ICCEPM 2024 https://dx.doi.org/10.6106/ICCEPM.2024.0658
The 10th International Conference on Construction Engineering and Project Management
Jul. 29-Aug.1, 2024, Sapporo

Automated Generation of BIM Models with Indoor Spaces Using Street View Façade Images

Joonho Jeong[1], Sohyun Kim[2], Junwoo Park[3], Jungmin Lee[4], Kwangbok Jeong[5], Jaewook Lee[6]*

[1] *Graduate Research Assistant, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, Seoul, 05006, Republic of Korea. E-mail address: 23110167@sju,ac.kr*
[2] *Graduate Research Assistant, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, Seoul, 05006, Republic of Korea. E-mail address: thgus960221@sju.ac.kr*
[3] *Graduate Research Assistant, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, Seoul, 05006, Republic of Korea. E-mail address: 21150053@sju.ac.kr*
[4] *Graduate Research Assistant, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, Seoul, 05006, Republic of Korea. E-mail address: jungmin914@sju.ac.kr*
[5] *Assistant Professor, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006, South Korea. E-mail address: kbjeong7@sejong.ac.kr*
[6] *Professor, Deep Learning Architecture Research Center, Department of Architectural Engineering, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006, South Korea. E-mail address: jaewook@sejong.ac.kr (*CORRESPONDING AUTHOR)*

Abstract: The importance of 3D city models for sustainable urban development and management is underscored, but existing models often overlook indoor spaces and attribute information. This issue can be tackled with BIM models, though the conventional method requires accurate and extensive information, incurring considerable time and cost in data collection and processing. To overcome these limitations, this study proposes a method to automatically generate BIM models that include indoor spaces using street view images. The proposed method uses YOLOv5 to identify façade elements and DBSCAN to normalize façade layouts, facilitating the generation of detailed BIM models with a parametric algorithm. To validate the method, a case study of a building in Korea was conducted. The results showed that indoor spaces similar to the actual building were generated, with an error rate of object quantities between 8.46% and 9.03%. This study is anticipated to contribute to the efficient generation of 3D city models that incorporate indoor spaces.

Key words: Street view image, Building information modeling, 3D city model, Façade elements, Parametric algorithm

1. INTRODUCTION

As urbanization progresses globally, more than 55% of the world's population now resides in urban areas, and this number is expected to increase further by 2050. This urbanization trend leads to various issues, such as urban environmental pollution, traffic congestion, and housing problems, which require sustainable development and efficient management of cities [1]. In particular, the importance of 3D city models has been emphasized to comprehensively understand and analyze complex urban environments with multiple buildings.Thus, research to create 3D city models has been actively conducted in recent years. For instance, Chun and Guldmann (2012) created a 3D city model by estimating the surface temperature and the building height and forming buildings based on data from satellite, LIDAR, and

GIS. This allowed them to analyze the urban heat island phenomenon by considering the spatial interaction between buildings [2]. Oh et al. (2007) generated a geometric model of a building based on building layers from existing digital maps and then matched the texture of the building extracted from spatial imagery (e.g., aerial photos, satellite images) to create a 3D city model [3]. However, one limitation of these studies is that they only consider the geometric information of the building (e.g., building shape and building texture) and do not reflect the attribute information (e.g., materials and building use).

Building information modeling (BIM) is a technology that digitally represents the information (e.g., cost, schedule, and structure) that occurs throughout the life cycle of a building, enabling the creation of a 3D building model that reflects the building's geometric and attribute information [4]. Park et al. (2022) performed semantic segmentation on point clouds obtained from laser scanners to classify objects and applied object space relationships to automatically generate BIM models [5]. In addition, Kim et al. (2023) proposed a BIM model generation method using point clouds and panorama images to consider the building material of an existing building. The 2D material data taken from the panorama image through material recognition and classification was converted into 3D data to generate a BIM model reflecting the material information [6]. However, the data used in BIM models generation methods is precise and large, leading to significant time and cost expenditures in the data collection and preprocessing.

