Energy & Buildings 360 (2026) 117384

Contents lists available at ScienceDirect

Energy & Buildings

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

Automated window detection for digital twin buildings: A generalized 'sandwich' model and practical guidelines

Siyuan Meng [a], Longyong Wu [a], Maosu Li [b], Anthony G. O. Yeh [c,d], Fan Xue [a,d,*]

[a] Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, China
[b] Thrust of Urban Governance and Design, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, China
[c] Department of Urban Planning and Design, The University of Hong Kong, Pokfulam, Hong Kong, China
[d] National Center of Technology Innovation for Digital Construction Hong Kong Branch, The University of Hong Kong, Pokfulam, Hong Kong, China

A R T I C L E I N F O

Keywords:
Digital twin building
Openings
Window detection
Decision tree
Systematic review

A B S T R A C T

Windows play a significant role in the livability and sustainability of buildings, such as energy efficiency, natural lighting, and ventilation; therefore, they are increasingly important subjects of surveys and analysis for both new and existing buildings. Many automated window detection (AWD) workflows have been studied to extract windows' properties; however, a generalized model of the underlying technologies and practical guidelines remains lacking for both newcomers and industry practitioners. Therefore, this paper aims to address the following questions: (1) the scope and general workflow of AWD; (2) the trends regarding AWD key components; and (3) the pathways of key component (e.g., data source, attribute, storage, and method) selections to fit the application scenario. Four decision trees, trained on 105 technical papers collected in accordance with the PRISMA standard, achieved an average F1-score of 80.3% and accuracy of 88.1%. This paper contributes the first comprehensive review of AWD, with a general 'sandwich' model, a comprehensive analysis of advanced technologies, trends, and future directions, as well as three-step guidelines for practitioners.

1. Introduction

Sustainable development of buildings is increasingly recognized worldwide. Buildings serve as primary hotspots of resource consumption and environmental impact [1], with windows constituting major components of buildings that significantly contribute to this overall impact. Windows are openings in walls, doors, or rooftops that serve both physical (e.g., light exchange, air circulation, enhancement of indoor thermal comfort, and sound transmission) and psychological (e.g., improvement of mental health, facade articulation, and aesthetic enjoyment) functions for residents [2–7]. Windows' potential environmental impacts include energy consumption and greenhouse gas emissions during the construction, operation, and demolition phases [1,8]. Additionally, windows are key elements in the analyses of flood, wind, and fire damage risks when evaluating potential environmental damage [9]. Therefore, proper design, construction, and maintenance of windows are essential for promoting sustainable building development and supporting the United Nations' Sustainable Development Goal 11 (SDG 11) [10].

Digital twin buildings (DTBs), virtual replicas of physical buildings, have been increasingly incorporated into building life-cycle projects in the past decade under the concept of Construction 4.0 [11]. Typical DTBs encompass acquisition, transmission, modeling, data/model integration, as well as service layers [12]. DTBs provide dynamic data inputs that, in conjunction with digital models, further support reasoning functionalities beyond those enabled by static models such as building information model (BIM) [13,14]. These reasoning functionalities include project progress monitoring, construction information exchange, building energy performance simulation and evaluation, and building maintenance optimization [15–17]. Digital twin windows, the extension of DTB, can therefore empower window information modeling and enable building life-cycle applications.

Automated window detection (AWD) refers to an automated workflow that aims to detect window-related information, such as position, shape, and condition. The dynamic digitalized windows can serve as the basis for model and data/model integration layers in DTBs, owing to the intermediary role of windows in real-time interaction with the built environment. AWD can further support DTB-based life-cycle reasoning in response to increasing sustainability demands, such as light

* Corresponding author.
E-mail addresses: siyuanm@connect.hku.hk (S. Meng), wulongyong@connect.hku.hk (L. Wu), maosuli@hkust-gz.edu.cn (M. Li), xuef@hku.hk (F. Xue).

https://doi.org/10.1016/j.enbuild.2026.117384
Received 8 December 2025; Received in revised form 6 February 2026; Accepted 24 March 2026
Available online 28 March 2026

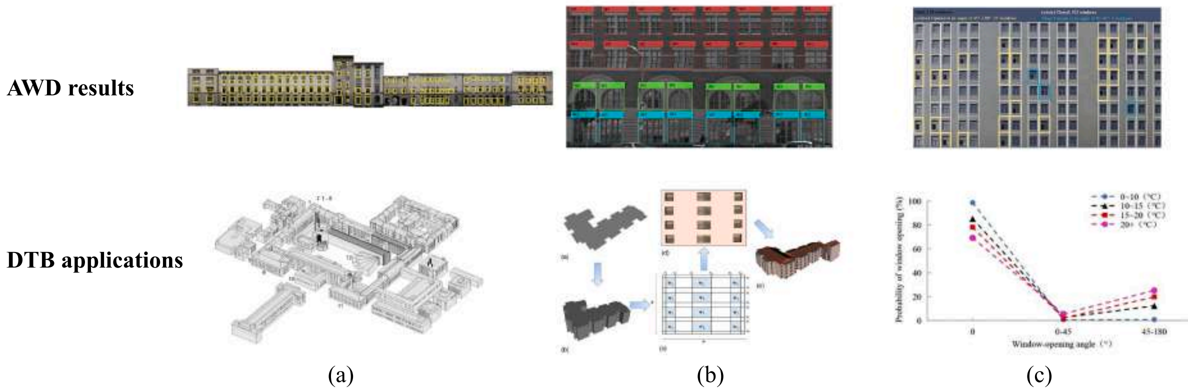

Fig. 1. Examples of detected window attributes and DTB applications in the literature. (a) Window sizes from a building in the temperate zone [21]; (b) Window types from a building in the arid zone [22], with the application of semantic enrichment modeling; (c) Window positions and opening statuses from a building in the cold zone [23], with the application of window open behavior analysis.

and visual comfort design in design phase, construction robotics monitoring in construction phase, and monitoring carbon dioxide emissions in O&M phase [18–20].

AWD has been studied in many cities using advanced remote sensing and detection technologies. For example, window positions and opening status were examined in Harbin City, which is located in a cold region, as shown in Fig. 1c. Various data sources and technologies are employed in the AWD workflow due to the significance of window-related attributes, such as position, size, and type, for DTB applications. Typical data sources include imagery, point clouds, and thermal cameras, which are used to collect information on windows' open or closed status, geometric shapes, types, materials, and other relevant details. Both rule-based and learning-based methods are applied across multiple disciplines to refine AWD workflows.

Various applications of AWD have been implemented in the architecture, engineering, and construction (AEC) industry. Multidisciplinary research from other fields, such as geoscience, remote sensing, and instrumentation, also involves AWD applications. The applications of AWD can be categorized into two types: semantic enrichment and window semantics for reasoning. Specifically, the shape of detected windows can first be enriched as semantic information in 3D BIM (i.e., Scan-to-BIM) and city information models (CIMs) for managing and analyzing building information [24,25]. Furthermore, window semantics also enable various DTB simulations and reasoning tasks. For example, the opening status of windows is utilized to monitor changes in window opening behavior [23]; the position of windows is combined with facade area to calculate the window-to-wall ratio [26]; the type of window is utilized to refine BIM models [22]; moreover, the shapes of windows can be utilized to monitor construction progress [27].

However, a new researcher or practitioner in related fields may find AWD perplexing despite its fundamental role in DTB-related functions. Existing AWD workflows, including objectives, methods, and outcomes of AWD tasks across various applications, remain fragmented. Besides, they differ substantially in their pipelines and level of automation, including data source selection, feature engineering, model storage, processing methods, and intended application. For example, data sources, such as LiDAR point clouds and imagery, vary across the construction project stages to support information management [28–30]. As for feature engineering, energy simulation primarily focuses more on window open/close status [31], whereas building information extraction only requires the existence and quantity of windows [32]. Despite its growing relevance, AWD lacks a systematic review to summarize the general information extraction process, and no structured guidelines are available to assist newcomers in the field. Existing literature usually considers AWD as a secondary element embedded in other industry-specific processes or in narrowly scoped, task-specific scenarios [24]. Therefore, this paper aims to contribute the first review that summarizes the AWD model, workflow, and guidelines in AEC industry, thereby filling a gap in the literature.

The knowledge gap in the literature between AWD and related applications, therefore, can be summarized as the following research questions:

- What is the scope and general workflow of AWD?
- What are the technological advancements and trends regarding AWD data sources, attributes, methods, and applications in the literature?
- How can a practitioner select appropriate data sources, attributes, storage, detection methods, and pathways to fit the application scenario?

This paper employs a mixed-methods research design to systematically answer the research questions. First, a conceptual 'sandwich' model of the AWD workflow is defined to summarize the scope and general processing steps. Then, a literature data collection process is conducted based on the conceptual model. To answer the second question, the subsequent comprehensive literature review examines the development trends of AWD data sources, attributes, methods, and applications. Finally, a decision tree approach is applied, with the authors' moderation, to derive concise and comprehensive AWD guidelines from the existing workflows in the literature.

2. Research methods

2.1. Definition and conceptual 'sandwich' model of AWD

Window detection refers to the task of identifying interested attributes–such as the position, shape, and condition of windows–within a building, as an extension of object detection [33]. AWD denotes the automated workflows for collecting and processing information to accomplish window detection tasks. In DTB, AWD typically comprises workflows for defining requirements, collecting data and information, and executing techniques across three main layers [34]. Thus, requirements, data and information, and execution constitute the key steps of AWD.

Fig. 2 presents a conceptual 'sandwich' model of requirement-driven AWD studies as reviewed in this paper. The conceptual model outlines the general steps of AWD identified in the literature, which comprise three layers. The top layer is the DTB requirement layer, which includes construction phases in the lifecycle and application needs. The 'meat' layer in the middle is the data and information layer, comprising data sources, attributes, and model storage. The base layer is the execution layer, which primarily involves rule-based and learning-based (including both traditional and deep learning) methods. In the model in Fig. 2, the initial application block in the requirement layer determines the

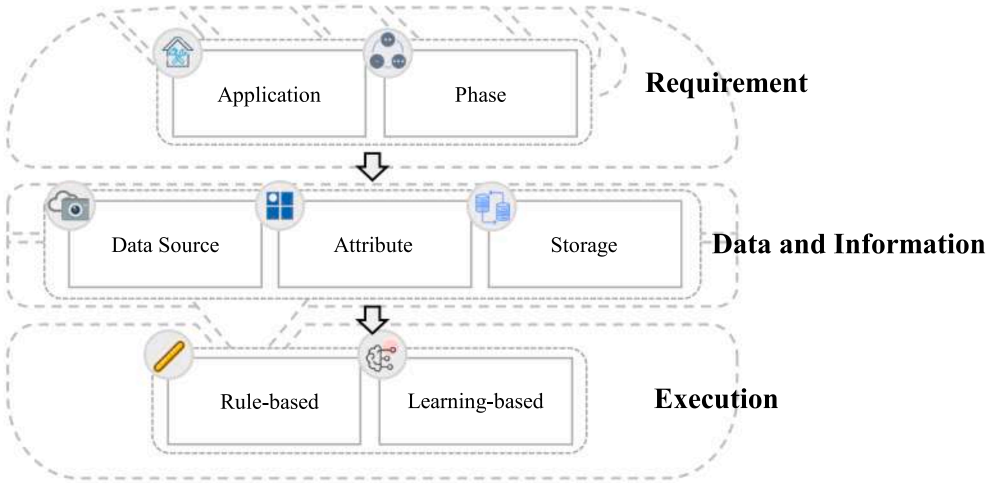

Fig. 2. Conceptual 'sandwich' model of AWD studies.

Table 1
Value and notes of the seven components in the 'sandwich' model's three layers.

Layers	Components	Value	Notes	Example
Requirement	Phase	Design	\	[35]
		Construction	\	[29]
		O&M	Operations and Management	[36]
	Application	Building information extraction	Extract Building high-level information (e.g., floor elevation, floor number)	[37]
		Construction robotics	Use automated robotics for building construction tasks.	[38]
		Damage simulation	Simulate damage impact on buildings (e.g., flood, winds).	[39]
		Energy simulation	Predict energy behaviour of buildings (e.g., thermal comfort, energy consumption)	[40]
		Heritage conservation	Technical activity towards historical buildings	[41]
		Semantic enrichment	Interchange and interoperate on a semantic level	[42]
Data & info.	Data source	Drawings	Floor plan images/vectors of buildings	[43]
		IRT	Infrared thermography	[44]
		RGB camera	Camera that captures light in red, green, and blue wavelengths (RGB)	[45]
		BIM	Building information model	[37]
		LiDAR	Light Detection And Ranging	[46]
		RGB-D	Depth camera that provide both depth and RGB color information	[47]
	Attribute	Dimension	Estimate size like width and height of window	[48]
		Position	Estimate the location of window	[49]
		Shape	Estimate the geometry patch of window	[50]
		Status	Estimate whether window is close, open or half-open	[23]
		Type	Classify the window type (e.g., style)	[22]
	Storage	2D polygon	Store in 2D image	[51]
		3D model	Store in 3D data like BIM and Mesh	[52]
		Raw	Not specified	[53]
Execution	Rule-based	\	Detect combined with rule and domain knowledge	[54]
	Learning-based	Classification	Distinguish type or style of windows	[28]
		GAN	Generative adversarial network	[55]
		Instance segmentation	Annotate windows boundary of each window	[56]
		Localization	Annotate windows position	[57]
		Semantic segmentation	Annotate windows boundary as a single category	[27]

contextual factors that inform the middle information layer of DTB. The bottom layer supports the data processing and the exchange of information between the three components in the middle.

Table 1 lists the possible values and notes for the components of the 'sandwich' model across the three layers depicted in Fig. 2. These values represent the component options utilized across the AWD literature.

In the top requirement layers, design, construction, as well as operation and maintenance (O&M) are included in the phase component, aligning with the building lifecycle. The application component encompasses six key analytical categories. These application scenarios utilize the AWD results and further extend their relevance in published works, including semantic enrichment, building information extraction, construction robotics, damage simulation, energy simulation, and heritage conservation.

The data and information 'meat' layer comprises six data sources (LiDAR, RGB camera, drawings, IRT, RGB-D, and BIM), five attribute types (shape, position, status, type, and dimension), and three storage formats (3D model, 2D polygon, and raw data). It is worth noting that RGB-D and LiDAR represent distinct data sources, as RGB-D technology is unsuitable for outdoor environments in AWD tasks.

In the base execution layer, the rule-based component refers to methods described in the literature that rely on domain knowledge and explicit rules (e.g., symmetry and similarity). In contrast, the learning-based component encompasses five prevalent deep learning approaches identified in the literature: classification, localization, semantic segmentation, instance segmentation, and generative adversarial networks (GANs).

2.2. Literature search

The literature data collection was conducted according to the conceptual model and the Preferred Reporting Items for Systematic Reviews

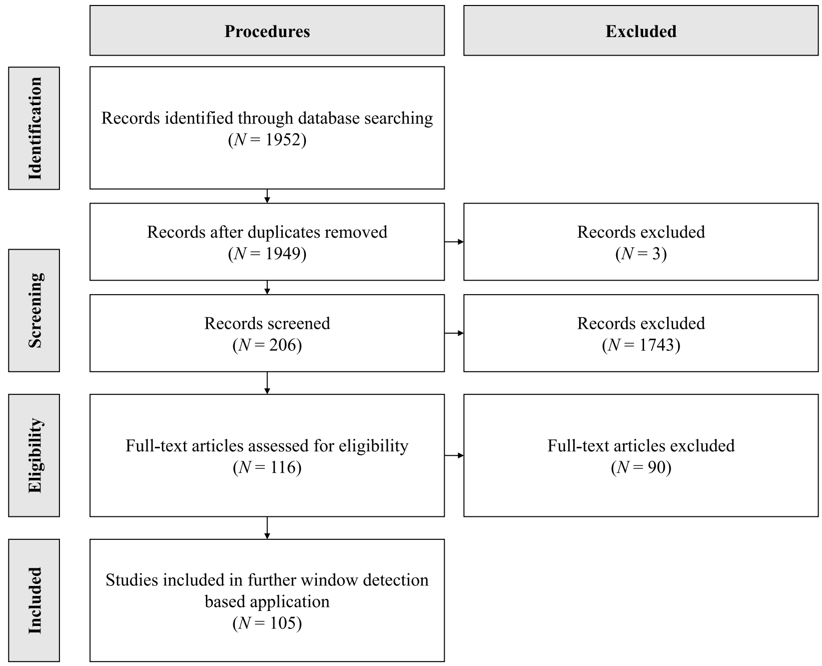

Fig. 3. PRISMA workflow.

and Meta-Analyses (PRISMA) standard, as illustrated in Fig. 3. The Web of Science academic database was used to identify relevant documents. The search strategy employed two groups of keywords: Group 1 targeted papers focused on window or opening detection ((window* OR opening*) AND (detection OR recognition OR segmentation OR extract* OR parsing OR reconstruct*)), while Group 2 was designed to narrow the scope of the AWD domain to buildings only (architecture OR city OR cities OR building* OR facade*). The literature was screened from 2000 to 2025 to reflect application trends.

