for large and noisy point clouds, in: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), IEEE, Kobe, Japan, 2009, pp. 3218–3223, https://doi.org/10.1109/ROBOT.2009.5152628.
[38] A.-V. Vo, L. Truong-Hong, D.F. Laefer, M. Bertolotto, Octree-based region growing for point cloud segmentation, ISPRS J. Photogramm. Remote Sens. 104 (2015) 88–100, https://doi.org/10.1016/j.isprsjprs.2015.01.011.
[39] F. Poux, C. Mattes, Z. Selman, L. Kobbelt, Automatic region-growing system for the segmentation of large point clouds, Autom. Constr. 138 (2022) 104250, https://doi.org/10.1016/j.autcon.2022.104250.
[40] J. Xiao, J. Zhang, B. Adler, H. Zhang, J. Zhang, Three-dimensional point cloud plane segmentation in both structured and unstructured environments, Robot. Auton. Syst. 61 (12) (2013) 1641–1652, https://doi.org/10.1016/j.robot.2013.07.001.
[41] A. Dimitrov, M. Golparvar-Fard, Segmentation of building point cloud models including detailed architectural/structural features and MEP systems, Autom. Constr. 51 (2015) 32–45, https://doi.org/10.1016/j.autcon.2014.12.015.
[42] M. Arikan, M. Schwärzler, S. Flöry, M. Wimmer, S. Maierhofer, O-snap: optimization-based snapping for modeling architecture, ACM Trans. Graph. 32 (1) (2013) 1–15, https://doi.org/10.1145/2421636.2421642.
[43] H. Macher, T. Landes, P. Grussenmeyer, Point clouds segmentation as base for as-built BIM creation, ISPRS Ann. Photogram. Remote Sens. Spat. Inform. Sci. II-5 (W3) (2015) 191–197, https://doi.org/10.5194/isprsannals-II-5-W3-191-2015.
[44] B. Bhanu, S. Lee, C.-C. Ho, T. Henderson, Range data processing: representation of surfaces by edges, in: Proceedings of The Eighth International Conference on Pattern Recognition, IEEE, Paris, France, 1986, pp. 236–238, last access: March 09, 2024. URL, https://api.semanticscholar.org/CorpusID:268107315.
[45] Q. Qiu, M. Wang, J. Guo, Z. Liu, Q. Wang, An adaptive down-sampling method of laser scan data for scan-to-BIM, Autom. Constr. 135 (2022) 104135, https://doi.org/10.1016/j.autcon.2022.104135.
[46] G. Vosselman, Advanced point cloud processing, in: Proceedings of the Photogrammetric Week, Heidelberg, Germany, 2009, pp. 137–146, last access: March 09, 2024. URL, https://phowo.ifp.uni-stuttgart.de/publications/phowo09/160Vosselman.pdf.
[47] A. Nuchter, H. Surmann, J. Hertzberg, Automatic model refinement for 3D reconstruction with mobile robots, in: Proceedings of Fourth International

Conference on 3-D Digital Imaging and Modeling (3DIM), IEEE, Banff, AB, Canada, 2003, pp. 394–401, https://doi.org/10.1109/IM.2003.1240274.
[48] J. Chen, B. Chen, Architectural modeling from sparsely scanned range data, Int. J. Comput. Vis. 78 (2008) 223–236, https://doi.org/10.1007/s11263-007-0105-5.
[49] F. Yu, K. Liu, Y. Zhang, C. Zhu, K. Xu, PartNet: A recursive part decomposition network for fine-grained and hierarchical shape segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Long Beach, CA, USA, 2019, pp. 9491–9500, https://doi.org/10.1109/CVPR.2019.00972.
[50] T. Czerniawski, B. Sankaran, M. Nahangi, C. Haas, F. Leite, 6D DBSCAN-based segmentation of building point clouds for planar object classification, Autom. Constr. 88 (2018) 44–58, https://doi.org/10.1016/j.autcon.2017.12.029.
[51] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, J. Xiao, 3D ShapeNets: A deep representation for volumetric shapes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA, 2015, pp. 1912–1920, https://doi.org/10.1109/CVPR.2015.7298801.
[52] R.Q. Charles, W. Liu, C. Wu, H. Su, J.G. Leonidas, Frustum PointNets for 3D object detection from RGB-D Data, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Salt Lake City, UT, USA, 2018, pp. 918–927, https://doi.org/10.1109/CVPR.2018.00102.
[53] D. Prokhorov, A convolutional learning system for object classification in 3-D lidar data, IEEE Trans. Neural Netw. 21 (5) (2010) 858–863, https://doi.org/10.1109/TNN.2010.2044802.