Meanwhile, street-view images taken by specially equipped vehicles or cameras are widely used in map services and applications to help users get directions, navigate buildings, and understand their surroundings [7]. With the advancement of computer vision and deep learning technologies, street-view images have been extensively used in various fields, especially in architecture, engineering, and construction (AEC) for research to generate 3D models of buildings. However, existing studies have mainly considered only building façade information (e.g., façade shape and building height) using street-view images, and there are limitations in generating BIM models, including indoor spaces of buildings [8-9].

Therefore, this study proposes a method to recognize façade elements (i.e., windows and doors) of a building using street-view images and automatically generate a BIM model, including the indoor space of a building through real estimation, zoning, etc., based on these elements. First, façade elements were recognized through deep learning-based detection from street-view images. Second, normalization using clustering was performed to extract façade layouts suitable for the generation of BIM models. Third, a parametric algorithm was used to generate a BIM model based on the normalized façade layout (Figure 1).

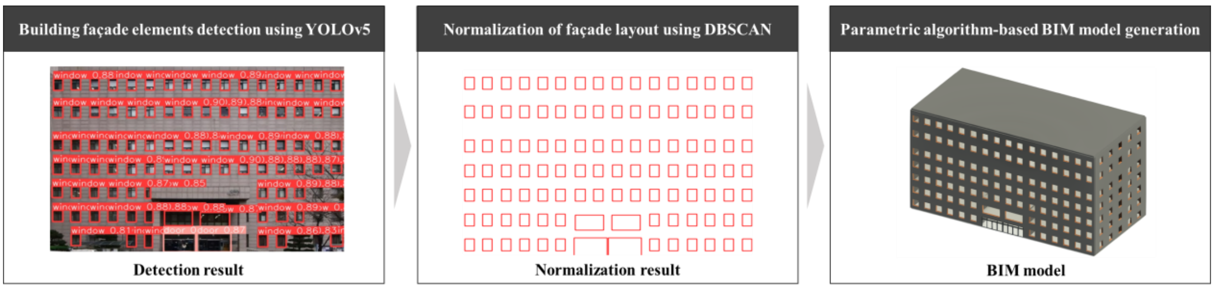


Figure 1. Research framework

2. Materials and Method

2.1 Building façade elements detection using YOLOv5

A street-view image contains a façade element, which includes the information needed to create a BIM model, such as the building's geometric shape, exterior wall finishes and structure, and the location of windows and doors. In particular, windows and doors are façade elements crucial to the floor height of a building and the division of indoor space. Thus, this study used the YOLO v5, a deep learning-based object detection model, to recognize windows and doors from street-view images. YOLO is a one-stage detection model that can detect objects in an image. It analyzes the entire image at once to calculate the class and confidence score of the detected object and provides its exact location and size through the bounding box surrounding it [10]. The model's performance was evaluated using the main metrics of Precision, Recall, and F1-score. Precision is the percentage of true predictions that are actually true, Recall is the percentage of true predictions predicted by the model that are actually true, and F1-score is the harmonized average of Precision and Recall (Eq. 3–Eq. 5).

$$Precision = \frac{True\ detections}{whole\ detections\ of\ an\ algorithm} = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{True\ detections}{whole\ detections\ of\ an\ algorithm} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

where, *TP* is the true positive, *FP* is the false positive, and *FN* is the false negative.

2.2 Normalization of façade layout using DBSCAN

Street-view images can be noisy due to various external factors, such as vegetation, structures, shadows, and light glare, and at times, noise interferes with the accurate determination of a building's façade layout. Specifically, external factors have the following effects when detecting façade elements from street-view images: First, façade features can be occluded by vegetation, structures, and the like, which can result in undetected façade features during detection (Figure 2(a)). Second, shadows and light smearing can complicate the extraction of the façade elements' exact size and location (Figure 2(b)). Consequently, generating a BIM model based on the noise-affected façade layout may change the actual building appearance. To address this, this study uses clustering to normalize the size and position of the façade elements in the façade layout. Clustering is a technique that classifies clusters based on the degree of similarity of data from unsupervised learning. This study particularly used density-based DBSCAN, one of the representative techniques of clustering, which forms a cluster of data within a certain radius if there is a minimum number of neighboring data (MinPts) to form a cluster within a given radius (Eps) for each data [11].