To identify relevant literature, an initial screening was conducted based on the title, abstract, and keywords to select articles that matched the search criteria. In addition to sources from the AEC industry, application-oriented literature from computer science and remote sensing fields related to buildings was included to provide a more comprehensive understanding of the diversity and trends of AWD across different scenarios. Three duplicate entries and 1,743 out-of-scope publications were excluded. Furthermore, publications lacking one or more layers of the conceptual model–for example, remote sensing studies without any analysis or application–were removed, totaling 90 publications, as illustrated in Fig. 3. The full texts of the remaining articles were examined to ensure comprehensive coverage. In total, 105 papers relevant to the proposed review topic were collected, with the earliest published in 2009; of these, 78 were journal articles and 27 were conference papers.

2.3. Literature analysis based on the conceptual model

The 105 selected publications were coded according to the components of the conceptual model. A heat map based on the corresponding author's affiliated institution for each publications was generated to reflect the distribution of AWD applications across different climate zones.

A Sankey chart was employed to visually represent the coded literature as an overview of the existing studies. The intertwined links in the Sankey chart denote the systematic and organized technological roadmaps in the literature. By examining the network representation of the shortlisted papers in relation to the conceptual models, we summarized the general workflow of AWD.

This review examined the latest technological trends in applications, data sources, attributes, and methods related to AWD. Each component, along with all values presented in the Sankey chart, is depicted using a stacked bar plot. These trends demonstrate the current choices for AWD within the conceptual model.

2.4. Guidelines compilation

A set of AWD guidelines is compiled based on the conceptual model and literature analysis. In the guidelines, each step was summarized based on the reviewed publications. The AWD guidelines provide practitioners with a comprehensive map to choose the optimal methods for window detection. For researchers, the proposed guideline, in conjunction with technological advancements and trends, serves as a foundation for future research to explore more innovative research concepts.

The selection procedures for components in the data and information layer and the execution layer were illustrated using a series of decision trees. Decision trees are interpretable methods widely used to derive insights from complex datasets across various disciplines [58]. Four decision trees were constructed through in-depth analysis using *OrangeML* (ver. 3.36) in *Python* (ver. 3.11). The decision tree algorithm's primary parameter was set to 'min_samples_split = 2, ' while the rest of the parameters were set to their default values. The input variables for the data and information layer were defined as requirement layer components (i.e., application and phase), while input variables for the execution layer were set as data source and storage. Literature published after 2020 was selected to ensure the methods reflect current practice. Unsuitable or outdated options were manually excluded based on findings from the literature analysis. Methods, processing steps, and applications were manually synthesized from the selected publications.

Literature published after 2020, indexed by Scopus but not by Web of Science, was collected to validate the recommended decision trees. Validation metrics, including accuracy, precision, recall, and F1-score, were employed to demonstrate the accuracy and robustness of the four decision trees.

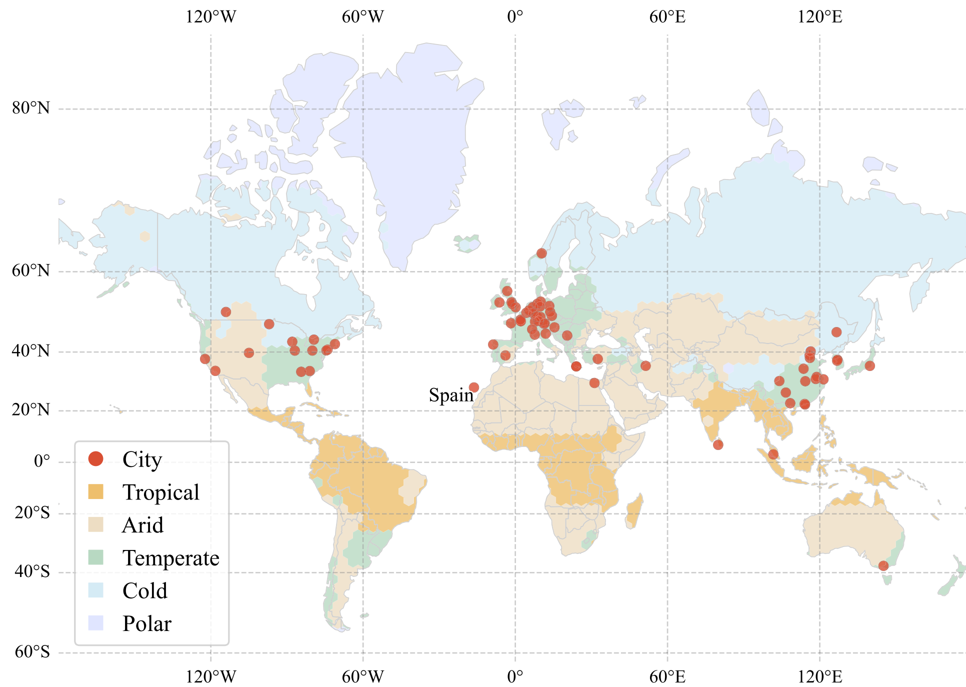

Fig. 4. Map of source cities with climate zone of the 105 publications in Table A.1.

3. Data extraction and analytical results

3.1. Overview of 105 AWD publications

A total of 105 records of AWD technological pipelines were collected, of which the values of the seven components of the 'sandwich' model were manually encoded. Fig. 4 shows the origin cities with climate zone of the 105 publications. The majority of papers had original cities in the temperate zone (totaling 78), potentially because of diverse building typologies and high-density population densities in these regions. The most prominent cluster is concentrated in the temperate zone of Europe. Arid, tropical, and cold zones contained 15, 8, and 4 papers, respectively. No studies originated from polar zones, possibly because of the small population size and low-rise, low-density buildings.

Fig. 5 presents a Sankey chart of the 105 technical pipelines, highlighting the frequency and relationships among the coded components. The rectangles containing numbers within each subcategory indicate the total number of publications. Scenarios related to construction and building technology (C&BT) account for 60–over half of the 105 papers on AWD workflows, whereas computer science (CS)-centric scenarios contribute approximately 11% of the methods. This contrast demonstrates that domain-specific datasets and application scenarios are crucial to the success of AWD. Additionally, researchers from remote sensing and computer science have primarily focused on applications in the operations and maintenance (O&M) phase, while publications from the C&BT domain span the entire building lifecycle, from design and construction to O&M.

The input data sources can be classified into four categories: 2D (e.g., RGB camera, drawings), 3D (e.g., LiDAR point clouds with or without RGB information, and BIM), thermal (infrared thermography [IRT]), and multi-source data (e.g., combinations of IRT and RGB camera). Among these, the most frequently used data sources were LiDAR and RGB cameras, which supported a range of applications from semantic enrichment modeling to energy simulation. Drawings and BIM, as manually created data, were primarily utilized for semantic enrichment modeling in the O&M and design phases [28,35].

The majority of attribute detection focused on the shape, position, and spatial dimensions of windows. Additionally, seven papers focused on detecting the status of windows to predict energy consumption or heat usage. Three papers focused on categorizing windows by type to enhance the semantic enrichment of BIM.

The outcomes of AWD were integrated with 3D reconstruction techniques, such as shape grammar and scan-to-BIM, resulting in a 3D model. 2D polygons are less utilized in the storage stage, with only limited applications. Raw data formats were exclusively used for storage purposes in 30 articles, as the primary requirement was to infer higher-level data from the existence of windows, such as building floors, point cloud generation, and energy monitoring [45,59,60].

In the execution layer, window detection methods included 55 studies using rule-based approaches, while 50 studies employed deep learning. The majority of windows on building facades are characterized by a high degree of repeatability and symmetry. Therefore, the predominant and more established method utilizes these rules to detect the existence of windows, which can be applied to various data types based on the extracted primitives (e.g., lines, patches) [26,61]. Deep learning methods can be categorized into instance segmentation, semantic segmentation, localization, classification, and generative models, depending on the data inputs and application scenarios of AWD. The main advantage of deep learning is that it can recognize objects in scenes in cases such as occlusion and irregular windows [45,51].

The application of AWD encompasses six topics in the requirement layer: building information extraction, heritage conservation, robotics, damage simulation, semantic enrichment, and energy simulation. The majority of applications were related to semantic enrichment for 3D building models (totaling 50), which was addressed in both the design phase and the O&M phase. The remaining applications pertained to the O&M phase, except for robotics, which was exclusively utilized in the construction phase. This suggests that AWD applications play a crucial role in the O&M phase compared to other phases. Energy simulation was the second most prevalent type of application, with a total of 25 papers, which is reasonable given that the presence of windows significantly influences the thermal performance of buildings [57].

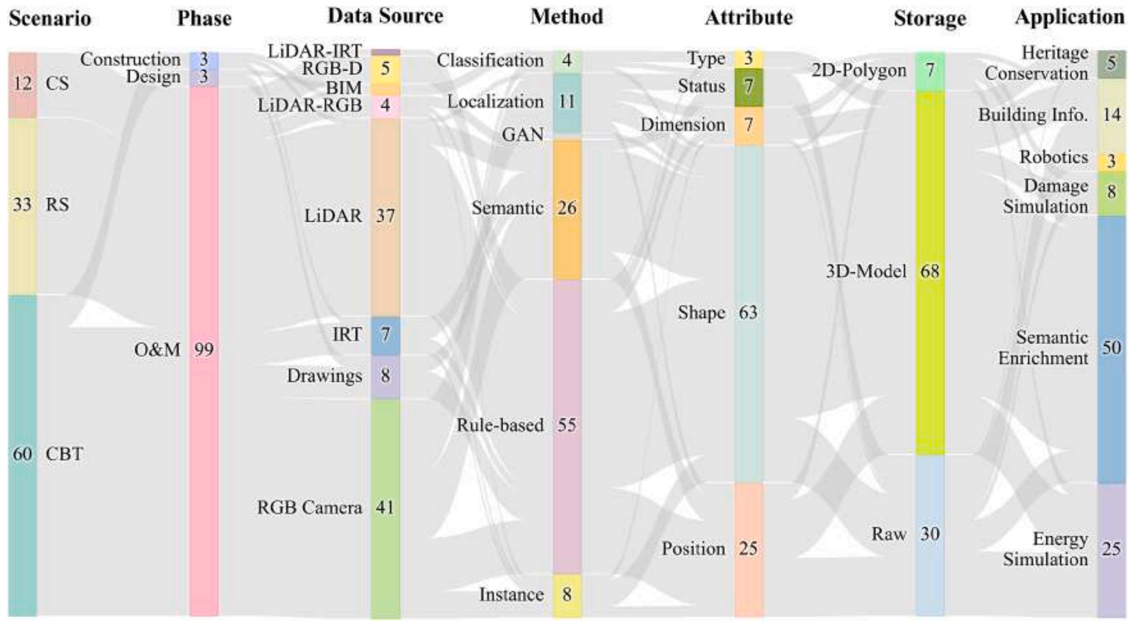

Fig. 5. Sankey diagram of the relationships between Scenario, Phase, Application, Data Source, Method, Attribute, and Storage.

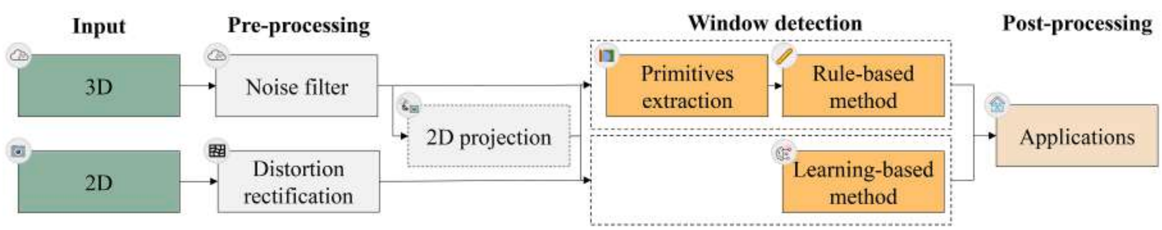

Fig. 6. General workflow of AWD.

3.2. Generalized workflow of AWD

Fig. 6 shows the general workflow of AWD. Before detection, pre-processing steps are essential to tailor the process to the specific attributes of interest. For rule-based methods, the pre-processing of the 2D and 3D data sources differs slightly. 2D images, typically captured from a human perspective of street view, require perspective correction to enhance the accuracy of AWD results [60]. 3D-based methods commonly involve noise filtering and plane detection as standard pre-processing procedures [62]. Some 3D methods may also integrate 2D projection since 2D-based methods are more mature. After pre-processing, the window detection algorithm is applied. Rule-based methods typically extract primitives such as lines and patches, while learning-based methods directly utilize the preprocessed data. Once the windows are extracted, they are further processed or analyzed to satisfy specific application requirements, such as 3D model reconstruction or window-to-wall ratio calculation.

3.3. Current trends

Fig. 7a illustrates the trend of AWD applications in the requirement layer. It is evident that AWD's use in semantic enrichment modeling, including BIM and CIM, has consistently attracted the most attention in the sector. This was primarily because semantic enrichment models provide significant value for subsequent research, such as energy simulation and building asset management [63,64]. The growth trajectory of semantic enrichment modeling slowed after 2018, while energy simulation has exhibited a more obvious upward trend. This trend shift underscores the increasing emphasis on the urban thermal effect in current literature. The remaining applications, such as damage simulation, heritage conservation, and construction robotics, have remained relatively limited in number. The diversity of these application types highlights the significant potential of AWD within the AEC industry.

The trend of the input data sources for AWD is illustrated in Fig. 7b. LiDAR point clouds have been the most popular data source for semantic enrichment modeling since the first paper in 2009, due to their ability to capture the three-dimensional characteristics of windows. Prior to 2022, the use of RGB cameras in AWD applications was less prevalent compared to LiDAR point clouds. However, after 2022, there was a significant increase in the use of RGB cameras, as evidenced by a rise in related applications, ranging from semantic enrichment modeling to energy simulation. Since 2021, IRT has been utilized for energy simulation, as reported in up to seven papers. In contrast, the use of drawings, although considerably niche, has remained steady at approximately one publication per year since 2014.

The yearly statistics and trend of applied methods, i.e., rule-based, semantic segmentation, instance segmentation, localization, classification, and generative adversarial networks, are demonstrated in Fig. 7c. Before 2018, rule-based methods were the predominant method. For instance, techniques such as utilizing symmetry and repetitiveness to identify buildings in aerial images [65], employing top-down rules inference to parse facade components [66], and detecting holes and edges of windows based on extracted primitives [49] were commonly adopted. Rule-based methods continue to be widely used after 2018, but focus more on 3D input data sources (19 out of 28). This shift was attributed to the scarcity and high cost of annotations of 3D training datasets for deep learning, in contrast to the abundance of publicly available 2D datasets.

Since 2018, with the advancement of computer vision technology, the application of deep learning in AWD has increased significantly.