[54] C.R. Qi, H. Su, K. Mo, L.J. Guibas, PointNet: Deep learning on point sets for 3D classification and segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, USA, 2017, pp. 77–85, https://doi.org/10.1109/CVPR.2017.16.
[55] H. Thomas, C.R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, L.J. Guibas, KPConv: Flexible and deformable convolution for point clouds, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Seoul, Korea (South), 2019, pp. 6411–6420, https://doi.org/10.1109/ICCV.2019.00651.
[56] M.M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst, Geometric deep learning: going beyond euclidean data, IEEE Signal Process. Mag. 34 (4) (2017) 18–42, https://doi.org/10.1109/MSP.2017.2693418.
[57] I. Armeni, O. Sener, A.R. Zamir, H. Jiang, I. Brilakis, M. Fischer, S. Savarese, 3D semantic parsing of large-scale indoor spaces, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Las Vegas, NV, USA, 2016, pp. 1534–1543, https://doi.org/10.1109/CVPR.2016.170.
[58] F.J. Lawin, M. Danelljan, P. Tosteberg, G. Bhat, F.S. Khan, M. Felsberg, Deep projective 3D semantic segmentation, in: Proceedings of 17th International Conference on Computer Analysis of Images and Patterns (CAIP), Springer, Ystad, Sweden, 2017, pp. 95–107, https://doi.org/10.1007/978-3-319-64689-3_8.
[59] A. Dai, M. Nießner, 3DMV: Joint 3D-multi-view prediction for 3D semantic scene segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV) vol. 11214, Springer, Munich, Germany, 2018, pp. 458–474, https://doi.org/10.1007/978-3-030-01249-6_28.
[60] H. Zhao, L. Jiang, J. Jia, P.H. Torr, V. Koltun, Point transformer, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Montreal, QC, Canada, 2021, pp. 16259–16268, https://doi.org/10.1109/ICCV48922.2021.01595.
[61] V. Pratt, Direct least-squares fitting of algebraic surfaces, ACM SIGGRAPH Comput. Graph. 21 (4) (1987) 145–152, https://doi.org/10.1145/37402.37420.
[62] S.-W. Kwon, F. Bosche, C. Kim, C.T. Haas, K.A. Liapi, Fitting range data to primitives for rapid local 3D modeling using sparse range point clouds, Autom. Constr. 13 (1) (2004) 67–81, https://doi.org/10.1016/j.autcon.2003.08.007.
[63] J.J. Park, P. Florence, J. Straub, R. Newcombe, S. Lovegrove, DeepSDF: Learning continuous signed distance functions for shape representation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Long Beach, CA, USA, 2019, pp. 165–174, https://doi.org/10.1109/CVPR.2019.00025.
[64] A. Adan, D. Huber, 3D reconstruction of interior wall surfaces under occlusion and clutter, in: Proceedings of International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, IEEE, Hangzhou, China, 2011, pp. 275–281, https://doi.org/10.1109/3DIMPVT.2011.42.
[65] V. Stamati, I. Fudos, On reconstructing 3D feature boundaries, Computer-Aided Design and Applications 5 (1) (2008) 316–324, https://doi.org/10.3722/cadaps.2008.316-324.
[66] S. Wang, Y. Tseng, Semi-automated CSG model-based building extraction from photogrammetric images, in: Proceedings of the XXth Congress of the International Society for Photogrammetry and Remote Sensing vol. 24, 2004, pp. 277–284, last access: March 09, 2024. URL, https://www.isprs.org/proceedings/XXXV/congress/comm3/papers/310.pdf.
[67] J. Xiao, Y. Furukawa, Reconstructing the World's museums, Int. J. Comput. Vis. 110 (2014) 243–258, https://doi.org/10.1007/s11263-014-0711-y.
[68] Q. Wu, K. Xu, J. Wang, Constructing 3D CSG models from 3D raw point clouds 37 (5) (2018) 221–232, https://doi.org/10.1111/cgf.13504.
[69] X. Liu, X. Wang, G. Wright, J. Cheng, X. Li, R. Liu, A state-of-the-art review on the integration of building information modeling (BIM) and geographic information system (GIS), ISPRS Int. J. Geo Inf. 6 (2) (2017) 53, https://doi.org/10.3390/ijgi6020053.
[70] M.S. Mafipour, S. Vilgertshofer, A. Borrmann, Deriving digital twin models of existing bridges from point cloud data using parametric models and metaheuristic algorithms, in: Proceedings of EG-ICE 2021 Workshop on Intelligent Computing in Engineering, University Press of the Technical University Berlin, Berlin, Germany, 2021, pp. 464–474, https://doi.org/10.14279/depositonce-12021.