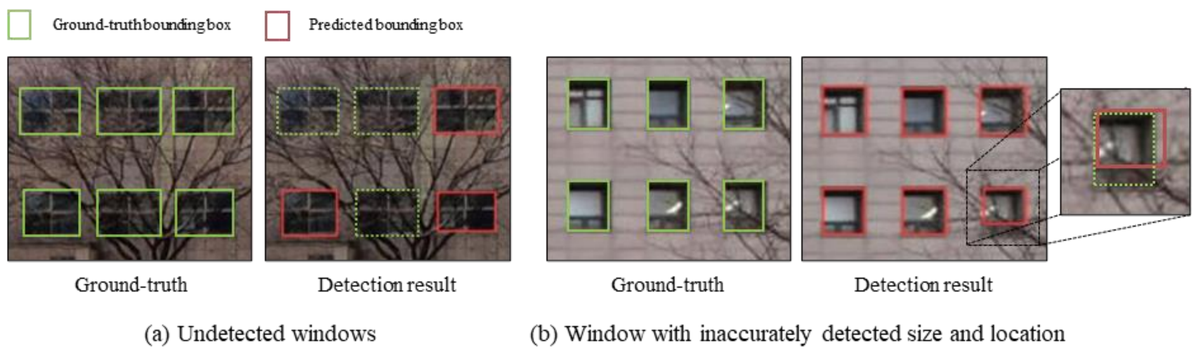
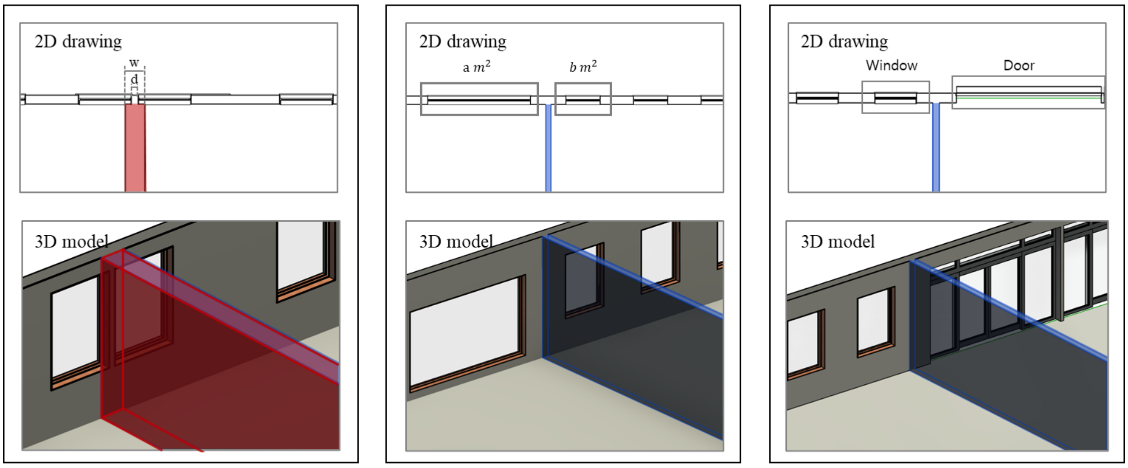


Figure 2. Impact of noise in the detection process

2.3 Parametric algorithm-based BIM model generation

This study developed a dynamo-based parametric algorithm to generate BIM models based on façade layout. This algorithm utilizes different details (i.e., class, location, size) of the bounding box that constitutes the façade layout and the building footprint provided by the spatial information open platform (i.e., V-World) [12] to generate a BIM models including the building's façade and indoor space. First, to generate the façade, the pixel-level façade layout was converted to the scale of the actual building. Then, the class, location, and size of the bounding box and the location of the slab were used to generate the façade of the BIM models with the façade elements.

