2023 European Conference on Computing in Construction
40th International CIB W78 Conference
Heraklion, Crete, Greece
July 10-12, 2023

HARBINGERS OF NeRF-TO-BIM: A CASE STUDY OF SEMATIC SEGMENTATION ON BUILDING STRUCTURE WITH NEURAL RADIANCE FIELDS

Shun Hachisuka, Alberto Tono, and Martin Fisher
Stanford University, Stanford, CA

Abstract

Scan-to-BIM applications rely on point clouds obtained by laser scan, which require expensive hardware and laborious tasks. To address this issue, we introduce a NeRF-to-BIM approach, exploiting recent advancements in computer vision with Neural Radiance Fields (NeRF). NeRF is a state-of-the-art (SOTA) for 3D scene reconstruction from 2D images but lacks specific applications in the architecture, engineering, and construction (AEC) domain. We propose a 3-step approach: (1) 3D reconstruction of buildings using NeRF. (2) Semantic segmentation by fine-tuning pre-trained deep learning (DL) algorithm. (3) Conversion from the semantic segmentation point cloud to BIM. Finally, qualitative and quantitative analyses are performed.

Introduction

The process of capturing the space as point cloud data and translating the data into Building Information Modeling (BIM) is generally called Scan-to-BIM. This process is typically used in the architecture, engineering, and construction (AEC) industry to create detailed, accurate, and up-to-date representations of existing buildings. Especially in the construction industry, using as-built models can be a powerful collaborative tool for planning, quality control, and real-time progress tracking with team members (Golparvar-Fard, Peña-Mora & Savarese 2011, Tserng et al. 2014, Matthews et al. 2015, Fobiri 2021, Fobiri et al. 2022). However, there are still challenges and limitations when capturing point clouds and converting them to BIM because the process requires experienced professionals and heavy human intervention (Golparvar-Fard, Bohn, Teizer, Savarese & Peña-Mora 2011). There is a growing interest in automating the Scan-to-BIM process to address these issues. The process of converting point cloud data to a BIM model can be divided into three key aspects: data acquisition and pre-processing, object segmentation, and reconstruction to BIM (Ma et al. 2020).

In the data acquisition step, laser scanning and photogrammetry are currently widely used to create digital representations of physical buildings and sites. While laser scanning allows for obtaining highly accurate data, it involves laborious tasks and requires specialized skills and expensive equipment. On the other hand, photogrammetry is generally faster and more cost-effective than laser scanning as it does not require specialized equipment. However, the accuracy of photogrammetry is lower than laser scanning (Moon et al. 2019), and it is a complex process that requires specialized software and expensive expertise.

After the point cloud data is collected, the point cloud is segmented to classify specific objects. The segmentation can be edge-based, region-growing, model-fitting, hybrid, and machine learning(Grilli, Menna & Remondino 2017). Once a point cloud has been segmented, each segment of points is labeled to identify the class. Based on the assigned semantic labels for individual point cloud objects, BIM objects can be modeled. The class labeling procedure is typically achieved using one of three approaches: supervised, unsupervised, and interactive (De Geyter et al. 2022).

Supervised learning algorithms are commonly used to classify architectural and structural elements of buildings. The use of deep learning (DL) methods for semantic segmentation of 3D point clouds is also a growing research topic. However, there are relatively few studies on deep learning (DL) applied to buildings, which results in a limited amount of available datasets related to buildings.

The final step is the reconstruction of the BIM model from a semantically segmented point cloud. The process of converting point cloud data to BIM models involves simplifying the data into a vector-based geometric model and adding relevant BIM metadata, as point cloud data typically contains a high level of geometric detail. (Qu et al. 2014). This process is referred to as the Scan-to-BIM process.