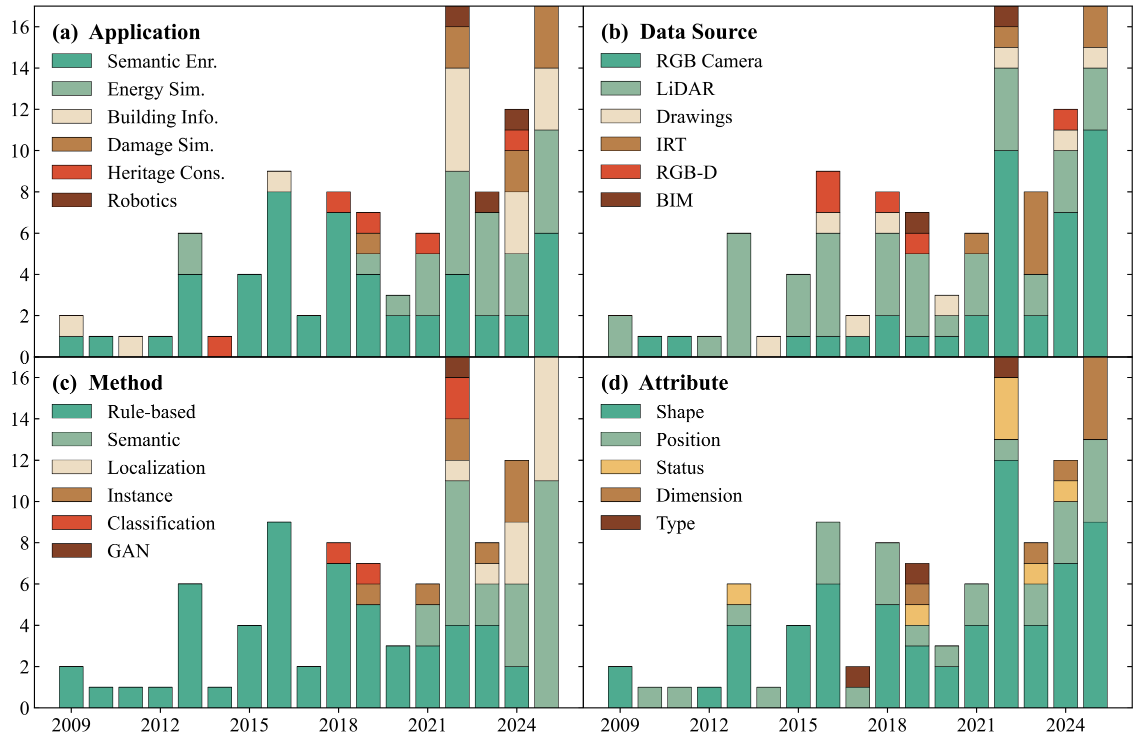

Fig. 7. Trend of 105 publications from 2009 to 2025. (a) Application, (b) Data Source, (c) Method, (d) Attribute.

Initially, the classification methods, such as 2D convolutional neural networks, were employed to identify and classify unknown objects as windows [28]. Semantic segmentation methods emerged as the mainstream approach since 2021 alongside rule-based methods, because they provide accurate pixel-level window detection results and have been applied in 26 studies [45]. In summary, there has been an increased interest in utilizing deep learning to detect windows using 2D input data sources.

Fig. 7d demonstrates the trend of attributes for AWD. Despite the continued popularity of 3D data sources for AWD, research focused on detecting window dimensions has been relatively scarce in the past five years. Only six papers were published during this period, addressing 3D thermal modeling, indoor building quality measurement, and energy simulation [48,64,67]. Shapes and positions have consistently been the primary attributes in this field. These findings underscore the inherently two-dimensional characteristics of windows in urban buildings.

In summary, AWD is being increasingly applied in building life-cycle management, consistent with the concept of DTB. The growing number of related publications and the diversity of application trends underscore the potential for developing novel applications of AWD. Input data sources are becoming increasingly diverse. Learning-based methods are gradually supplanting rule-based methods in the execution layer.

4. Guidelines for AWD and further application

4.1. A three-step guideline

Fig. 8 presents the three-step guidelines for AWD proposed in this paper. The guidelines were developed primarily based on the technical workflows reviewed in the 105 publications. Additional moderation was made by the authors to enhance clarity and to incorporate complementary aspects that were insufficiently addressed in the existing literature.

Step 1: To confirm the data and information 'meat' layer to support the target application. The data and information 'meat' layer should effectively (correctly and accurately) and efficiently (less cost and waste) support the top application layer, according to the proposed sandwich model. To support a target application scenario, one should carefully determine the three key factors–namely, data source, attributes, and storage format–within the data and information layer. Three decision trees for guiding the selection of the factors are illustrated and explained as follows.

Fig. 9 illustrates the process of selecting an input data source. The primary recommended input data sources include RGB cameras, LiDAR point clouds, and drawings. LiDAR point clouds are primarily recommended for applications that require highly accurate 3D positioning and geometric information in the construction and O&M phases [67–69]. RGB cameras are recommended for energy simulation, damage assessment, and building information extraction, especially during the O&M phase. They provide high-resolution, vivid color images that can capture a window's real-time position, shape, and status. Therefore, for applications that only require calculating window-related indicators or directly classifying window status, 2D imagery is a cost-effective choice [23,59]. One is recommended to use BIM, drawings, or IRT for updating the semantic information of BIM components, semantic enrichment modeling during the design phase, and heat monitoring, respectively, as discussed in the literature [28,70,71].

Fig. 10 illustrates the window attributes to be prioritized in AWD. The typical detection attributes include shape, position, and status. Detecting windows' 2D/3D shapes is sufficient and recommended for most applications, as indicated by the default 'otherwise' branches. There are two main exceptions. In energy simulation applications, status is employed to calculate window opening behaviors [72] and shape is used to estimate building energy consumption [73]. For heritage conservation during the O&M phase and semantic enrichment modeling at the design stage, detection of window position is recommended. However, if 3D reconstruction is required, shape detection should be prioritized [74]. Additionally, if higher-level semantic information is needed (e.g., window type), other studies should be consulted [22,37,75].

Fig. 11 displays the selection of a storage plan for the AWD system. Since 2009, the primary storage format has been 3D models, as indicated by the default 'otherwise' branches in Fig. 11. This trend is expected to

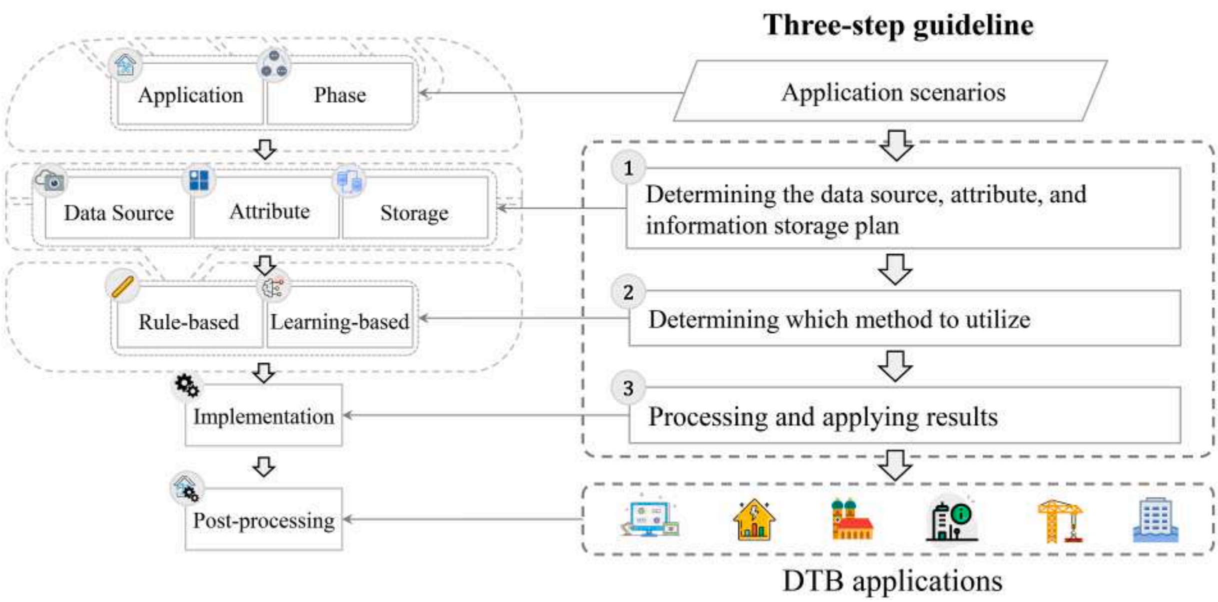

Fig. 8. Three-step guidelines for selecting *a posteriori* AWD systems.

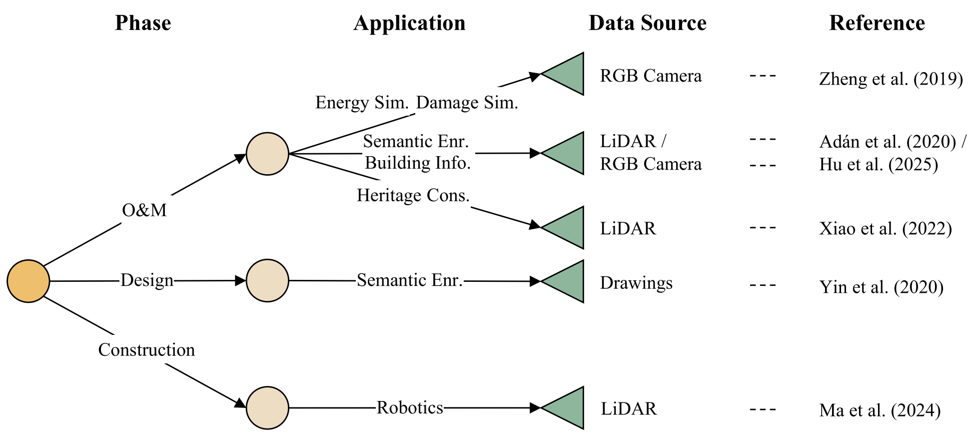

Fig. 9. Decision tree for AWD data source selection.

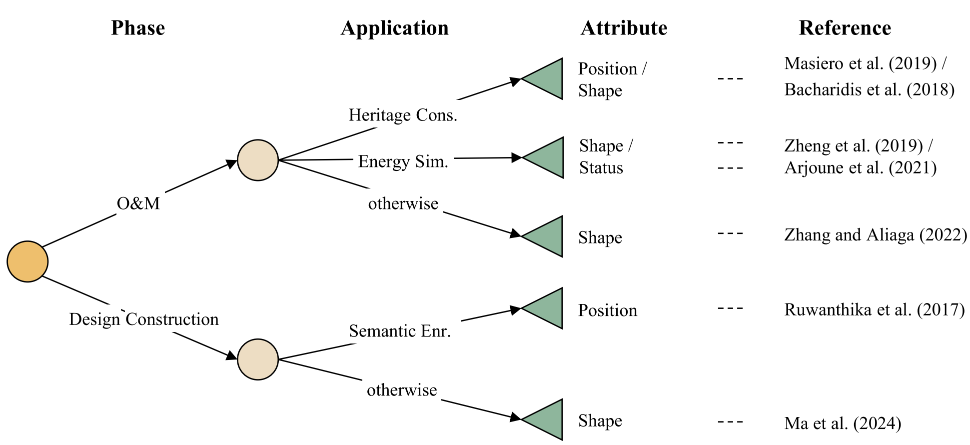

Fig. 10. Decision tree for AWD attributes selection.

continue due to the extensive application support offered by semantically rich BIM and CIM models [25]. Additionally, 2D polygons and raw formats are recommended in damage simulation, robotics, and energy simulation applications. For example, flood risk possibilities can be precisely calculated based on extracted window 2D polygons [9]; real-time robot movement paths can be detected [27]; and window opening behavior can be monitored using raw results [76].

Step 2: To build the 'base' execution layer. The method in the execution layer can be determined based on the chosen data source and storage plan in Step 1. The guidelines for this step notably incorporate

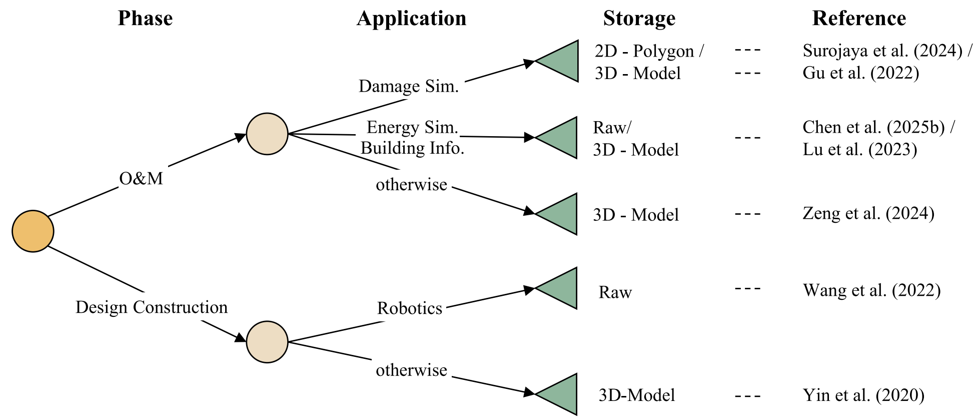

Fig. 11. Decision tree for AWD storage selection.

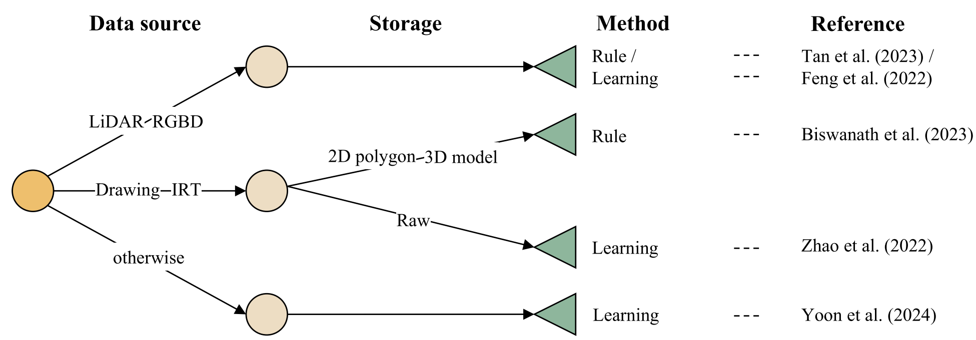

Fig. 12. Decision tree for AWD method selection.

the authors' moderation. In the existing studies, rule-based methods account for 55 out of 105, which is more than 52%. However, learning-based methods have pushed the boundaries of computer vision [77] and have become increasingly popular in the AEC industry over the past five years (47 out of 63), as shown in 7c. Therefore, the guidelines below give more credit to the learning-based methods.

Fig. 12 shows the decision tree for the method to execute. Based on the data source and storage plan, the default recommendation method is learning-based. The recommendations for data sources of drawings and IRT vary according to the storage plan. More specifically, the use of rule-based methods was advocated for both 2D polygon and 3D models, whereas for raw format, the recommendation was to utilize learning-based methods. The recommended methods for LiDAR and RGBD cameras include both rule-based and learning-based. The final choice between rule- or learning-based methods may vary depending on the availability of training datasets (e.g., Ma et al. [27], Wei et al. [78]) and the complexity of the input data sources.

In general, rule-based methods offer a practical choice when there is a limited amount of training data. The edges of windows have stronger image gradients compared to other facade elements, which can be utilized as a 2D-based data source [79]. Besides, pattern recognition based on similarity and symmetry can effectively detect windows [65]. Using rule-based methods with 3D-based inputs is more challenging than with 2D-based inputs. This is because the invalid returns and reflections of LiDAR data result in holes in the scanned facades [80]. Nevertheless, several rule-based methods can still be applied to detect windows in 3D data. For example, splitting the 3D point cloud into slices to estimate the boundary points of windows [81], utilizing ray casting [82] to identify open areas in 3D space [83], and projecting 3D data onto 2D image spaces followed by applying edge detection algorithms [84] to identify window edges [52].

Learning-based methods have been well-proven for 2D data sources. The existing AWD methods can be categorized into localization, semantic segmentation, and instance segmentation, ranging from coarse to fine. The localization-based methods (e.g., Faster-RCNN [85], YOLO [86]) can simultaneously provide the window category [22] and its position(i.e., bounding box) on the facade. However, such methods are insufficient for representing the detailed geometry of windows. For example, if the facade images are affected by perspective projection, rectangular windows may be represented inaccurately. The main reason is that the parallel edges of windows undergo perspective projection and intersect at an invisible vanishing point [87]. In contrast, semantic segmentation-based methods (e.g., U-Net [88], Deeplab [89]) can more precisely estimate the pixels belonging to windows, though they may fail to distinguish between individual adjacent windows. By combining the above two approaches, instance segmentation-based methods (e.g., Mask-RCNN [90], SOLOv2 [91]) not only estimate the detailed geometry of windows but also identify distinct window instances by parsing the window mask image.

The further selection for a specific method from the learning-based category varies with the level of requirements and window attributes. Localization-based methods are suitable for studies that need to monitor the status of window openings (e.g., open, closed, or half-open) [72] and analyze the relationship between window opening behavior and environmental factors [23]. Semantic segmentation-based methods are viable approaches to address tasks such as calculating the window-to-wall ratio, which do not require explicit estimation of the position and shape of windows [26]. However, if one wants reverse engineering on the facade [92], instance segmentation-based methods would be a better choice [56,93]. Similarly, for both vectorized and rasterized drawings, instance segmentation-based methods [94–96] can effectively provide accurate information on the position and geometric details of window symbols.