Also, to generate the indoor space based on the façade of the BIM models, the indoor space was estimated according to the location of the façade elements. Based on the relationship between façade elements in the estimated space, the space was zoned according to three cases (Figure 3). First, if the distance (d) between two elements is smaller than the thickness (w) of the interior wall (case 1), no wall is created between the elements. Second, if two neighboring elements have different sizes (case 2), a wall is created between them. Third, if two neighboring façade elements are of different types (case 3), a wall is created between them. In the case of the corridor, it is created to be located in the center of the building as it is a passage connecting different spaces inside the building and is usually long in shape. After this process, a BIM model including the façade and interior was created based on the façade layout from the street-view image.



(a) Case 1 (b) Case 2 (c) Case 3

Figure 3. Three cases of indoor space partitioning

3. Result and discussion

3.1. Validation of façade elements detection using YOLOv5

To validate the proposed method, this study conducted a case study on street-view images collected in South Korea. First, 240 building façade images were gathered from the street view provided by NAVER Maps [13] to build data for learning. Then, windows and doors were labeled to build a learning dataset, which was used to train the YOLO v5 model. The trained YOLO v5 model detected building façade elements (i.e., windows and doors) in the street-view images. As a result, the F1-score, the harmonized average value of precision and recall, showed an average accuracy of 87.8% (i.e., precision: 89.1%, recall: 86.7%). When examined by façade elements, windows were 9.7% more accurate than doors (i.e., window: 92.7%, door: 83.0%) because the number of windows in a building is relatively higher than the number of doors, so more training data was utilized to detect windows, and the YOLOv5 model learned the characteristics of windows more accurately. These results indicate that detecting façade elements of buildings from street-view images with a high level of accuracy is feasible.

Table 1. Detection results of facade elements using YOLOv5

Façade elements	Precision(%)	Recall(%)	F1-score(%)
Window	94.2	91.3	92.7
Door	83.9	82.1	83.0
Average	89.1	86.7	87.8

3.2. Validation of normalization for facade layout using DBSCAN

DBSCAN was used to extract the normalized façade layout from the detection results, conducted in the following steps (Figure 5). In the first step, the size of the incorrectly detected bounding box was corrected according to the bounding boxes of the same class by clustering. The second step has the position of the incorrectly detected bounding boxes corrected by clustering to classify the rows and columns of the façade layout. The third and last step involves the bounding boxes of the undetected façade elements generated by placing a grid based on the size and position corrected bounding boxes. As a result, the positions and sizes of the façade elements were normalized, implying that a normalized façade layout can be extracted from the façade elements detected by the YOLO v5 model using DBSCAN.