In 2020, NeRF was introduced (Mildenhall et al. 2020). NeRF uses neural networks to create realistic view synthesis on a collection of 2D images as input. NeRF has demonstrated higher performance in capturing 3D objects, especially in cases where photogrammetric techniques do not allow satisfactory results to be achieved. (Condorelli et al. 2021)

To address the data acquisition challenges in Scan-to-BIM, such as the need for expensive hardware and labor-intensive tasks, we propose a new approach called NeRF-to-BIM. This method uses semantic segmentation algorithms on point cloud data created with NeRF technology. After implementing this approach, the results are evaluated and the potential for obtaining a BIM representation while reducing manual efforts is discussed.

Related Works

In this section, the related works for the following key steps are discussed: (1) NeRF (Data acquisition), (2) Semantic segmentation, (3) BIM reconstruction.

(1) NeRF (Data Acquisition)

NeRF was first introduced by Mildenhall et al. (2020) in 2020, a ground-breaking method that represents realistic 3D scenes from a sparse set of input collections of 2D images. This original NeRF algorithm had several disadvantages, such as slow training caused by querying a deep MultiLayer Perceptron (MLP) millions of times (Reiser et al. 2021). In recent years, there has been significant research in the computer vision community to improve upon the original NeRF algorithm, with more than 50 papers being published on the topic in 2022 alone (Gao, Gao, He, Lu, Xu & Li 2022).

While many new NeRF models are published, NVIDIA and Luma Labs translated the research into practical applications to create 3D scenes easier. NVIDIA offers Instant-NeRF (Müller et al. 2022) as software, which features a graphical user interface (GUI) allowing users to control visualization options. The Instant-NeRF algorithm shortens rendering time through a technique called multi-resolution hash grid encoding. This novel input encoding technique produces high-quality output with a fast-running, compact neural network. The Luma AI app was developed by Luma Labs in 2022, simplifying the entire process of rendering 3D scenes using NeRF technology. From start to finish, the entire process can now be managed be managed using only a smartphone, even for 3D images captured casually with a smartphone camera.

(2) Semantic segmentation on 3D point clouds

3D point clouds semantic segmentation classifies point clouds into multiple regions. The classified points in the same region can be derived with similar properties, essential for as-built BIM reconstruction. While semantic segmentation in 2D image analysis has advanced, segmentation in point clouds remains challenging due to high redundancy, uneven sampling density, and a scarcity of labeled point cloud data (Gao, Pan, Li, Geng & Zhao 2022).

At this time, while there are annotated datasets such as Replica (Straub et al. 2019) and Habitat-Matterport 3D (HM3D) (Ramakrishnan et al. 2021), the Stanford Large-Scale 3D Indoor Spaces (S3DIS) dataset (Xu et al. 2020) is the most trained in various studies among the dataset related to building. The S3DIS dataset comprises a large-scale indoor environment, including six indoor areas with 271 rooms for a total of 695 million points. Each point in the scene point cloud is annotated with one of the 13 semantic categories, which are structural elements (ceiling, floor, wall, beam, column, window, and door), furniture (table, chair, sofa, bookcase, and board) and clutter for all other elements. Given this shortage of datasets, the focus is set on deep learning algorithms trained on the S3DIS dataset.

PointNet (Qi, Su, Mo & Guibas 2017) is the first promising algorithm that feeds point clouds directly into the DL architecture. This algorithm is a ground-breaking solution that addresses the challenges associated with large data in a point cloud format. PointNet++ (Qi, Yi, Su & Guibas 2017) resolves the disadvantage of PointNet, which fails to capture the local structure and generalize to complex scenes. PointNeXt (Qian et al. 2022) improved training strategies based on the classical PointNet++ through a systematic study of model training and scaling strategies. The inverted residual bottleneck design and separable MLPs into PointNet++ enable efficient and effective model scaling. PointNeXt established a new state-of-the-art performance with 74.9% mean Intersection-over-Union (IoU) on S3DIS in September 2022.