[71] M.S. Mafipour, S. Vilgertshofer, A. Borrmann, Creating digital twins of existing bridges through AI-based methods, in: Proceedings of the IABSE Symposium – Challenges for Existing and Oncoming Structures, Prague, Czech, 2022, pp. 727–734, https://doi.org/10.2749/prague.2022.0727.
[72] A. Borrmann, M. Flurl, J.R. Jubierre, R.-P. Mundani, E. Rank, Synchronous collaborative tunnel design based on consistency-preserving multi-scale models, Adv. Eng. Inform. 28 (4) (2014) 499–517, https://doi.org/10.1016/j.aei.2014.07.005.
[73] M. Elsayed, O. Tolba, A. El Antably, Architectural space planning using parametric modeling, in: Proceedings of the 8th Arab Society for Computer Aided Architectural Design Conference (ASCAAD), London, UK, 2016, p. 45, last access: March 09, 2024. URL, https://papers.cumincad.org/data/works/att/ascaad2016_007.pdf.
[74] M. Keshavarzi, M. Rahmani-Asl, GenFloor: interactive generative space layout system via encoded tree graphs, Front. Architect. Res. 10 (4) (2021) 771–786, https://doi.org/10.1016/j.foar.2021.07.003.
[75] A.J. Trevor, S. Gedikli, R.B. Rusu, H.I. Christensen, Efficient organized point cloud segmentation with connected components, in: Proceedings of the semantic perception and mapping exploration, Karlsruhe, Germany, 2013, pp. 1–6, last access: March 09, 2024. URL, https://cs.gmu.edu/kosecka/ICRA2013/spme13_trevor.pdf.
[76] Y. Wu, J. Shang, F. Xue, RegARD: symmetry-based coarse registration of smartphone's colorful point clouds with CAD drawings for low-cost digital twin buildings, Remote Sens. 13 (10) (2021) 1882, https://doi.org/10.3390/rs13101882.
[77] N. Chehata, L. Guo, C. Mallet, Airborne lidar feature selection for urban classification using random forests, in: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVIII-3/W8, 2009, pp. 207–212, last access: March 09, 2024. URL, https://www.isprs.org/proceedings/xxxviii/3-w8/papers/207_laserscanning09.pdf.
[78] E. Grilli, E. Farella, A. Torresani, F. Remondino, Geometric features analysis for the classification of cultural heritage point clouds, in: The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XLII-2/W15, 2019, pp. 541–548, https://doi.org/10.5194/isprs-archives-XLII-2-W15-541-2019.
[79] Y. Pan, A. Braun, A. Borrmann, I. Brilakis, 3D deep-learning-enhanced void-growing approach in creating geometric digital twins of buildings, Smart Infrastruct. Construct. 176 (1) (2023) 24–40, https://doi.org/10.1680/jsmic.21.00035.
[80] C. Mura, O. Mattausch, A. Jaspe-Villanueva, E. Gobbetti, R. Pajarola, Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts, Comput. Graph. 44 (2014) 20–32, https://doi.org/10.1016/j.cag.2014.07.005.
[81] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD), Vol. 96, Portland Oregon, USA, 1996, pp. 226–231, last access: March 09, 2024. URL, https://cdn.aaai.org/KDD/1996/KDD96-037.pdf.
[82] J. Nelder, R. Mead, A simplex method for function minimization, Comput. J. 7 (4) (1965) 308–313, https://doi.org/10.1093/comjnl/7.4.308.
[83] T.F. Chan, G.H. Golub, R.J. LeVeque, Updating formulae and a pairwise algorithm for computing sample variances, in: Proceedings in 5th Symposium of Computational Statistics COMPSTAT 1982, Toulouse, France, 1982, pp. 30–41, https://doi.org/10.1007/978-3-642-51461-6_3.
[84] A. Ben Hur, D. Horn, H. Siegelmann, V. Vapnik, Support vector clustering, J. Mach. Learn. Res. 2 (2001) 125–137, last access: March 09, 2024. URL, https://www.jmlr.org/papers/volume2/horn01a/horn01a.pdf.
[85] W.M. Rand, Objective criteria for the evaluation of clustering methods, J. Am. Stat. Assoc. 66 (336) (1971) 846–850, https://doi.org/10.2307/2284239.
[86] Y. Cai, L. Fan, An efficient approach to automatic construction of 3D watertight geometry of buildings using point clouds, Remote Sens. 13 (10) (2021) 1947, https://doi.org/10.3390/rs13101947.