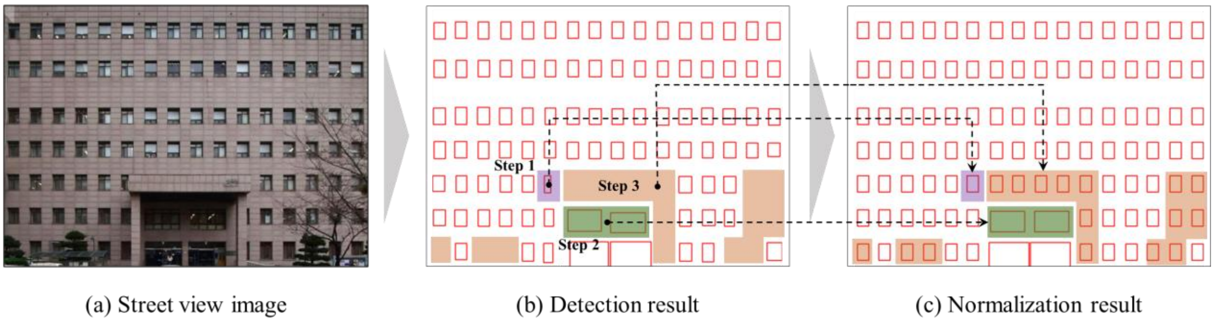


Figure 5. Facade layout normalization result

3.3. Validation of automated BIM model generation

Based on the façade layout extracted from the street-view image, a BIM model was generated using a parametric algorithm and compared with the 2D drawing–based BIM model (Figure 6). As a result, the BIM model generated from the street-view image was verified to have windows and doors in the same number and positions as the actual building. In addition, four rooms and one corridor were generated for the indoor space, similar to the 2D drawing–based BIM model. In the case of the corridor, it borders all rooms and is connected to the entrance and exit. However, because the parametric algorithm generates indoor spaces based on rules related to the elements' location and size, it has a limitation in reflecting structural features, such as the core.

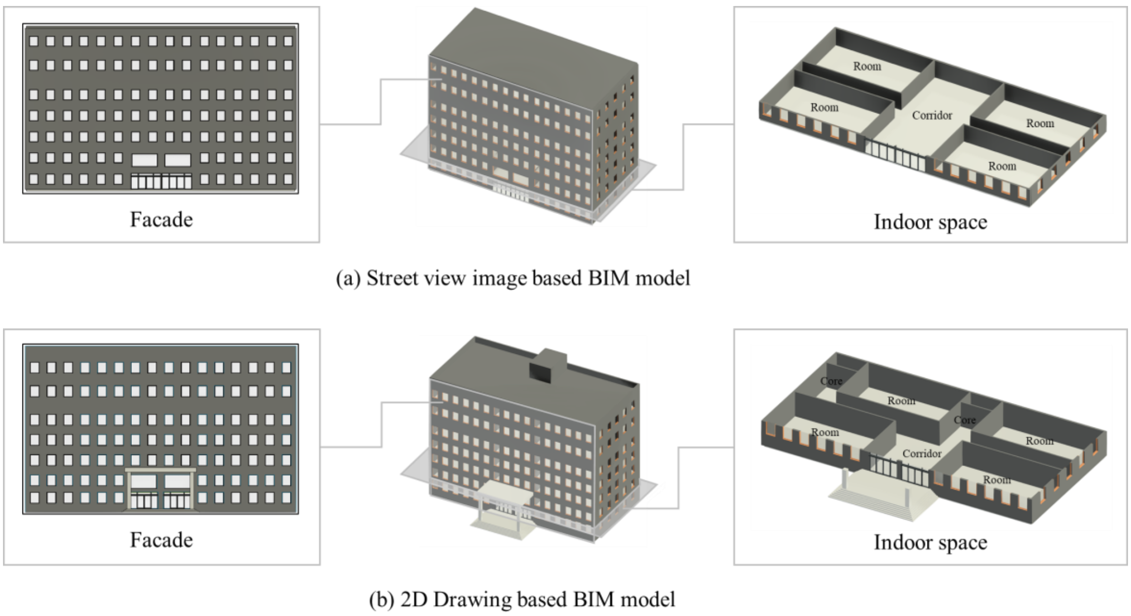


Figure 6. Comparison of BIM models based on street view images and 2D drawings

Meanwhile, to quantitatively analyze the BIM model generation accuracy of the proposed method in this study, the quantities of two BIM model types (i.e., street view image–based BIM model and 2D drawing–based BIM model) were compared. The quantities were calculated for opening objects (i.e., window and door) and structural objects (i.e., wall and slab). Table 2 shows the result of comparing the volume of the 2D drawing–based BIM model with the volume of the street view image–based BIM model. The mean absolute percentage error (MAPE) values, calculated as the average of the absolute percentage differences, exhibited lower values for opening objects at 9.03% and for structural objects at 8.46%. In particular, the differences in quantities for doors and slabs were minimal, at 2.77% and

2.29%, respectively, when compared to the 2D drawing–based BIM model, suggesting that this study's proposed method makes it possible to generate a BIM model using street-view images that closely resemble the BIM model created based on 2D drawings.