In the AEC industry, semantic segmentation methods are becoming a crucial area of focus for building implementation. The initial implementation of deep learning algorithms is for parsing building facades (Corydon et al. 2016, Liu et al. 2017) and urban scenes segmentation (Niemeyer et al. 2016, Grilli, Barabás, Michalska-Smith & Allesina 2017, Weinmann et al. 2015, Hackel et al. 2016).

Following these studies, Murtiyoso et al. (2022) developed the implementation of a deep learning-based semantic image segmentation method for photogrammetric 3D reconstruction and classification workflows. Cao & Scaioni (2022) proposed a pre-training method for 3D building point cloud that learns from a large source dataset and evaluates the proposed method by employing four fully supervised networks as backbones. The end-to-end deep learning method is proposed by Perez-Perez et al. (2021), named Scan2BIM-NET. The method was trained and tested for semantically segmenting the structural, architectural, and mechanical components in point cloud data. The study achieved an average accuracy of 86.13% using 83 rooms from point cloud data representing real-world industrial and commercial buildings.

As mentioned above, acquiring a labeled point cloud dataset is challenging.The use of synthetic datasets created from 3D models is gaining interest as a means of acquiring training data for semantic segmentation. The dataset generated from a virtual environment can be produced with lower costs and less manual effort for data creation.

Fedorova et al. (2021) constructed a field-specific synthetic 3D data generation pipeline to tackle the problem of insufficient 3D scanning and modeling datasets. This framework is suitable for multiple deep learning tasks, including geometric deep learning that requires direct 3D supervision. However, the effort required for 3D modeling should also be taken into account. Ma et al. (2020) investigates the viability of using synthetic point clouds generated from BIM to train deep neural networks to perform semantic segmentation of point clouds on building interiors. The performance increased by 7.1% IOU when synthetic point clouds were used for training, compared to training the classifier on the real data alone.

(3) Conversion from point cloud to BIM

The conversion process in the construction industry still needs to be fully automated and the research on this topic is in its infancy. (Wang & Xiong 2021). There are several studies that propose different approaches to automate the process.

Jung et al. (2016) proposes a methodology for creating a 3D wire-frame model of indoor surfaces. In this study, two indoor scenes are captured using a laser scanner with an accuracy range of ±2mm @25m. The wall boundary is traced from the point cloud projected onto 2D x-y coordinates, and the boundary lines are extracted using the least-squares method. The height is estimated using a 2D floor boundary map obtained through the incorporation of RANdom sample consensus (RANSAC). Armeni et al. (2016) proposes a detection-based semantic parsing method for large-scale building point clouds. In the study, space dividers (i.e. walls) are first detected as the boundaries of a room to parse a raw point cloud into disjoint spaces. To detect the wall from the point cloud, a one-dimensional histogram of the density of points along the x-axis is analyzed. Then, the wall feature is identified with the signature of two peaks with an empty space in-between. After this process, those spaces are parsed into their semantic structural (e.g. floor, walls, etc.) and building (e.g. furniture) elements. Chen et al. (2019) identifies and categorizes building elements from laser scans by applying the deep learning method. The classified point cloud is converted into bounding boxes and matched with relevant BIM entities. Croce et al. (2021) presents a semi-automatic approach to the 3D reconstruction from scanned point clouds based on machine learning techniques. The approach is reconstructing 3D geometry from point clouds via the RANSAC algorithm included in the built-in option of Cloud Compare software after generating semantically labeled data. Bassier et al. (2020) implements the reconstruction of the wall geometry from point clouds with 3D and 2D reconstruction methods. The parameters, including the orientation, thickness, location, and boundary of the walls, were extracted based on the 3D point clusters and histograms, respectively. The two methods detected wall geometry with high precision.