Step 3: To implement and apply the constructed AWD. The third step is to realize and apply the AWD workflow. The primary principle of sensor placement is to obtain the best coverage of target windows, such as capturing adequate lighting for the RGB camera and conducting multiple scans for the LiDAR. Sensor carriers encompass vehicles for large-scale low-rise building windows; drones for large-scale outdoor windows; Internet of Things devices for monitoring window/building status; and surveyors for individual buildings/scenes data collection [67,97,98]. The dataset for training and execution can be obtained either by utilizing off-the-shelf datasets or by annotating using tools, such as labelImg [72].

The execution layer's learning-based methods can often be obtained from official algorithm repositories, such as GitHub or HuggingFace, which typically contain pre-trained learning models and auto-pipelines. Training for 2D data sources can be conducted locally because of the relatively low computational cost, whereas LiDAR-based training is recommended to utilize online GPU rental services to accommodate higher computational demands. One should carefully examine the target attributes before selecting a performance metric from classification metrics (primarily for instance segmentation), mean Intersection over Union (for semantic segmentation), and error metrics (commonly reported in rule-based methods).

Applying AWD results with post-processing DTB applications may need further transformation. For heritage conservation and semantic enrichment modeling, windows function as digital assets directly [73]. In energy simulation, raw window data facilitate the calculation of window-to-wall ratios by orientation and climate-based daylight metrics [99]. Some scenes parsing applications convert detected results into graphs to incorporate human knowledge [65] or classify room types [100]. Therefore, the detected results must be further transformed according to the final intended use. The detected results can be stored on local/cloud disks or in BIM based on the transformed results.

4.2. Validation

Table 2 presents the validation results for the four decision trees. The validation set shown in Table A.2 comprised 21 AWD papers from Scopus but excluded by Web of Science; so the validation set contained no items in the training data. The 21 papers pertain to the DTB life-cycle phases, too. The overall accuracy was satisfactory, with a mean accuracy of 88.1%, precision of 92.1%, and F1-score of 80.3%. Therefore, the proposed guidelines and decision trees are reliable and robust for practitioners conducting AWD and DTB applications.

Table 2
Validation results of the guidelines consisting of four decision trees on 21 more AWD papers from Scopus.

Metric	Data Source	Attribute	Storage	Method	Average
Accuracy (%)	85.7	81.0	85.7	100.0	88.1
Precision (%)	88.9	88.9	90.7	100.0	92.1
Recall (%)	92.3	71.4	60.0	100.0	80.9
F1-score (%)	89.0	73.8	58.4	100.0	80.3

The majority of cases were correctly classified. Some focused on building energy in the O&M stage across five DTB applications. For example, for enhanced building energy simulation, Kim et al. [101] utilized aerial imagery to detect windows and store them as LoD-3 3D-Model, as shown in Fig. 13d.

In data source selection, three cases mismatched our guidelines. Two papers aimed at building information extraction in the construction phase, and one for heritage conservation in the O&M phase [102–104].

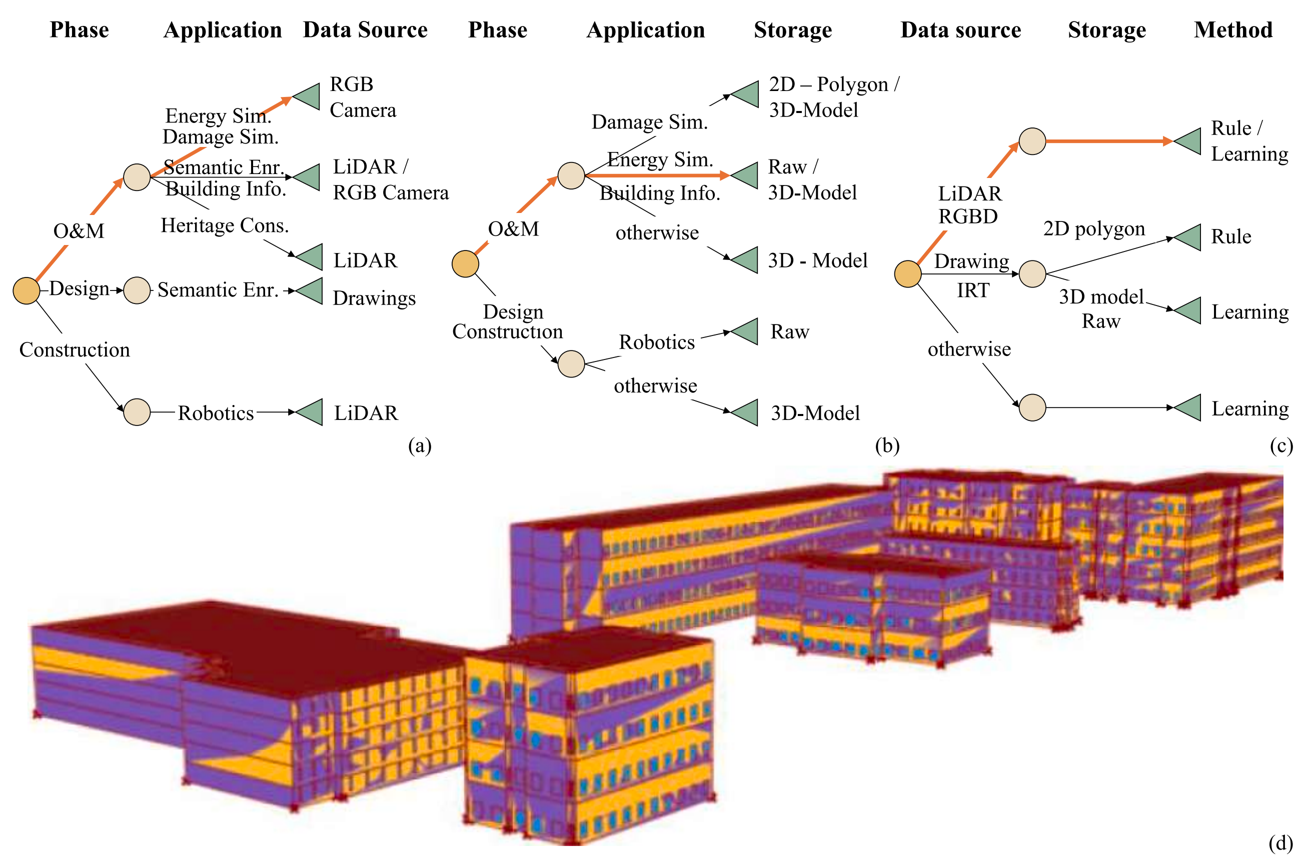

Fig. 13. Demonstrative case of decision trees. (a) decision tree for data source; (b) decision tree for storage; (c) decision tree for method; (d) final storage illustration of building energy model [101].

A reason was the training data had no camera-based cases for construction monitoring. This gap highlights the diversity, growing demand, and importance of AWD.

The decision tree for attribute selection achieved the lowest accuracy (81.0%) and F1-score (58.4%), primarily because the position information derived from localization-based methods remains relatively infrequent across the 105 publications, especially in energy simulation [101,105]. Further, three misclassified cases in storage selection predominantly involved semantic enrichment in the O&M phase, damage simulation in the O&M phase, and building information extraction in the construction phase [102,106,107]. These misclassified cases, nevertheless, provide valuable examples that enrich the potential data sources, attributes, and storage options available to non-expert layman users.

5. Discussion

5.1. Comprehensive review of AWD

5.1.1. General observations

The diversity of DTB application categories does not inherently reflect a corresponding level of innovation in AWD techniques. Specifically, most studies have focused on the end-use aspects of DTB applications. However, AWD methodologies are often repeatedly applied or yield only incremental advancements.

Input data sources are typically limited to unimodal and small-scale datasets. Most studies primarily utilize a single modality for feature extraction, with any additional sources serving only as supplementary information, although multiple data sources are discussed in the literature. Furthermore, input data are frequently limited to small-scale samples, rather than being applied at the scale of urban buildings.

The literature predominantly emphasizes attributes intrinsic to the window itself, rather than contextual information, such as the associated unit, floor level, or material properties. The predominant storage format is three-dimensional, which is favored due to its portability and versatility. In addition, several proprietary models, such as building energy models augmented with window-based energy simulation data, have been developed to address specific, niche applications [64].

Deep learning, as an emerging technology, is increasingly replacing traditional rule-based methods. These learning-based approaches are primarily supervised and necessitate carefully annotated datasets. Additionally, it is noteworthy that the execution methods utilized in AWD tend to lag behind the most recent advancements in computer science, such as vision-language models.

The validation of detected windows is contingent upon the execution methods utilized and the role of windows within DTB applications. Notably, 29% of studies did not report window-level evaluation metrics, indicating that window detection is frequently regarded as a secondary outcome in DTB applications. Nevertheless, AWD results continue to impact the assessment of overall DTB application performance, as reflected by metrics such as the root-mean-square error of model distance in semantic enrichment modeling.

5.1.2. Challenges

This section highlights the common issues and challenges identified in AWD across the reviewed studies, as informed by the proposed conceptual model.

1) In the Requirement layer: Limited to individual buildings, lack of cross-style detection, and dependence on domain expertise.

   The DTB applications of AWD are mainly focused on individual buildings. However, comprehensive window digital assets at the city scale are required to support data-driven analysis and resident-centric applications at a fine-grained level. Moreover, few studies have investigated window detection across different building types and regions. The creation of AWD relies on domain expert knowledge, semantic annotation, and algorithm execution, which increases the challenge of broadly adopting AWD in DTB applications.

2) In the Data and Information layer: Lack of multimodal data fusion, limited 3D dataset availability, low reusability of detected results, and insufficient contextual information regarding the window.

   The detection accuracy of AWD is limited due to its reliance on a single input data source and corresponding limitations (e.g., occlusion in 2D, glass reflection in 3D, and distortion in photorealistic CIM). The existing literature does not sufficiently address window detection through data fusion [27]. Additionally, current research primarily focuses on AWD utilizing 2D input; however, there is a lack of available datasets for 3D semantic segmentation of windows using deep learning techniques. The reusability of detected AWD results should be improved to facilitate the broader adoption of AWD. Finally, the absence of supplementary contextual information about the window, such as the associated unit and floor, poses a challenge for accurately locating the window within the DTB application.

3) In the Execution layer: Low robustness and applicability across window types and dataset sizes, lack of zero- or few-shot learning, and high computational costs.

   The robustness of execution methods requires improvement. For instance, both rule-based and learning-based approaches struggle to reliably extract windows when window features (e.g., coordinates, color) are incomplete. The applicability of these methods is further challenged by small sample sizes and imbalanced data distributions, such as windows in rural areas or with uncommon shapes and styles. Furthermore, reducing the amount of required training data and minimizing dependence on domain expertise are necessary for advancing DTB applications. Finally, current methods are limited in terms of computational efficiency and power consumption in window detection, which constrains the deployment of large-scale window detection platforms, such as drones.

5.1.3. Future directions

Fig. 14 illustrates three groups of future directions for AWD, based on the findings from the literature analysis, practical guidelines, the conceptual model, and the identified challenges.

- **Mapping emerging needs and demands**: The influence of windows on urban residents is increasing due to the rapid and vertical expansion of urban environments. Drones are expected to significantly affect the daily lives of city dwellers in the near future [108]. Window views, noise, privacy perception, and real-time instance-level window operation represent potential DTB applications aimed at measuring and enhancing residents' comfort [109–111]. In addition, building facade health monitoring and renovation also depend on the digitization of windows [60]. These emerging directions can broaden the requirements layer and further expand DTB applications.
- **Cross-style window detection**: Windows across different building types, such as commercial and rural buildings, represent feasible detection targets. For instance, commercial buildings are typically characterized by glass curtain walls with small openable panels. In contrast, rural buildings may exhibit irregular window patterns that are challenging for current execution methods to detect. Accurate detection of windows across diverse styles will enhance DTB applications, including assessments of energy consumption and indoor thermal insulation [112].
- **Crowd-sourced city-wide cloud platform of windows**: Although CityGML has established standards for various levels of detail (LoD) in building modeling, these standards are difficult to modify and have limited widespread adoption and application [113]. Advanced technologies for semantic mapping, geometry reconstruction, and specialized software integration can be incorporated into such cloud platforms. Additionally, data sharing protocols will facilitate large-scale integration and reusability of AWD results, promoting the adoption of DTBs throughout the entire building lifecycle.
- **Multi-source, multi-modal data fusion**: Both 2D and 3D window features offer unique advantages. For instance, 2D data captures con-

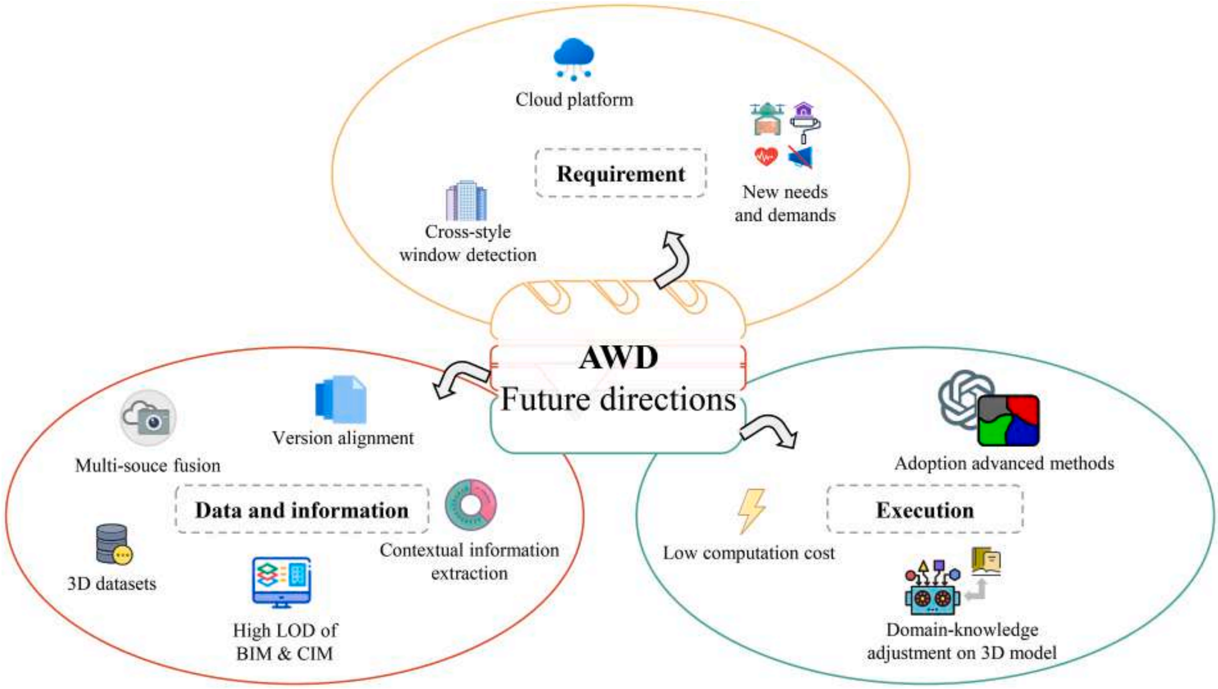

Fig. 14. Potential future directions of AWD.

tinuous and complete glass surfaces, while 3D data provides distinct three-dimensional characteristics, such as window sills and frames. Integrating both modalities through deep learning and 3D Gaussian splatting may represent an effective approach for enhancing window detection.

- **Version alignment of data sources and window assets**: The same window assets across different data sources (e.g., photogrammetry models from Google Earth and governmental databases) may exhibit discrepancies, such as incomplete features in one source while being complete in another. Aligning the detected window versions from various platforms can yield more comprehensive and accurate window detection results.
- **Encourage open-source 3D benchmark datasets**: High-quality 3D benchmark datasets for window detection are needed. Such datasets will facilitate baseline model training and the evaluation of execution methods, thereby supporting the development of new algorithms and reducing the learning curve for non-experts.
- **Reconstruction of CIM and BIM models with high LoD**: High LoD CIM or BIM models, with window details, serve as the foundational geometry of DTB applications [114]. Moreover, in BIM's automated modeling, the low LoD of building components hampers building simulations across the entire lifecycle [115]. Thus, developing automated and cost-effective methods for high LoD CIM/BIM modeling is necessary to promote AWD and DTB applications.
- **Extraction of contextual information for windows**: Different building data sources encompass a range of information, from intrinsic attribute data to external contextual information. For example, photorealistic CIMs represent the up-to-date and vivid appearances of building facades. In contrast, architectural drawings depict indoor layouts but are less digitized and may be misaligned with the constructed buildings. The automated integration of window information based on indoor and outdoor data sources (such as drawings and photorealistic CIM) represents a promising research avenue for achieving comprehensive semantic enrichment of windows.
- **Adoption of advanced window detection methods**: New deep learning models are emerging with the development of computer vision. The hybrid method can consist of the following potential directions to improve the performance of window detection results: (i) Utilizing up-to-date 2D semantic segmentation models, such as Segment Anything [116]; (ii) Introducing vision-language models to assist in identifying errors in post-data processing steps, because of the multi-modality nature and large-scale pretraining on existing human digitalized knowledge; (iii) Utilizing interactive labeling to implement few-shot learning to dynamically improve the model's segmentation capabilities, given the similarity of windows on a single building [117].
- **Improvement of window detection methods based on domain knowledge**: The 3D deep learning model architecture can be improved by domain-characteristic-based adjustments, such as loss function design based on symmetry and repeatability, synthetic data generation and training, and contrastive learning to enhance 3D scene understanding [118]. Additionally, 3D scene completion can be used to mitigate the effects of occlusion and glass reflections.
- **Development of low-cost, low-power algorithms**: The large-scale window detection requires devices such as drones to collect city-wide window information. Considering the high computational cost of current execution methods, algorithms with low computational cost and power consumption are needed.