Table 2. Comparison of quantity: street view image vs. 2D drawing-based BIM models

Object	Class	2D Drawing	Street view image	Diffrence(%)	MAPE(%)
Opening	Window	311.2 m^2	263.63 m^2	(-)15.29	9.03
	Door	25.31 m^2	26.01 m^2	(+)2.77	
Structure	Wall	1567.81 m^3	1338.39 m^3	(-)14.63	8.46
	Slab	1658.47 m^3	1696.46 m^3	(+)2.29	

4. Conclusion

This study proposes a method for automatically generating BIM models with indoor spaces using street view images. To validate the proposed method, accuracy analysis of detection, analysis of normalization results, and comparison of BIM models quantities were conducted. The results revealed that street-view images can be used to generate BIM models with indoor spaces similar to real buildings. The study’s major findings are as follows: First, using easily accessible street- view images can solve the costly and time-consuming problem of the conventional development of BIM models. Second, BIM models including indoor spaces can be generated based on the location, size, and other aspects of the façade elements. However, since this study only considers façade elements to generate a BIM model, it cannot generate indoor space by considering the space associated with structural characteristics (i.e., core). Therefore, future research is recommended to generate BIM model that considers not only façade elements but also other vital factors, such as structural characteristics, building shape, and building use. This study proposes a novel approach for generating BIM models from street view images, which is expected to contribute to the creation of 3D city models that can aid in the sustainable development and efficient management of cities.

ACKNOWLEGEMENTS

This research was supported by grants from the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (NRF-2020R1A2C1010421) and the Ministry of Education (RS-2023-00271991) of the Korean government.

REFERENCES

[1] A. Sharifi, Z. Allam, S. Bibri, A. Khavarian-Garmsir, "Smart cities and sustainable development goals (SDGs): A systematic literature review of co-benefits and trade-offs", Cities, vol. 146, 104659, 2024.

[2] B. Chun, J. Guldmann, "Spatial analysis of the urban heat island using a 3-D city model", Spatial Information Research, 20(4), 1-16. 2012.

[3] J. Oh, S. Shin, J. Park, H. Lee, "3D Building Model Texture Extraction from Multiple Spatial Imagery for 3D City Modeling", Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, 25(4), 347-354, 2007.

[4] R. Volk, J. Stengel, F. Schultmann, "Building Information Modeling (BIM) for existing buildings—Literature review and future needs", Automation in construction, 38, 109-127, 2014.

[5] J. Park, J. Kim, D. Lee, K. Jeong, J. Lee, H. Kim, T. Hong, "Deep learning–based automation of scan-to-BIM with modeling objects from occluded point clouds", Journal of Management in Engineering, 38(4), 04022025, 2022.

[6] S. Kim, K. Jeong, T. Hong, J. Lee, J. Lee, "Deep Learning–Based Automated Generation of Material Data with Object–Space Relationships for Scan to BIM", Journal of Management in Engineering, 39(3), 04023004, 2023.

[7] C. Wang, S. Antos, L. Triveno, "Automatic detection of unreinforced masonry buildings from street view images using deep learning-based image segmentation", Automation in Construction, 132, 103968, 2021.

[8] Y. Yan, B. Huang, "Estimation of building height using a single street view image via deep neural networks", ISPRS Journal of Photogrammetry and Remote Sensing, 192, 83-98, 2022.

[9] J. Kang, M. Körner, Y. Wang, H. Taubenböck, X. Zhu, "Building instance classification using street view images", ISPRS journal of photogrammetry and remote sensing, 145, 44-59, 2018.

[10] J. Redmon, S. Divvala, R. Girshick, A, Farhadi, "You only look once: Unified, real-time object detection", In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788), 2016.

[11] M. Ester, H. Kriegel, J. Sander, X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise", In kdd (Vol. 96, No. 34, pp. 226-231), 1996

[12] V-World. Available online : https://www.vworld.kr (Jan. 10. 2024)

[13] Naver Maps. Available online : https://map.naver.com (Jan. 10. 2024)