Research Methodology

We propose a NeRF-to-BIM method, which is a 3-step approach (Figure 1): (1) 3D reconstruction of buildings using the Neural Radiance Fields (NeRF) algorithm. (2) Semantic segmentation using pre-trained and fine-tuned deep learning algorithms. (3) Conversion of the semantically segmented point cloud into BIM. The aim of this study is to verify the feasibility of the NeRF-to-BIM approach by implementing it on simple structural objects consisting of columns and beams to simplify the testing and evaluation process.

(1) 3D reconstruction of buildings using NeRF

In this step, two NeRF-based applications, Instant NeRF and Luma AI, are tested to create point clouds and evaluate the performance of these applications. After generating the 3D scenes with these applications, the point cloud

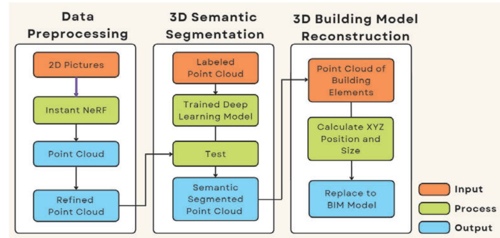

Figure 1: Framework of NeRF-to-BIM

is segmented and labeled as the ground truth with structural elements, such as floor, beam, column, and ceiling, to evaluate and analyze the data.

(2) Semantic segmentation

In the semantic segmentation step, the PointNeXt algorithm is used, a SOTA trained on the S3DIS dataset. Presently, the S3DIS dataset is the most trained in the 3D point cloud segmentation field. However, the S3DIS dataset used in this study needs to provide more quality results since the structural features of the buildings differ from the interior features present in the S3DIS dataset. Fine-tuning is performed on pre-trained deep learning (DL) algorithms to address this minor dataset issue to extract robust features from unlabelled building data. When the dataset is relatively small, fine-tuning is a fundamental approach to improving the model's performance by training a pre-trained model on a new dataset. In the AEC field, synthetic point clouds generated from BIM are expected to improve the network's performance as a new dataset (Ma et al. 2020). Therefore, the point cloud is tested using pre-trained weights trained with S3DIS and then the PointNeXt model is fine-tuned using labeled point clouds generated from 3D models. To evaluate our approach, the semantic segmentation accuracy is tested based on two different capturing methods and two different trained datasets to evaluate the performance (Figure 2).

(3) Conversion from point cloud to BIM

After semantic segmentation, the classified building elements are transformed into a vector-based geometric model with BIM metadata (Qu & Sun 2015). In the Scan-to-BIM process, as shown in the related works section, the point cloud is captured using laser scanning in related works and used fitting techniques such as least-squares and RANSAC (Jung et al. 2016, Armeni et al. 2016, Croce et al. 2021). Since these studies are based on highly accurate point cloud data gained with the laser scanner, the conversion method in this study must be properly selected for the NeRF-to-BIM process by observing the feature of the NeRF output.

Experimental Results

(1) Reconstruction of the 3D scenes using NeRF

The first step is to obtain a simple structural object from 2D pictures by applying Instant-NeRF and Luma AI. For Instant-NeRF, 104 images captured with a smartphone

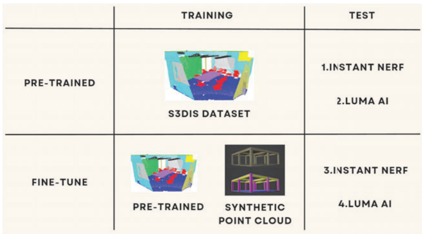

Figure 2: Methodology of test

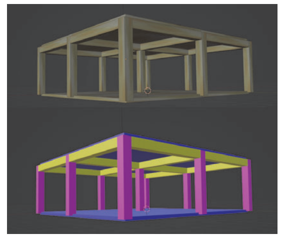
Figure 3: 3D model for training

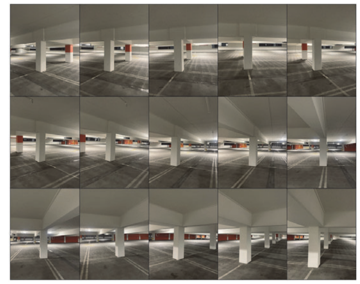
Figure 4: Partial set of pictures for NeRF