5.2. *The sandwich model and guidelines*

5.2.1. *Implications*

The sustainable development of building life-cycle projects is increasingly emphasized across many fields. Window detection plays a crucial role in various building life-cycle activities (e.g., design, construction, and management) that align with the Sustainable Development Goals (SDGs). The growing number of new buildings underscores the necessity of Automated Window Detection (AWD) at both the microscale (e.g., indoor facilities management) and the macroscale (e.g., urban energy consumption), with attributes being essential for both indoor and outdoor windows.

A systematic analysis of the current literature was conducted based on coded classification results, general workflows, and emerging trends. This paper further proposes a three-step guideline for newcomers aiming to develop applications based on AWD results. Four decision trees–covering data sources, attributes, storage plans, and methods–demonstrate suitable options for users addressing specific application tasks. This guideline, combined with the conceptual model, reveals potential connections between the three layers and seven components of application-driven AWD. It also provides new insights relevant to fields such as remote sensing, building engineering, building energy consumption, and urban planning.

5.2.2. Limitations and future work

This paper has several limitations that should be addressed in future research. First, the literature review covers publications only up to December 2025. Given the rapid evolution of large models, regular updates–ideally every five years–are recommended to keep pace with advancements in the field. Second, the literature search was limited to the Web of Science and Scopus databases, which may introduce selection bias. Third, only papers written in English were included, potentially overlooking significant contributions from non-Anglophone research communities. Although the applications span the entire building life cycle, there is a disproportionate focus on studies related to the Operation and Maintenance (O&M) phase, which limits insights into the integration of design and construction phases using the proposed decision trees. The impact of rapid technological advancement on the conceptual model and guidelines needs to be assessed. Future research should validate the robustness and accuracy of the proposed generic conceptual model and guidelines across diverse building life-cycle scenarios and applications in different regions.

6. Conclusion

The digitalization of building assets provides a foundational platform for enabling applications throughout the building lifecycle in alignment with the Sustainable Development Goals (SDGs). Windows, as integral elements of the building envelope, play a pivotal role in functions such as air exchange and heat transfer. Consequently, accurate window detection is essential for deriving high-level building information, estimating energy consumption, and monitoring human activities, among other applications. Although numerous studies have explored advanced applications that leverage automated window detection (AWD) results, the nature of the relationship between these applications and AWD remains insufficiently addressed.

Therefore, this paper establishes a clear definition of automated window detection (AWD) and introduces a conceptual 'sandwich' model comprising three layers–requirement as 'top', data and information as 'meat', and execution as 'base'–and six components: application, phase, data source, attribute, storage, and method. Utilizing the PRISMA standard and the proposed 'sandwich' model, 105 relevant publications were systematically collected, analyzed, and evaluated to assess a *posteriori* status and trends of each component. Furthermore, a set of three-step guidelines is presented to assist prospective users: (1) selecting the input data source, attributes of interest, and information storage strategy; (2) determining the appropriate method to employ; and (3) processing and transforming the results. The four decision trees achieved an average accuracy of 88.1% in the validation set, pinpointing the reliability and robustness of our proposed guidelines and decision trees.

To the best of our knowledge, this paper presents the first comprehensive review of AWD applications in the AEC-related industry and provides full coverage of the literature from 2009 to 2025. It delineates the current status, prevailing trends, and key developments that inform the selection of critical factors for AWD-based applications. The proposed three-step guidelines offer a practical framework for future researchers and industry practitioners seeking to implement AWD-based solutions. Recommended future research directions include: investigating novel application domains, enhancing window detection across diverse building typologies, developing cloud-based platforms, integrating multi-source data fusion, harmonizing heterogeneous data versions, promoting the development of 3D public datasets, refining high LoD CIM and BIM models, expanding window information repositories, adopting innovative detection methodologies, and advancing 3D learning through domain-specific knowledge. This review is limited to literature published up to 2025 and relies exclusively on sources indexed in the Web of Science and Scopus databases and written in English. It is recommended that systematic reviews be conducted every five years to capture the evolving landscape of AWD research and validate its applicability across various stages of the building life cycle.

CRediT authorship contribution statement

Siyuan Meng: Writing – original draft, Visualization, Software, Methodology, Data curation, Conceptualization; **Longyong Wu:** Writing – original draft, Visualization, Software, Data curation; **Maosu Li:** Writing – review & editing, Visualization; **Anthony G. O. Yeh:** Resources, Project administration, Writing - review & editing; **Fan Xue:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Data availability

No data was used for the research described in the article.

Declaration of interests

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.

Declaration of generative AI in writing

The authors declare the following ways of interactions with generative AI which may be considered as potential uses in writing or editing the content in this paper:

GenAI-assisted proofreading. A university self-hosted GenAI LLM (ChatGPT-4.1) was used to assist with proofreading the work to eliminate grammatical and connection errors in the main text, and to improve readability of Abstract, Introduction, and Conclusion sections. We declare no use of GenAI to produce any new content in the work.

Acknowledgments

The work presented in this paper was supported by the Hong Kong Research Grants Council (Nos. 17201325, T22-504/21-R), and in part by the Department of Science and Technology of Guangdong Province (GDST) (No. 2020B1212030009).

Appendix A. Literature database

Table A.1
List of 105 cases of application-driven AWD.

No.	Reference	Scenarios*	Phase†	Application‡	Data Source§	Method	Attribute	Storage
1	Becker [66]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
2	Chung et al. [65]	C.S.	O&M	Building Info.	LiDAR	Rule	Shape	3D-Model
3	Meixner and Leberl [53]	R.S.	O&M	Semantic Enr.	RGB Camera	Rule	Position	Raw
4	Meixner and Leberl [119]	C.&B. T.	O&M	Building Info.	RGB Camera	Rule	Position	3D-Model
5	Truong-Hong et al. [120]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
6	Demisse et al. [121]	C.S.	O&M	Energy Sim.	LiDAR	Rule	Status	Raw
7	Dumitru et al. [122]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Position	Raw
8	Truong-Hong et al. [123]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
9	Tuttas and Stilla [21]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
10	Wang et al. [124]	C.&B. T.	O&M	Energy Sim.	LiDAR	Rule	Shape	3D-Model
11	Xiong et al. [83]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
12	Riedinger et al. [125]	C.S.	O&M	Heritage Cons.	Drawings	Rule	Position	3D-Model
13	Frommholz et al. [126]	R.S.	O&M	Semantic Enr.	RGB Camera	Rule	Shape	3D-Model
14	Hong et al. [127]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
15	Liu and Qin [128]	R.S.	O&M	Semantic Enr.	LiDAR-RGB	Rule	Shape	2D-Polygon
16	Sadeghi et al. [129]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
17	Arnaud et al. [130]	C.S.	O&M	Semantic Enr.	RGB-D	Rule	Position	3D-Model
18	Colleu and Benitez [131]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Position	3D-Model
19	Díaz-Vilariño et al. [132]	R.S.	O&M	Building Info.	LiDAR	Rule	Shape	3D-Model
20	Gimenez et al. [133]	C.&B. T.	O&M	Semantic Enr.	Drawings	Rule	Position	3D-Model
21	Jung et al. [134]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
22	Lachat et al. [135]	R.S.	O&M	Semantic Enr.	RGB-D	Rule	Shape	3D-Model
23	Nguatem et al. [136]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
24	Paravolidakis et al. [137]	C.S.	O&M	Semantic Enr.	RGB Camera	Rule	Shape	3D-Model
25	Valero et al. [52]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
26	Oskouie et al. [22]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Rule	Type	3D-Model
27	Ruwanthika et al. [35]	C.&B. T.	Design	Semantic Enr.	Drawings	Rule	Position	2D-Polygon
28	Bacharidis et al. [74]	R.S.	O&M	Heritage Cons.	RGB Camera	Rule	Shape	3D-Model
29	Chen et al. [138]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	Raw
30	Franz et al. [139]	C.&B. T.	O&M	Semantic Enr.	RGB-D	Rule	Shape	3D-Model
31	Hao et al. [49]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Position	3D-Model
32	Jung et al. [140]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
33	Lu et al. [75]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Classification	Shape	3D-Model
34	Previtali et al. [141]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Position	3D-Model
35	Wu et al. [142]	C.&B. T.	O&M	Semantic Enr.	Drawings	Rule	Position	Raw
36	Cui et al. [143]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
37	Kim et al. [28]	C.&B. T.	Design	Semantic Enr.	BIM	Classification	Type	3D-Model
38	Masiero et al. [68]	R.S.	O&M	Heritage Cons.	LiDAR	Rule	Position	Raw
39	Shi et al. [62]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
40	Tang et al. [144]	R.S.	O&M	Semantic Enr.	RGB-D	Rule	Shape	3D-Model
41	Van Ackere et al. [51]	R.S.	O&M	Damage Sim.	LiDAR-RGB	Instance	Dimension	2D-Polygon
42	Zheng et al. [79]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Rule	Status	Raw
43	Adán et al. [50]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
44	Dochev et al. [36]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Rule	Position	3D-Model
45	Yin et al. [70]	C.&B. T.	Design	Semantic Enr.	Drawings	Rule	Shape	3D-Model
46	Arjoune et al. [73]	R.S.	O&M	Energy Sim.	IRT	Instance	Shape	Raw
47	Fan et al. [145]	R.S.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
48	Pexman et al. [61]	R.S.	O&M	Heritage Cons.	LiDAR	Rule	Position	3D-Model
49	Touzani et al. [97]	R.S.	O&M	Energy Sim.	RGB Camera	Semantic	Shape	3D-Model
50	Voříšek et al. [146]	C.S.	O&M	Semantic Enr.	LiDAR	Rule	Position	3D-Model
51	Wang et al. [147]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Semantic	Shape	Raw
52	Feng et al. [9]	C.&B. T.	O&M	Damage Sim.	LiDAR	Instance	Shape	3D-Model
53	Golombek and Marshall [54]	R.S.	O&M	Building Info.	LiDAR	Rule	Shape	3D-Model
54	Gu et al. [55]	C.&B. T.	O&M	Damage Sim.	RGB Camera	GAN	Shape	3D-Model
55	Jarząbek-Rychard and Maas [148]	R.S.	O&M	Semantic Enr.	RGB Camera	Instance	Shape	3D-Model
56	Jiang et al. [149]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Semantic	Shape	3D-Model
57	Li et al. [150]	C.&B. T.	O&M	Energy Sim.	IRT	Semantic	Shape	Raw
58	Luong et al. [31]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Rule	Status	Raw
59	Murtiyoso et al. [59]	R.S.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
60	Pantoja-Rosero et al. [151]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Semantic	Shape	3D-Model
61	Sun et al. [23]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Classification	Status	Raw
62	Szcześniak et al. [26]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Rule	Position	3D-Model
63	Tien et al. [57]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Status	Raw
64	Wang et al. [38]	C.&B. T.	Con.	Robotics	LiDAR	Rule	Shape	Raw
65	Wysocki et al. [152]	R.S.	O&M	Semantic Enr.	LiDAR	Semantic	Shape	3D-Model
66	Xiao et al. [37]	C.&B. T.	O&M	Building Info.	BIM	Classification	Type	3D-Model
67	Zhang and Aliaga [45]	C.S.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
68	Zhao et al. [100]	C.&B. T.	O&M	Building Info.	Drawings	Semantic	Shape	Raw
69	Adán et al. [64]	C.&B. T.	O&M	Energy Sim.	LiDAR-IRT	Rule	Dimension	3D-Model

Table A.1
Continue.

70	Biswanath et al. [44]	R.S.	O&M	Energy Sim.	IRT	Rule	Shape	3D-Model
71	Johns et al. [29]	C.&B. T.	Con.	Robotics	IRT	Rule	Shape	Raw
72	Liu et al. [72]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Status	2D-Polygon
73	Lu et al. [60]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Instance	Position	Raw
74	Walter et al. [71]	C.&B. T.	O&M	Energy Sim.	IRT	Semantic	Position	Raw
75	Wei et al. [78]	R.S.	O&M	Semantic Enr.	LiDAR	Semantic	Shape	3D-Model
76	Zeng et al. [69]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
77	Drobnyi et al. [47]	C.&B. T.	O&M	Building Info.	RGB-D	Rule	Shape	3D-Model
78	Forth et al. [30]	C.&B. T.	O&M	Energy Sim.	LiDAR-RGB	Semantic	Shape	3D-Model
79	Gu et al. [98]	C.&B. T.	O&M	Damage Sim.	RGB Camera	Semantic	Shape	3D-Model
80	Knechtel et al. [43]	C.&B. T.	O&M	Semantic Enr.	Drawings	Semantic	Position	Raw
81	Ma et al. [27]	C.&B. T.	Con.	Robotics	LiDAR-RGB	Semantic	Shape	Raw
82	Ogawa et al. [56]	R.S.	O&M	Semantic Enr.	RGB Camera	Instance	Shape	3D-Model
83	Pérez and Sánchez [41]	C.S.	O&M	Heritage Cons.	RGB Camera	Instance	Shape	3D-Model
84	Surojaya et al. [96]	C.S.	O&M	Damage Sim.	RGB Camera	Instance	Position	2D-Polygon
85	Tan et al. [67]	R.S.	O&M	Building Info.	LiDAR	Rule	Dimension	3D-Model
86	Wang et al. [76]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Status	Raw
87	Xia and Gong [32]	C.&B. T.	O&M	Building Info.	RGB Camera	Localization	Position	2D-Polygon
88	Yoon et al. [153]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Shape	3D-Model
89	Hu et al. [154]	C.&B. T.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
90	Wang et al. [155]	R.S.	O&M	Semantic Enr.	RGB Camera	Semantic	Shape	3D-Model
91	Chen et al. [156]	C.S.	O&M	Building Info.	RGB Camera	Localization	Position	Raw
92	Chang et al. [157]	C.S.	O&M	Semantic Enr.	Drawings	Semantic	Dimension	3D-Model
93	Su et al. [158]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Semantic	Shape	3D-Model
94	Yoon et al. [159]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Shape	3D-Model
95	Suppa et al. [40]	R.S.	O&M	Energy Sim.	RGB Camera	Localization	Dimension	2D-Polygon
96	Xie et al. [160]	C.&B. T.	O&M	Building Info.	RGB Camera	Semantic	Position	Raw
97	Mirzabeigi et al. [161]	C.&B. T.	O&M	Energy Sim.	IRT	Semantic	Dimension	Raw
98	Liang et al. [46]	C.S.	O&M	Semantic Enr.	LiDAR	Semantic	Shape	Raw
99	Ariss et al. [39]	C.&B. T.	O&M	Damage Sim.	RGB Camera	Localization	Shape	3D-Model
100	Zhang et al. [48]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Dimension	3D-Model
101	Yavariabdi et al. [162]	C.&B. T.	O&M	Damage Sim.	RGB Camera	Semantic	Position	Raw
102	Hany et al. [163]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Semantic	Shape	3D-Model
103	Fan et al. [164]	C.&B. T.	O&M	Damage Sim.	IRT	Semantic	Position	Raw
104	Zhao et al. [42]	R.S.	O&M	Semantic Enr.	LiDAR	Localization	Shape	3D-Model
105	Mehraban et al. [165]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Semantic	Shape	3D-Model

∗ RS = Remote Sensing, CS = Computer Science, C&BT = Construction & Building Technology
†O&M = Operations and Management, Con. = Construction
‡Enr. = Enrichment, Info. = Information extraction, Sim. = Simulation, Cons. = Conservation
§IRT = Infrared Temperature Sensor

Table A.2
List of 21 cases for validation collected from Scopus.

No.	Reference	Scenarios∗	Phase†	Application‡	Data Source§	Method	Attribute	Storage
1	Zhu et al. [166]	C.S.	O&M	Semantic Enr.	RGB Camera	Semantic	Shape	Raw
2	Lu et al. [167]	C.S.	Design	Semantic Enr.	Drawings	Semantic	Position	3D-Model
3	Pan et al. [168]	C.&B. T.	O&M	Semantic Enr.	LiDAR-RGB	Semantic	Shape	3D-Model
4	Xing et al. [106]	C.&B. T.	Design	Building Info.	Drawings	Localization	Position	Raw
5	Yu et al. [169]	R.S.	O&M	Building Info.	LiDAR	Localization	Position	Raw
6	Haznedar et al. [170]	C.&B. T.	O&M	Heritage Cons.	LiDAR-RGB	Semantic	Shape	Raw
7	Pantoja-Rosero et al. [171]	C.&B. T.	O&M	Damage Sim.	RGB Camera	Semantic	Shape	3D-Model
8	Chua and Cheah [102]	C.&B. T.	Con.	Building Info.	RGB Camera	Localization	Shape	2D-Polygon
9	Dai et al. [172]	C.&B. T.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
10	Escalada [173]	C.&B. T.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
11	Fotsing et al. [174]	C.&B. T.	O&M	Semantic Enr.	LiDAR-RGB	Rule	Shape	3D-Model
12	Li et al. [175]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Rule	Shape	3D-Model
13	Mehranfar et al. [176]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Localization	Shape	3D-Model
14	Ro and Gong [107]	C.&B. T.	O&M	Damage Sim.	RGB Camera	Semantic	Shape	Raw
15	Chowdhury et al. [105]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Position	3D-Model
16	Duan et al. [177]	C.&B. T.	O&M	Semantic Enr.	RGB Camera	Localization	Position	3D-Model
17	Hou et al. [104]	C.&B. T.	O&M	Heritage Cons.	RGB Camera	Semantic	Shape	Raw
18	Lai et al. [103]	C.&B. T.	Con.	Building Info.	RGB Camera	Localization	Position	Raw
19	Zhong et al. [178]	C.&B. T.	O&M	Building Info.	RGB Camera	Semantic	Shape	Raw
20	Forouzandeh et al. [179]	C.&B. T.	O&M	Semantic Enr.	LiDAR	Semantic	Shape	3D-Model
21	Kim et al. [101]	C.&B. T.	O&M	Energy Sim.	RGB Camera	Localization	Position	3D-Model

∗ RS = Remote Sensing, CS = Computer Science, C&BT = Construction & Building Technology
†O&M = Operations and Management, Con. = Construction
‡Enr. = Enrichment, Info. = Information extraction, Sim. = Simulation, Cons. = Conservation
§IRT = Infrared Temperature Sensor

References

[1] M. Lotteau, P. Loubet, M. Pousse, E. Dufrasnes, G. Sonnemann, Critical review of life cycle assessment (LCA) for the built environment at the neighborhood scale, Build. Environ. 93 (2015) 165–178. https://doi.org/10.1016/j.buildenv.2015.06.029
[2] A. Baranzini, C. Schaerer, A sight for sore eyes: Assessing the value of view and land use in the housing market, J. Hous. Econ. 20 (3) (2011) 191–199. https://doi.org/10.1016/j.jhe.2011.06.001
[3] R.K. Raanaas, G.G. Patil, T. Hartig, Health benefits of a view of nature through the window: a quasi-experimental study of patients in a residential rehabilitation center, Clin. Rehabil. 26 (1) (2012) 21–32. https://doi.org/10.1177/0269215511412800
[4] L. Wang, S. Greenberg, Window operation and impacts on building energy consumption, Energy Build. 92 (2015) 313–321. https://doi.org/10.1016/j.enbuild.2015.01.060
[5] C. Rubino, S. Liuzzi, G. Fusaro, F. Martellotta, C. Scrosati, M. Garai, Balancing ventilation and sound insulation in windows by means of metamaterials: a review of the state of the art, Build. Environ. (2025) 112780. https://doi.org/10.1016/j.buildenv.2025.112780
[6] F. Abd-Alhamid, M. Kent, Y. Wu, Quantifying window view quality: a review on view perception assessment and representation methods, Build. Environ. 227 (2023) 109742. https://doi.org/10.1016/j.buildenv.2022.109742
[7] Y. Chen, N. Li, X. Tian, X. Hu, Improving thermal comfort and working performance in windowless environment: relieving thermal response by introducing artificial view windows, Energy Build. 348 (2025) 116444. https://doi.org/10.1016/j.enbuild.2025.116444
[8] J. Souviron, G. Van Moeseke, A.Z. Khan, Analysing the environmental impact of windows: a review, Build. Environ. 161 (2019) 106268. https://doi.org/10.1016/j.buildenv.2019.106268
[9] Y. Feng, Q. Xiao, C. Brenner, A. Peche, J. Yang, U. Feuerhake, M. Sester, Determination of building flood risk maps from LiDAR mobile mapping data, Comput. Environ. Urban Syst. 93 (2022) 101759. https://doi.org/10.1016/j.compenvurbsys.2022.101759
[10] UNDESA, Synergy Solutions for Climate and SDG Action: Bridging the Ambition Gap for the Future we Want, Technical Report, United Nations, Department of Economic and Social Affairs, 2024. https://sdgs.un.org/sites/default/files/2024-08/UN%20Synergy%20Solutions%20for%20Climate%20and%20SDG%20Action-4_0.pdf#page=58.08.
[11] S. Kosse, P. Hagedorn, M. König, Semantic digital twins in construction: developing a modular system reference architecture based on information containers, Adv. Eng. Inf. 67 (2025) 103483. https://doi.org/10.1016/j.aei.2025.103483
[12] Q. Lu, A.K. Parlikad, P. Woodall, G. Don Ranasinghe, X. Xie, Z. Liang, E. Konstantinou, J. Heaton, J. Schooling, Developing a digital twin at building and city levels: case study of West Cambridge campus, J. Manage. Eng. 36 (3) (2020) 05020004. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000763
[13] L. Su, H. Yan, Y. Sun, D. Cao, X. Luo, Deciphering the trajectory and challenges of BIM research in China: a data-driven review of topics and innovation trends, Adv. Eng. Inf. 68 (2025) 103567. https://doi.org/10.1016/j.aei.2025.103567
[14] K. You, K. Chen, F. Xue, Automated scan-to-BIM for construction digital transformation: Conceptual framework, processing methods and best-practice guidelines, Autom. Constr. 182 (2026) 106755. https://doi.org/10.1016/j.autcon.2025.106755
[15] P. Pauwels, R. de Koning, B. Hendrikx, E. Torta, Live semantic data from building digital twins for robot navigation: Overview of data transfer methods, Adv. Eng. Inf. 56 (2023) 101959. https://doi.org/10.1016/j.aei.2023.101959
[16] I.C. Akyol, Y.M. Karadag, S. Ucar, I. Talaz, F.E. Gursoy, I.G. Dino, S. Kalkan, Transfer learning and parameter-efficient fine-tuning for heating energy consumption prediction using urban building energy models (UBEM), Adv. Eng. Inf. 68 (2025) 103576. https://doi.org/10.1016/j.aei.2025.103576
[17] W. Xie, Y. Zeng, X. Zhang, H.Y. Wong, T. Zhang, Z. Wang, X. Wu, J. Shi, X. Huang, F. Xiao, et al., AIoT-powered building digital twin for smart firefighting and super real-time fire forecast, Adv. Eng. Inf. 65 (2025) 103117. https://doi.org/10.1016/j.aei.2025.103117
[18] I. Acosta, M.a. Campano, J.F. Molina, Window design in architecture: analysis of energy savings for lighting and visual comfort in residential spaces, Appl. Energy 168 (2016) 493–506. https://doi.org/10.1016/j.apenergy.2016.02.005
[19] A. Arsiwala, F. Elghaish, M. Zoher, Digital twin with machine learning for predictive monitoring of CO2 equivalent from existing buildings, Energy Build. 284 (2023) 112851. https://doi.org/10.1016/j.enbuild.2023.112851
[20] K. Guerrero Ramírez, C. Nuevo-Gallardo, J.M. Santamaría Ulecia, B. Montalbán Pozas, C. Fernández Bandera, Digital twin implementation based on a white-box building energy model: a case study on blind control for passive heating, Energy Build. 349 (2025) 116454. https://doi.org/10.1016/j.enbuild.2025.116454
[21] S. Tuttas, U. Stilla, Reconstruction of façades in point clouds from multi aspect oblique ALS, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. II-3-W3 (2013) 91–96. https://doi.org/10.5194/isprsannals-II-3-W3-91-2013
[22] P. Oskouie, B. Becerik-Gerber, L. Soibelman, Automated recognition of building façades for creation of as-is mock-up 3D models, J. Comput. Civ. Eng. 31 (6) (2017) 04017059. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000711
[23] C. Sun, X. Guo, T. Zhao, Y. Han, Real-time detection method of window opening behavior using deep learning-based image recognition in severe cold regions, Energy Build. 268 (2022) 112196. https://doi.org/10.1016/j.enbuild.2022.112196
[24] A. Klimkowska, S. Cavazzi, R. Leach, S. Grebby, Detailed three-dimensional building façade reconstruction: a review on applications, data and technologies, Remote Sens. 14 (11) (2022) 2579. https://doi.org/10.3390/rs14112579
[25] F. Xue, L. Wu, W. Lu, Semantic enrichment of building and city information models: a ten-year review, Adv. Eng. Inf. 47 (2021) 101245. https://doi.org/10.1016/j.aei.2020.101245
[26] J.T. Szcześniak, Y.Q. Ang, S. Letellier-Duchesne, C.F. Reinhart, A method for using street view imagery to auto-extract window-to-wall ratios and its relevance for urban-level daylighting and energy simulations, Build. Environ. 207 (2022) 108108. https://doi.org/10.1016/j.buildenv.2021.108108
[27] J. Ma, X. Jiang, D. Zheng, X. Liao, J. Lu, Y. Zhao, TransWallNet: high-performance semantic segmentation of large-scale and multifeatured point clouds of building gables, J. Constr. Eng. Manage. 150 (8) (2024) 04024092. https://doi.org/10.1061/JCEMD4.COENG-14827
[28] J. Kim, J. Song, J.-K. Lee, Recognizing and classifying unknown object in BIM using 2D CNN, in: J.-H. Lee (Ed.), Computer-Aided Architectural Design. "Hello, Culture", Springer, Singapore, 2019, pp. 47–57. https://doi.org/10.1007/978-981-13-8410-3_4
[29] B. Johns, E. Abdi, M. Arashpour, Crane payload localisation for curtain wall installation: a markerless computer vision approach, Measurement 221 (2023) 113459. https://doi.org/10.1016/j.measurement.2023.113459
[30] K. Forth, F. Noichl, A. Borrmann, LCA calculation of retrofitting scenarios using geometric model reconstruction and semantic enrichment of point clouds and images, in: Computing in Civil Engineering 2023, 2023, pp. 390–397. https://doi.org/10.1061/9780784485231.047
[31] D. Luong, R. Richman, M. Touchie, Towards window state detection using image processing in residential and office building facades, Build. Environ. 207 (2022) 108486. https://doi.org/10.1016/j.buildenv.2021.108486
[32] J. Xia, J. Gong, Computer vision based first floor elevation estimation from mobile LiDAR data, Autom. Constr. 159 (2024) 105258. https://doi.org/10.1016/j.autcon.2023.105258
[33] Z. Zou, K. Chen, Z. Shi, Y. Guo, J. Ye, Object detection in 20 years: a survey, Proc. IEEE 111 (3) (2023) 257–276. https://doi.org/10.1109/JPROC.2023.3238524
[34] S. Yoon, Building digital twinning: data, information, and models, J. Build. Eng. 76 (2023) 107021. https://doi.org/10.1016/j.jobe.2023.107021
[35] R.G.N. Ruwanthika, P.A.D.B.M. Amarasekera, R.U.I.B. Chandrasiri, D.M.A.I. Rangana, A. Nugaliyadde, Y. Mallawarachchi, Dynamic 3D model construction using architectural house plans, in: 2017 6th National Conference on Technology and Management (NCTM), 2017, pp. 181–184. https://doi.org/10.1109/NCTM.2017.7872850
[36] I. Dochev, P. Gorzalka, V. Weiler, J. Estevam Schmiedt, M. Linkiewicz, U. Eicker, B. Hoffschmidt, I. Peters, B. Schröter, Calculating urban heat demands: an analysis of two modelling approaches and remote sensing for input data and validation, Energy Build. 226 (2020) 110378. https://doi.org/10.1016/j.enbuild.2020.110378
[37] M. Xiao, Z. Chao, R.F. Coelho, S. Tian, Investigation of classification and anomalies based on machine learning methods applied to large scale building information modeling, Appl. Sci. 12 (13) (2022) 6382. https://doi.org/10.3390/app12136382
[38] Y. Wang, L. Xie, H. Wang, W. Zeng, Y. Ding, T. Hu, T. Zheng, H. Liao, J. Hu, Intelligent spraying robot for building walls with mobility and perception, Autom. Constr. 139 (2022) 104270. https://doi.org/10.1016/j.autcon.2022.104270
[39] M. Ariss, B.G. Pantoja-Rosero, F. Duarte, M. Klimenka, C. Ratti, Seismic assessment of unreinforced masonry façades from images using macroelement-based modeling, Commun. Eng. 4 (1) (2025) 155. https://doi.org/10.1038/s44172-025-00487-2
[40] A.R. Suppa, A. Aliberti, M.C. Bottero, V. Corrado, Detecting window-to-wall ratio for urban-scale building simulations using deep learning with street view imagery and an automatic classification algorithm, Build. Simul. 18 (8) (2025) 2175–2199. https://doi.org/10.1007/s12273-025-1301-3
[41] F.N. Pérez, I.B. Sánchez, Automatic detection and 3D reconstruction of buildings from historical maps, Appl. Artif. Intell. 38 (1) (2024) 2378274. https://doi.org/10.1080/08839514.2024.2378274
[42] T. Zhao, T. Xiong, M. Li, Z. Li, Automatic reconstruction of 3D building models from ALS point clouds based on façade geometry, ISPRS Int. J. Geo-Inf. 14 (12) (2025). https://doi.org/10.3390/ijgi14120462
[43] J. Knechtel, P. Rottmann, J.-H. Haunert, Y. Dehbi, Semantic floorplan segmentation using self-constructing graph networks, Autom. Constr. 166 (2024) 105649. https://doi.org/10.1016/j.autcon.2024.105649
[44] M.K. Biswanath, L. Hoegner, U. Stilla, Thermal mapping from point clouds to 3D building model facades, Remote Sens. 15 (19) (2023) 4830. https://doi.org/10.3390/rs15194830
[45] X. Zhang, D. Aliaga, RFCNet: enhancing urban segmentation using regularization, fusion, and completion, Comput. Vision Image Understanding 220 (2022) 103435. https://doi.org/10.1016/j.cviu.2022.103435
[46] X. Liang, Y. Liu, J. Wang, Z. Qin, L. Chen, H. Jia, C. Dai, Automatic 3D reconstruction of indoor building point clouds under occlusion based on gestalt rules, Int. J. Digital Earth 18 (1) (2025) 2545582. https://doi.org/10.1080/17538947.2025.2545582
[47] V. Drobnyi, S. Li, I. Brilakis, Digitization of existing buildings with arbitrary shaped spaces from point clouds, J. Comput. Civ. Eng. 38 (5) (2024) 04024027. https://doi.org/10.1061/JCCEE5.CPENG-5853
[48] C. Zhang, P.-J. Hoes, B. Tian, R. Zhang, R. Loonen, Enhancing neighborhood-scale building performance simulation through building classification and automated data acquisition: Supporting the Dutch heating transition, Sustainable Cities Soc. 132 (2025) 106811. https://doi.org/10.1016/j.scs.2025.106811
[49] W. Hao, Y. Wang, W. Liang, Slice-based building facade reconstruction from 3D point clouds, Int. J. Remote Sens. 39 (20) (2018) 6587–6606. https://doi.org/10.1080/01431161.2018.1463113
[50] A. Adán, B. Quintana, S.A. Prieto, F. Bosché, An autonomous robotic platform for automatic extraction of detailed semantic models of buildings, Autom. Constr. 109