Table 1: Feature comparison between NeRF and Laser scan

Method	Cost	Rebuild speed	Accuracy
NeRF	0*	20–60 s(2)	10mm< @560mm
Laser scan	$3,813 - $103,700(1)	20–30 min(2)	2mm @25m(3)

* Require smartphone for picture. Source (1) Bi et al. (2021) (2) Pei et al. (2022) (3) Jung et al. (2016)

were prepared (Figure 4). The pictures' original resolutions are 4,032 by 3,024 pixels. A higher picture resolution and Axis Aligned Bounding Box (AABB) scale result in insufficient GPU memory, while a lower AABB scale and picture resolution create scarce point clouds and lead to more noise in the output. To be processed well with NVIDIA GeForce RTX 3070 Laptop GPU, the pictures were resized to 1,440 by 1,080 pixels and used 8 for the AABB. The camera position and output results are shown in Figure 5 and 6, respectively. There is excessive noise in the background around the target object due to insufficient photos in such areas, while Instant-NeRF generates higher-quality beams and columns that were taken from multiple angles.

In terms of the Luma AI application, the same object as Instant-NeRF was captured. The application recognizes the camera's movement and position and automatically captures pictures at a certain distance. The pictures are sent to a web server, which returns the output after calculations. Figure 7 shows the output result of our object. Compared to the output of Instant-NeRF, there are far fewer artifacts behind the surface, and the point cloud is smoother. Table 1 compares NeRF and laser scanning methods. NeRF reduces the time and cost of generating 3D representations, while laser scanning has higher accuracy. However, two main challenges have been identified for NeRF-to-BIM. The first challenge is that NeRF creates points in the middle of space and behind the surface of objects, which does not occur when using laser scanning devices. The second challenge is that the extent of light and shadow affects the shape. NeRF creates points more accurately in areas of darker colors, whereas it creates points more sparsely in areas of lighter colors. In this case study, since the concrete structure frame is painted with glossy white paint, the output result of the ceiling and beams is particularly poor (Figure 6). The result was cropped for the semantic segmentation process to exclude the noisy area.

Figure 8 demonstrates the dimensional accuracy by showcasing the size of the columns. The results indicate that Luma AI produces dimensions that are approximately 100-200mm smaller than their actual values, while Instant-NeRF exhibits remarkable precision at the corners of the columns. Luma AI technical support indicated that this problem is caused by ARKit, which is a software development kit (SDK) developed by Apple that enables the creation of augmented reality (AR) applications for iOS devices and recommended putting a scale marker in the scene. Furthermore, both results show that the accuracy is lower in the plane sides of the column compared to the corner points. For Instant-NeRF, the measured distances between the corner points are 557mm (-3mm) and 553mm (-7mm), whereas the true value is 560mm. This difference in accuracy may be attributed to the fact that corners can be captured from multiple angles, while flat faces tend to reflect light, resulting in a loss of detail. These results indicate that NeRF output contains different accuracy levels of each point in one point cloud data, whereas the point cloud captured by laser scan produces a more consistently accurate result. If the same fitting techniques used in previous studies are utilized for the conversion to the BIM process, it is expected that the performance will be lower since the algorithms will include lower accuracy points. In other words, to enhance the performance of NeRF-to-BIM, it is the key to detect and select higher accuracy points for simplified geometric models, excluding lower accuracy point clouds.

(2) Applying semantic segmentation algorithm

To proceed with semantic segmentation training, six synthetic point clouds were prepared using the built-in option of the Cloud Compare software by scaling the model (shown in Figure 3) from 1.0 to 1.5. The data acquired in Step 1 was labeled into four structural elements: beams, columns, floors, and ceilings for evaluation. The two-point cloud datasets obtained in step 1 were tested using the PointNeXt algorithm with both pre-trained weights and fine-tuned weights using the additional synthetic dataset.