(2020) 102963. https://doi.org/10.1016/j.autcon.2019.102963
[51] S. Van Ackere, J. Verbeurgt, L. De Sloover, A. De Wulf, N. Van de Weghe, P. De Maeyer, Extracting dimensions and locations of doors, windows, and door thresholds out of mobile lidar data using object detection to estimate the impact of floods, Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-3-W8 (2019) 429–436. https://doi.org/10.5194/isprs-archives-XLII-3-W8-429-2019
[52] E. Valero, A. Adán, F. Bosché, Semantic 3D reconstruction of furnished interiors using laser scanning and RFID technology, J. Comput. Civ. Eng. 30 (4) (2016) 04015053. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000525
[53] P. Meixner, F. Leberl, Interpreting Building Façades from Vertical Aerial Images Using the Third Dimension, Int. Ann. Photogramm. Remote Sens. Spatial Inf. Sci. 38 (4) (2010).
[54] Y. Golombek, W. Marshall, High-density mobile LiDAR for measuring urban streetscape features, URBAN DESIGN International 27 (1) (2022) 3–17. https://doi.org/10.1057/s41289-021-00176-4
[55] D. Gu, W. Chen, X. Lu, Automated assessment of wind damage to windows of buildings at a city scale based on oblique photography, deep learning and CFD, J. Build. Eng. 52 (2022) 104355. https://doi.org/10.1016/j.jobe.2022.104355
[56] Y. Ogawa, R. Nakamura, G. Sato, H. Maeda, Y. Sekimoto, End-to-end framework for the automatic matching of omnidirectional street images and building data and the creation of 3D building models, Remote Sens. 16 (11) (2024) 1858. https://doi.org/10.3390/rs16111858
[57] P.W. Tien, S. Wei, J.K. Calautit, J. Darkwa, C. Wood, Real-time monitoring of occupancy activities and window opening within buildings using an integrated deep learning-based approach for reducing energy demand, Appl. Energy 308 (2022) 118336. https://doi.org/10.1016/j.apenergy.2021.118336
[58] Y.Y. Song, Y. Lu, Decision tree methods: applications for classification and prediction, Shanghai Arch. Psychiatry 27 (2) (2015) 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044
[59] A. Murtiyoso, E. Pellis, P. Grussenmeyer, T. Landes, A. Masiero, Towards semantic photogrammetry: generating semantically rich point clouds from architectural close-range photogrammetry, Sensors 22 (3) (2022) 966. https://doi.org/10.3390/s22030966
[60] Y. Lu, W. Wei, P. Li, T. Zhong, Y. Nong, X. Shi, A deep learning method for building façade parsing utilizing improved SOLOv2 instance segmentation, Energy Build. 295 (2023) 113275. https://doi.org/10.1016/j.enbuild.2023.113275
[61] K. Pexman, D.D. Lichti, P. Dawson, Automated storey separation and door and window extraction for building models from complete laser scans, Remote Sens. 13 (17) (2021) 3384. https://doi.org/10.3390/rs13173384
[62] W. Shi, W. Ahmed, N. Li, W. Fan, H. Xiang, M. Wang, Semantic geometric modelling of unstructured indoor point cloud, ISPRS Int. J. Geo-Inf. 8 (1) (2018) 9. https://doi.org/10.3390/ijgi8010009
[63] M. Al-Kasasbeh, O. Abudayyeh, H. Liu, An integrated decision support system for building asset management based on BIM and work breakdown structure, J. Build. Eng. 34 (2021) 101959. https://doi.org/10.1016/j.jobe.2020.101959
[64] A. Adán, A. Ramón, J.L. Vivancos, A. Vilar, C. Aparicio-Fernández, Automatic generation of as-is BEM models of buildings, J. Build. Eng. 73 (2023) 106865. https://doi.org/10.1016/j.jobe.2023.106865
[65] Y.-C. Chung, T.X. Han, Z. He, Building recognition using sketch-based representations and spectral graph matching, in: 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 2014–2020. https://doi.org/10.1109/ICCV.2009.5459444
[66] S. Becker, Generation and application of rules for quality dependent façade reconstruction, ISPRS J. Photogramm. Remote Sens. 64 (6) (2009) 640–653. https://doi.org/10.1016/j.isprsjprs.2009.06.002
[67] Y. Tan, X. Liu, S. Jin, Q. Wang, D. Wang, X. Xie, A terrestrial laser scanning-based method for indoor geometric quality measurement, Remote Sens. 16 (1) (2023) 59. https://doi.org/10.3390/rs16010059
[68] A. Masiero, F. Chiabrando, A.M. Lingua, B.G. Marino, F. Fissore, A. Guarnieri, A. Vettore, 3D modeling of Girifalco fortress, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2-W9 (2019) 473–478. https://doi.org/10.5194/isprs-archives-XLII-2-W9-473-2019
[69] R. Zeng, J.J.S. Shi, C. Wang, T. Lu, Integrating as-built BIM model from point cloud data in construction projects, Eng. Constr. Archit. Manage. 31 (9) (2024) 3557–3574. https://doi.org/10.1108/ECAM-12-2022-1196
[70] M. Yin, L. Tang, T. Zhou, Y. Wen, R. Xu, W. Deng, Automatic layer classification method-based elevation recognition in architectural drawings for reconstruction of 3D BIM models, Autom. Constr. 113 (2020) 103082. https://doi.org/10.1016/j.autcon.2020.103082
[71] I. Walter, M. Tanasković, M. Stanković, IR building analysis with extraction of elements using image segmentation and RetinaNet, Buildings 13 (1) (2022) 109. https://doi.org/10.3390/buildings13010109
[72] Y. Liu, W.T. Chong, Y.H. Yau, D. Han, M. Qin, F. Deng, X. Wang, S. Pan, A new algorithm for real-time detection of window opening area in residential buildings, Build. Environ. 244 (2023) 110817. https://doi.org/10.1016/j.buildenv.2023.110817
[73] Y. Arjoune, S. Peri, N. Sugunaraj, A. Biswas, D. Sadhukhan, P. Ranganathan, An instance segmentation and clustering model for energy audit assessments in built environments: a multi-stage approach, Sensors 21 (13) (2021) 4375. https://doi.org/10.3390/s21134375
[74] K. Bacharidis, F. Sarri, V. Paravolidakis, L. Ragia, M. Zervakis, Fusing georeferenced and stereoscopic image data for 3D building façade reconstruction, ISPRS Int. J. Geo-Inf. 7 (4) (2018) 151. https://doi.org/10.3390/ijgi7040151
[75] Q. Lu, S. Lee, L. Chen, Image-driven fuzzy-based system to construct as-is IFC BIM objects, Autom. Constr. 92 (2018) 68–87. https://doi.org/10.1016/j.autcon.2018.03.034
[76] Z. Wang, H. Tang, H. Zhang, D. Jiang, Y. Leng, Y. Wu, J. Calautit, L. Xia, H.N. Chaudhry, Multi-objectives occupant-centric control of thermostats and natural ventilation systems in cold climate conditions using real-time occupant-related information, Build. Environ. 263 (2024) 111882. https://doi.org/10.1016/j.buildenv.2024.111882
[77] N. ÓMahony, S. Campbell, A. Carvalho, S. Harapanahalli, G.V. Hernandez, L. Krpalkova, D. Riordan, J. Walsh, Deep learning vs. traditional computer vision, in: Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1, Springer, 2020, pp. 128–144. https://doi.org/10.1007/978-3-030-17795-9_10
[78] J. Wei, H. Wu, H. Yue, S. Jia, J. Li, C. Liu, Automatic extraction and reconstruction of a 3D wireframe of an indoor scene from semantic point clouds, Int. J. Digital Earth 16 (1) (2023) 3239–3267. https://doi.org/10.1080/17538947.2023.2246937
[79] H. Zheng, F. Li, H. Cai, K. Zhang, Non-intrusive measurement method for the window opening behavior, Energy Build. 197 (2019) 171–176. https://doi.org/10.1016/j.enbuild.2019.05.052
[80] T.S. Fong, W.Y. Yan, Signs on glasses: LiDAR data voids, hotspot effect, and reflection artifacts, Autom. Constr. 169 (2025) 105877. https://doi.org/10.1016/j.autcon.2024.105877
[81] S.M.I. Zolanvari, D.F. Laefer, A.S. Natanzi, Three-dimensional building façade segmentation and opening area detection from point clouds, ISPRS J. Photogramm. Remote Sens. 143 (2018) 134–149. https://doi.org/10.1016/j.isprsjprs.2018.04.004
[82] S.D. Roth, Ray casting for modeling solids, Comput. Graphics Image Process. 18 (2) (1982) 109–144. https://doi.org/10.1016/0146-664X(82)90169-1
[83] 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
[84] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. (6) (1986) 679–698. https://doi.org/10.1109/TPAMI.1986.4767851
[85] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (6) (2016) 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031
[86] J. Redmon, You only look once: unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91
[87] C. Rother, A new approach to vanishing point detection in architectural environments, Image Vis. Comput. 20 (9-10) (2002) 647–655. https://doi.org/10.1016/S0262-8856(02)00054-9
[88] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-assisted Intervention, Springer, 2015, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
[89] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2017) 834–848. https://doi.org/10.1109/TPAMI.2017.2699184
[90] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2961–2969. https://doi.org/10.1109/ICCV.2017.322
[91] X. Wang, R. Zhang, T. Kong, L. Li, C. Shen, SOLOv2: dynamic and fast instance segmentation, Adv. Neural Inf. Process. Syst. 33 (2020) 17721–17732. https://doi.org/10.5555/3495724.3497211
[92] A.D. Styliadis, Historical photography-based computer-aided architectural design: demolished buildings information modeling with reverse engineering functionality, Autom. Constr. 18 (1) (2008) 51–69. https://doi.org/10.1016/j.autcon.2008.04.010
[93] Y. Sun, S. Malihi, H. Li, M. Maboudi, Deepwindows: windows instance segmentation through an improved mask R-CNN using spatial attention and relation modules, ISPRS Int. J. Geo-Inf. 11 (3) (2022) 162. https://doi.org/10.3390/ijgi11030162
[94] S. Dou, X. Jiang, L. Liu, L. Ying, C. Shan, Y. Shen, X. Dong, Y. Wang, D. Li, C. Zhao, Hierarchical recognizing vector graphics and a new chart-based vector graphics dataset, IEEE Trans. Pattern Anal. Mach. Intell. (2024). https://doi.org/10.1109/TPAMI.2024.3394298
[95] Y. Zhao, X. Deng, H. Lai, Reconstructing BIM from 2D structural drawings for existing buildings, Autom. Constr. 128 (2021) 103750. https://doi.org/10.1016/j.autcon.2021.103750
[96] A. Surojaya, N. Zhang, J.R. Bergado, F. Nex, Towards fully autonomous UAV: damaged building-opening detection for outdoor-indoor transition in urban search and rescue, Electronics 13 (3) (2024) 558. https://doi.org/10.3390/electronics13030558
[97] S. Touzani, M. Wudunn, S. Fernandes, A. Zakhor, R. Najibi, J. Granderson, A machine learning approach to estimate windows-to-wall ratio using drone imagery, in: N. Chrysoulakis, T. Erbertseder, Y. Zhang (Eds.), Remote Sensing Technologies and Applications in Urban Environments VI, SPIE, Online Only, Spain, 2021, p. 20. https://doi.org/10.1117/12.2602157
[98] D.L. Gu, Q.W. Shuai, N. Zhang, N. Jin, Z.X. Zheng, Z. Xu, Y.J. Xu, Multi-view street view image fusion for city-scale assessment of wind damage to building clusters, Comput.-Aided Civ. Infrastruct. Eng. 40 (2) (2025) 198–214. https://doi.org/10.1111/mice.13324
[99] K.A. Al-Sallal, A.R. AbouElhamd, M.B. Dalmouk, UAE heritage buildings converted into museums: evaluation of daylighting effectiveness and potential risks on artifacts and visual comfort, Energy Build. 176 (2018) 333–359. https://doi.org/10.1016/j.enbuild.2018.06.067

[100] X. Zhao, T.-H. Wang, C. Peng, Automatic room type classification using machine learning for two-dimensional residential building plans, in: Proceedings of the 40th Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2022), 2, Ghent, Belgium, 2022, pp. 593–600. https://doi.org/10.52842/conf.ecaade.2022.2.593

[101] Y. Kim, J. Yoon, S. Lee, M. Shin, Automated building geometry modeling using UAV imagery and deep learning for energy simulation, J. Build. Eng. 117 (2026) 114949. https://doi.org/10.1016/j.jobe.2025.114949

[102] W.P. Chua, C.C. Cheah, Deep-learning-based automated building construction progress monitoring for prefabricated prefinished volumetric construction, Sensors 24 (21) (2024). https://doi.org/10.3390/s24217074

[103] Y. Lai, Z. Lin, Z. Ye, J. Zhang, Development of advanced progress recognition algorithms for construction monitoring, PLOS ONE 20 (10) (2025) e0333262. https://doi.org/10.1371/journal.pone.0333262

[104] J. Hou, J. Zhou, Y. He, B. Hou, J. Li, Automatic reconstruction of semantic façade model of architectural heritage, Heritage Sci. 12 (1) (2024) 400. https://doi.org/10.1186/s40494-024-01506-9

[105] S. Chowdhury, S.N. Kandelan, K.S. Grewal, Harnessing machine learning for rapid and cost-efficient 3D geometry generation in neighborhood energy modeling, Results Eng. 28 (2025) 107599. https://doi.org/10.1016/j.rineng.2025.107599

[106] J. Xing, Q. Luo, Y. Wu, J. Shang, Self-optimization for parsing floor plans, J. Comput. Civ. Eng. 36 (6) (2022) 04022037. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001048

[107] S.H. Ro, J. Gong, Scalable approach to create annotated disaster image database supporting AI-driven damage assessment, Nat. Hazards 120 (13) (2024) 11693–11712. https://doi.org/10.1007/s11069-024-06641-x

[108] S. Xiang, A. Xie, M. Ye, X. Yan, X. Han, H. Niu, Q. Li, H. Huang, Autonomous eVTOL: a summary of researches and challenges, Green Energy Intell. Transp. 3 (1) (2024) 100140. https://doi.org/10.1016/j.geits.2023.100140

[109] T. Van Renterghem, D. Botteldooren, View on outdoor vegetation reduces noise annoyance for dwellers near busy roads, Landscape Urban Plann. 148 (2016) 203–215. https://doi.org/10.1016/j.landurbplan.2015.12.018

[110] M. Li, F. Xue, Y. Wu, A.G.O. Yeh, A room with a view: automatic assessment of window views for high-rise high-density areas using city information models and deep transfer learning, Landscape Urban Plann. 226 (2022) 104505. https://doi.org/10.1016/j.landurbplan.2022.104505

[111] M. Li, S. Guo, F. Duarte, A. Kumar, N. Kobori, F. Xue, W. Zhuang, A.G.O. Yeh, C. Ratti, Influence of objective and perceived exposures to urban nature on people's happiness, npj Urban Sustainability 6 (1) (2026) 6. https://doi.org/10.1038/s42949-025-00306-9

[112] E. Cuce, S.B. Riffat, C.-H. Young, Thermal insulation, power generation, lighting and energy saving performance of heat insulation solar glass as a curtain wall application in Taiwan: a comparative experimental study, Energy Convers. Manage. 96 (2015) 31–38. https://doi.org/10.1016/j.enconman.2015.02.062

[113] J. Shi, Z. Pan, L. Jiang, X. Zhai, An ontology-based methodology to establish city information model of digital twin city by merging BIM, GIS and IoT, Adv. Eng. Inf. 57 (2023) 102114. https://doi.org/10.1016/j.aei.2023.102114

[114] B.G. Pantoja-Rosero, A. Rusnak, F. Kaplan, K. Beyer, Generation of LOD4 models for buildings towards the automated 3D modeling of BIMs and digital twins, Autom. Constr. 168 (2024) 105822. https://doi.org/10.1016/j.autcon.2024.105822

[115] S.-H. Chen, F. Xue, Automatic BIM detailing using deep features of 3D views, Autom. Constr. 148 (2023) 104780. https://doi.org/10.1016/j.autcon.2023.104780

[116] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, et al., Segment anything, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 4015–4026. https://doi.org/10.1109/ICCV51070.2023.00371

[117] C. Lee, S. Park, H. Song, J. Ryu, S. Kim, H. Kim, S. Pereira, D. Yoo, Interactive multi-class tiny-object detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 14136–14145. https://doi.org/10.1109/CVPR52688.2022.01374

[118] S. Meng, X. Su, G. Sun, M. Li, F. Xue, From 3D pedestrian networks to wheelable networks: an automatic wheelability assessment method for high-density urban areas using contrastive deep learning of smartphone point clouds, Comput. Environ. Urban Syst. 117 (2025) 102255. https://doi.org/10.1016/j.compenvurbsys.2025.102255