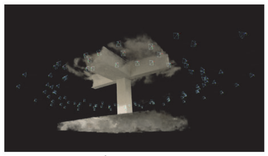

Figure 5: The camera position in 3D scene

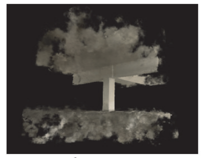

Figure 6: Execution of Instant-NeRF

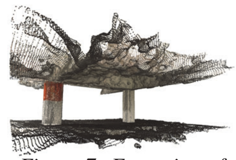

Figure 7: Execution of Luma AI

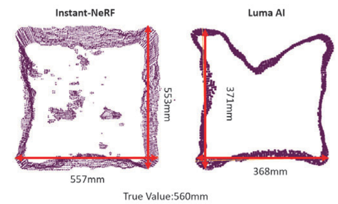

Figure 8: The comparison of accuracy on column size

Figure 9 illustrates the outputs tested on each training method and ground truth. When comparing the results obtained from the pre-trained deep learning model, the algorithm misclassifies columns and beams as other elements and fails to detect floors. On the other hand, the results from the fine-tuned deep learning model show an improved classification performance. Taking a closer look, table 2 shows the resulting accuracy percentage for each structural element in semantic segmentation. As can be seen from the mean Intersection-over-Union (mIoU) values, the performance significantly improved after fine-tuning. The Instant-NeRF point cloud results in lower performance for beams and columns. This is likely due to the significant amount of artifacts present around the borders between the column, beam, and ceiling. The point cloud created with Luma AI performed the best among our tests.

(3) Conversion from the semantic segmentation point cloud to BIM

The results of step 2 indicate that Luma AI achieved the best semantic segmentation performance. Even though it generates a model with a different scale, the result of Luma AI was used for the conversion process because of the higher semantic segmentation performance and the captured larger area.

The conversion process for this case study should be tailored in a different way from previous research, as the focus object and method of this research are different. Assuming NeRF-to-BIM can be used on the construction site, the standard grid-based column and beam structure for this study was used. Considering the features of the NeRF output as well, a four-step process is proposed: (1) Detect and set up the levels based on the semantically segmented point cloud of the floor and ceiling. (2) Set up the grid by detecting the center points of columns from the sliced horizontal sections. (3) Rotate the point clouds along the grid system and identify the beam size from the sliced section. (4) Reconstruct the BIM elements using the level and grid system, as well as the detected element sizes. In the first step, the floor and ceiling level is extracted from the resulting outputs of the point cloud segmented as 'Floor' and 'Ceiling' elements. Upon observing the section of the floor and ceiling point cloud height, the points were created even underneath the surface of objects. Considering tolerance and artifacts of NeRF output, it is assumed that a one-dimensional histogram of the density of points along the z-axis (as proposed by (Armeni et al. 2016)) would perform better in detecting the floor level than the fitting method, which would take into account tolerance and artifacts. By counting the number of Z-coordinates, the levels of the floor and ceiling were extracted at 0.19 m and 2.16 m, respectively.

In the second step, the center points of columns from the sliced horizontal section were extracted. In this case, the slice area is set from the floor level + 0.5 m to 1.5 m. To identify the center coordinates of columns, the points outside of the column are excluded with the same method in step 1, which counts the number of the same (x, y) coordinates and excludes the low number of (x, y) coordinates. After cleaning the point cloud, the Open Source Computer Vision Library (OpenCV) was used to detect the column shape. OpenCV is an open-source software library for image and video processing, including object detection and recognition, face detection, motion detection, and image segmentation. The detected column shapes are shown in Figure 10. Using these center points, the grids were created and rotated the whole point cloud alining the grids. In this case, two columns were used to find grids, however, when there are more columns, averaging and optimizing multiple grids are required.