[119] P. Meixner, F. Leberl, 3-Dimensional building details from aerial photography for internet maps, Remote Sens. 3 (4) (2011) 721–751. https://doi.org/10.3390/rs3040721

[120] L. Truong-Hong, D.F. Laefer, T. Hinks, H. Carr, Flying voxel method with delaunay triangulation criterion for façade/feature detection for computation, J. Comput. Civ. Eng. 26 (6) (2012) 691–707. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000188

[121] G.G. Demisse, D. Borrmann, A. Nüchter, Interpreting thermal 3D models of indoor environments for energy efficiency, in: 2013 16th International Conference on Advanced Robotics (ICAR), 2013, pp. 1–8. https://doi.org/10.1109/ICAR.2013.6766550

[122] R.-C. Dumitru, D. Borrmann, A. Nüchter, Interior reconstruction using the 3D Hough transform, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-5-W1 (2013) 65–72. https://doi.org/10.5194/isprsarchives-XL-5-W1-65-2013

[123] L. Truong-Hong, D.F. Laefer, T. Hinks, H. Carr, Combining an angle criterion with voxelization and the flying voxel method in reconstructing building models from LiDAR data, Comput.-Aided Civ. Infrastruct. Eng. 28 (2) (2013) 112–129. https://doi.org/10.1111/j.1467-8667.2012.00761.x

[124] C. Wang, Y.K. Cho, M. Gai, As-is 3D thermal modeling for existing building envelopes using a hybrid LIDAR system, J. Comput. Civ. Eng. 27 (6) (2013) 645–656. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000273

[125] C. Riedinger, M. Jordan, H. Tabia, 3D models over the centuries: From old floor plans to 3D representation, in: 2014 International Conference on 3D Imaging (IC3D), 2014, pp. 1–8. https://doi.org/10.1109/IC3D.2014.7032583

[126] D. Frommholz, M. Linkiewicz, H. Meissner, D. Dahlke, A. Poznanska, Extracting semantically annotated 3D building models with textures from oblique aerial imagery, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-3-W2 (2015) 53–58. https://doi.org/10.5194/isprsarchives-XL-3-W2-53-2015

[127] S. Hong, J. Jung, S. Kim, H. Cho, J. Lee, J. Heo, Semi-automated approach to indoor mapping for 3D as-built building information modeling, Comput. Environ. Urban Syst. 51 (2015) 34–46. https://doi.org/10.1016/j.compenvurbsys.2015.01.005

[128] Y. Liu, S. Qin, Facade model refinement by fusing terrestrial laser data and image, in: MIPPR 2015: Pattern Recognition and Computer Vision, 9813, SPIE, 2015, pp. 291–296. https://doi.org/10.1117/12.2214179

[129] F. Sadeghi, H. Arefi, A. Fallah, M. Hahn, 3D building façade reconstruction using handheld laser scanning data, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-1-W5 (2015) 625–630. https://doi.org/10.5194/isprsarchives-XL-1-W5-625-2015

[130] A. Arnaud, J. Christophe, M. Gouiffes, M. Ammi, 3D reconstruction of indoor building environments with new generation of tablets, in: Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, VRST '16, Association for Computing Machinery, New York, NY, USA, 2016, pp. 187–190. https://doi.org/10.1145/2993369.2993403

[131] T. Colleu, S. Benitez, Ray-based detection of openings in urban areas using mobile LiDAR data, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B3 (2016) 591–598. https://doi.org/10.5194/isprs-archives-XLI-B3-591-2016

[132] L. Díaz-Vilariño, P. Boguslawski, K. Khoshelham, H. Lorenzo, L. Mahdjoubi, Indoor navigation from point clouds: 3D modelling and obstacle detection, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B4 (2016) 275–281. https://doi.org/10.5194/isprs-archives-XLI-B4-275-2016

[133] L. Gimenez, S. Robert, F. Suard, K. Zreik, Automatic reconstruction of 3D building models from scanned 2D floor plans, Autom. Constr. 63 (2016) 48–56. https://doi.org/10.1016/j.autcon.2015.12.008

[134] J. Jung, S. Hong, S. Yoon, J. Kim, J. Heo, Automated 3D wireframe modeling of indoor structures from point clouds using constrained least-squares adjustment for as-built BIM, J. Comput. Civ. Eng. 30 (4) (2016) 04015074. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000556

[135] E. Lachat, T. Landes, P. Grussenmeyer, Combination of TLS point clouds and 3D data from kinect V2 sensor to complete indoor models, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B5 (2016) 659–666. https://doi.org/10.5194/isprs-archives-XLI-B5-659-2016

[136] W. Nguatem, M. Drauschke, H. Mayer, Automatic generation of building models with levels of detail 1-3, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B3 (2016) 649–654. https://doi.org/10.5194/isprsarchives-XLI-B3-649-2016

[137] V. Paravolidakis, K. Bacharidis, F. Sarri, L. Ragia, M. Zervakis, Reduction of building façade model complexity using computer vision, in: 2016 IEEE International Conference on Imaging Systems and Techniques (IST), 2016, pp. 454–459. https://doi.org/10.1109/IST.2016.7738269

[138] J. Chen, Y.K. Cho, K. Kim, Region proposal mechanism for building element recognition for advanced scan-to-BIM process, in: Construction Research Congress 2018, 2018, pp. 221–231. https://doi.org/10.1061/9780784481264.022

[139] S. Franz, R. Irmler, U. Rüppel, Real-time collaborative reconstruction of digital building models with mobile devices, Adv. Eng. Inf. 38 (2018) 569–580. https://doi.org/10.1016/j.aei.2018.08.012

[140] J. Jung, C. Stachniss, S. Ju, J. Heo, Automated 3D volumetric reconstruction of multiple-room building interiors for as-built BIM, Adv. Eng. Inf. 38 (2018) 811–825. https://doi.org/10.1016/j.aei.2018.10.007

[141] M. Previtali, L. Díaz-Vilariño, M. Scaioni, Indoor building reconstruction from occluded point clouds using graph-cut and ray-tracing, Appl. Sci. 8 (9) (2018) 1529. https://doi.org/10.3390/app8091529

[142] Q. Wu, P. Meng, G. Liu, Reconstruction of 3D building model based on the information in floor plan, Traitem. Signal 35 (3-4) (2018) 303–316. https://doi.org/10.3166/TS.35.303-316

[143] Y. Cui, Q. Li, B. Yang, W. Xiao, C. Chen, Z. Dong, Automatic 3-D reconstruction of indoor environment with mobile laser scanning point clouds, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12 (8) (2019) 3117–3130. https://doi.org/10.1109/JSTARS.2019.2918937

[144] S. Tang, Y. Zhang, Y. Li, Z. Yuan, Y. Wang, X. Zhang, X. Li, Y. Zhang, R. Guo, W. Wang, Fast and automatic reconstruction of semantically rich 3D indoor maps from low-quality RGB-D sequences, Sensors 19 (3) (2019) 533. https://doi.org/10.3390/s19030533

[145] H. Fan, Y. Wang, J. Gong, Layout graph model for semantic façade reconstruction using laser point clouds, Geo-spatial Inf. Sci. 24 (3) (2021) 403–421. https://doi.org/10.1080/10095020.2021.1922316

[146] J. Voříšek, B. Patzák, E. Dvořáková, D. Rypl, Automated BIM entity reconstruction from unstructured 3d pointclouds, Acta Polytech. CTU Proc. 30 (2021) 126–130. https://doi.org/10.14311/APP.2021.30.0126

[147] C. Wang, S. Wei, S. Du, D. Zhuang, Y. Li, X. Shi, X. Jin, X. Zhou, A systematic method to develop three dimensional geometry models of buildings for urban building energy modeling, Sustainable Cities Soc. 71 (2021) 102998. https://doi.org/10.1016/j.scs.2021.102998

[148] M. Jarząbek-Rychard, H.G. Maas, Automatic enrichment of indoor 3D models using a deep learning approach based on single images with unknown camera poses, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. VIII-5-W1-2022 (2022) 1–7. https://doi.org/10.5194/isprs-annals-VIII-5-W1-2022-1-2022

[149] Y. Jiang, S. Han, Y. Bai, Scan4Façade: automated as-is façade modeling of historic high-rise buildings using drones and AI, J. Archit. Eng. 28 (4) (2022) 04022031.

https://doi.org/10.1061/(ASCE)AE.1943-5568.0000564
[150] D. Li, T. Yu, Y. Yang, Y. Zhao, Z. Chen, J. Liu, Intelligent heating monitoring method based on infrared image segmentation and target detection: poster abstract, in: Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, BuildSys '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 301–302. https://doi.org/10.1145/3563357.3567755
[151] B.G. Pantoja-Rosero, R. Achanta, M. Kozinski, P. Fua, F. Perez-Cruz, K. Beyer, Generating LOD3 building models from structure-from-motion and semantic segmentation, Autom. Constr. 141 (2022) 104430. https://doi.org/10.1016/j.autcon.2022.104430
[152] O. Wysocki, E. Grilli, L. Hoegner, U. Stilla, Combining visibility analysis and deep learning for refinement of semantic 3D building models by conflict classification, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. X-4-W2-2022 (2022) 289–296. https://doi.org/10.5194/isprs-annals-X-4-W2-2022-289-2022
[153] J. Yoon, Y. Kim, S. Lee, M. Shin, UAV-based automated 3D modeling framework using deep learning for building energy modeling, Sustainable Cities Soc. 101 (2024) 105169. https://doi.org/10.1016/j.scs.2023.105169
[154] Y. Hu, J. Wei, S. Zhang, S. Liu, FDIE: a graph-based framework for extracting design information from annotated building facade images, J. Asian Archit. Build. Eng. 24 (4) (2025) 2530–2553. https://doi.org/10.1080/13467581.2024.2373835
[155] D. Wang, J. Liu, H. Jiang, P. Liu, Q. Jiang, Existing buildings recognition and BIM generation based on multi-plane segmentation and deep learning, Buildings 15 (5) (2025). https://doi.org/10.3390/buildings15050691
[156] Y. Chen, T. Wang, G. Chen, K. Zhu, X. Tan, J. Wang, W. Guo, Q. Wang, X. Luo, X. Zhang, BFA-YOLO: A balanced multiscale object detection network for building façade elements detection, Adv. Eng. Inf. 65 (2025) 103289. https://doi.org/10.1016/j.aei.2025.103289
[157] J. Chang, Y. Lv, J. Wang, H. Pang, Y. Liu, Raster image-based house-type recognition and three-dimensional reconstruction technology, Buildings 15 (7) (2025). https://doi.org/10.3390/buildings15071178
[158] A.J. Su, A. Ren, K.C. Xu, T. Dogan, Scalable building reconstruction and window detection for urban building energy modelling applications, J. Build. Perform. Simul. 0 (0) (2025) 1–19. https://doi.org/10.1080/19401493.2025.2501151
[159] J. Yoon, S. Hwang, K. Kim, S. Lee, Construction and recording method of a three-dimensional model to automatically manage thermal abnormalities in building exteriors, Buildings 15 (9) (2025). https://doi.org/10.3390/buildings15091558
[160] J. Xie, M. Li, J. Wu, X. Zhang, J. Zhang, Semantic segmentation of building façade materials and colors for urban conservation, npj Heritage Sci. 13 (1) (2025) 378. https://doi.org/10.1038/s40494-025-01888-4
[161] S. Mirzabeigi, R. Razkenari, P. Crovella, Building envelope thermal anomaly detection using an integrated vision-based technique and semantic segmentation, Buildings 15 (15) (2025). https://doi.org/10.3390/buildings15152672
[162] A. Yavariabdi, M.F. Asik, G. Dogan, M.H. Arslan, Vision-based analysis of soft story, short columns, and vertical geometry in RC structures, Structures 80 (2025) 109951. https://doi.org/10.1016/j.istruc.2025.109951
[163] Y. Hany, W. Ahmed, A. Elshazly, A.M. Senousi, W. Darwish, Automated 3D reconstruction of interior structures from unstructured point clouds, ISPRS Int. J. Geo-Inf. 14 (11) (2025). https://doi.org/10.3390/ijgi14110428
[164] Y. Fan, J. Mai, F. Xue, S.S.Y. Lau, S. Jiang, Y. Tao, X. Zhang, W.C. Tsang, UAV and deep learning for automated detection and visualization of façade defects in existing residential buildings, Sensors 25 (23) (2025). https://doi.org/10.3390/s25237118
[165] M.H. Mehraban, S. Mirzabeigi, M. Wang, R. Liu, S.M.E. Sepasgozar, Automated image-to-BIM using neural radiance fields and vision-language semantic modeling, Buildings 15 (24) (2025). https://doi.org/10.3390/buildings15244549
[166] P. Zhu, W.R. Para, A. Frühstück, J. Femiani, P. Wonka, Large-scale architectural asset extraction from panoramic imagery, IEEE Trans. Visual Comput. Graphics 28 (2) (2022) 1301–1316. https://doi.org/10.1109/TVCG.2020.3010694
[167] Z. Lu, T. Wang, J. Guo, W. Meng, J. Xiao, W. Zhang, X. Zhang, Data-driven floor plan understanding in rural residential buildings via deep recognition, Inf. Sci. 567 (2021) 58–74. https://doi.org/10.1016/j.ins.2021.03.032
[168] Y. Pan, A. Braun, A. Borrmann, I. Brilakis, 3D deep-learning-enhanced void-growing approach in creating geometric digital twins of buildings, Proc. Inst. Civ. Eng. Smart Infrastruct. Constr. 176 (1) (2022) 24–40. https://doi.org/10.1680/jsmic.21.00035
[169] B. Yu, J. Hu, X. Dong, K. Dai, D. Xiao, B. Zhang, T. Wu, Y. Hu, B. Wang, A robust automatic method to extract building facade maps from 3D point cloud data, Remote Sens. 14 (16) (2022). https://doi.org/10.3390/rs14163848
[170] B. Haznedar, R. Bayraktar, A.E. Ozturk, Y. Arayici, Implementing PointNet for point cloud segmentation in the heritage context, Heritage Sci. 11 (1) (2023) 2. https://doi.org/10.1186/s40494-022-00844-w
[171] B.G. Pantoja-Rosero, R. Achanta, K. Beyer, Automated image-based generation of finite element models for masonry buildings, Bull. Earthquake Eng. 22 (7) (2024) 3441–3469. https://doi.org/10.1007/s10518-023-01726-7
[172] M. Dai, J. Jurczyk, H. Arbabi, R. Mao, W. Ward, M. Mayfield, G. Liu, D.D. Tingley, Component-level residential building material stock characterization using computer vision techniques, Environ. Sci. Technol. 58 (7) (2024) 3224–3234. https://doi.org/10.1021/acs.est.3c09207
[173] M. Escalada, Leveraging deep learning segmentation techniques and connected component analysis to automate high-level cost estimates of facade retrofits using 2D images, VITRUVIO Int. J. Archit. Technol. Sustainability 9 (2) (2024). https://doi.org/10.4995/vitruvio-ijats.2024.22421
[174] C. Fotsing, W.C. Tchuitcheu, L.I. Besong, D.W. Cunningham, C. Bobda, A specialized pipeline for efficient and reliable 3D semantic model reconstruction of buildings from indoor point clouds, J. Imaging 10 (10) (2024). https://doi.org/10.3390/jimaging10100261
[175] S. Li, F. Yan, K. Ma, Q. Hu, F. Wang, W. Liu, Optimal feature-guided position-shape dual optimization for building point cloud facade detail enhancement, Remote Sens. 16 (22) (2024). https://doi.org/10.3390/rs16224324
[176] M. Mehranfar, M.A. Vega-Torres, A. Braun, A. Borrmann, Automated data-driven method for creating digital building models from dense point clouds and images through semantic segmentation and parametric model fitting, Adv. Eng. Inf. 62 (2024) 102643. https://doi.org/10.1016/j.aei.2024.102643
[177] Z. Duan, K. Lu, M. Li, S. Zhang, B. Lin, Z. Wang, Automating exterior BIM for existing buildings using GIS data and UAV-based 3D modeling, Autom. Constr. 178 (2025) 106422. https://doi.org/10.1016/j.autcon.2025.106422
[178] Y. Zhong, J. Jiang, W. Quan, M. Zhao, D.-m. Yan, Distinctive learning of latent space feature for occlusion-aware facade parsing, Build. Environ. 279 (2025) 112955. https://doi.org/10.1016/j.buildenv.2025.112955
[179] N. Forouzandeh, J. Huang, L. Nan, E. Brembilla, J. Stoter, Semi-automated indoor geometry reconstruction for daylight simulation, Build. Environ. 289 (2026) 114045. https://doi.org/10.1016/j.buildenv.2025.114045