Finally, the height and size of the beam between columns were detected by using the same method when the column shapes were detected. The bottom level and size of the beam were extracted as 1.59 m and (518mm, 570mm), respectively. All extracted numbers, such as levels, grids, size, and position of elements, were input to Dynamo and exported into the Revit model as shown in Figure 11.

Limitation and ethical considerations

This study focuses on a simplified demonstration of the NeRF-to-BIM process. The limitations of this research are that this process is implemented on the limited size of the column and beam building scene. There is a noted difficulty in acquiring data to implement the NeRF-to-BIM process on a larger scale due to the requirement for high-performance GPUs and the substantial amount of data generated. Additionally, the data transfer involved in the three-step approach is performed manually, necessitating additional effort to establish a streamlined pipeline. Furthermore, the NeRF model's quality significantly affects the conversion process, particularly concerning the accuracy of element shape and size extraction in areas where multi-angle photographs are not available. Finally, data security and confidentiality is also an issue to consider. The data generated with NeRF should have proper confidentiality measures to be restricted to only authorized persons with

Table 2: Semantic segmentation results (mIoU: mean Intersection-over-Union)

Experiments	mIoU	Ceiling	Floor	Beam	Column
1 Instant-NeRF & Pre-train	2.46	9.84	0.00	0.00	0.00
2 Luma AI& Pre-train	24.19	47.27	49.47	0.00	0.00
3 Instant-NeRF & Finetune	35.19	30.86	74.94	10.28	24.68
4 Luma AI& Finetune	57.53	59.52	72.2	57.87	40.54

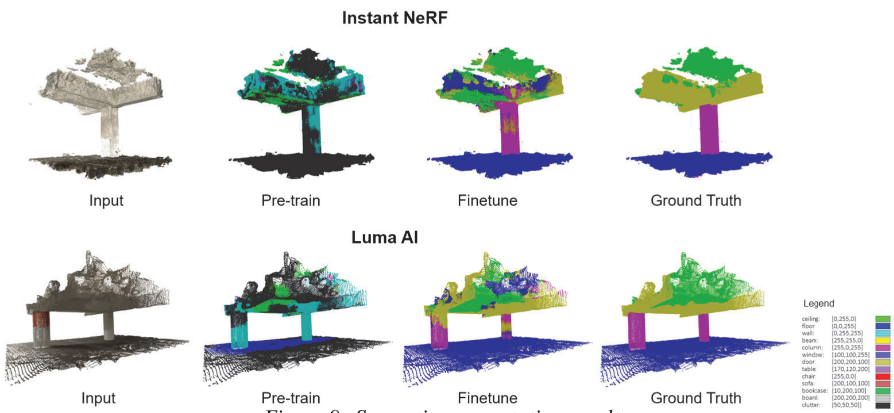

Figure 9: Semantic segmentation results

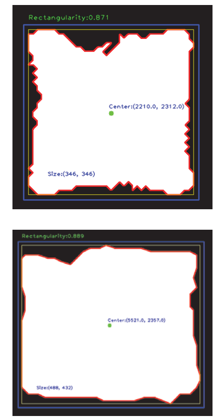

Figure 10: Detected column shapes

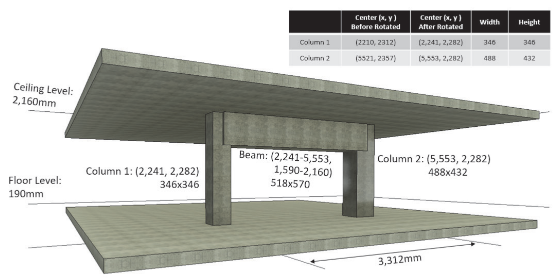

Figure 11: Excution to BIM

access to the information.

Conclusion and Future Work

This study proposed a novel framework for NeRF-to-BIM instead of Scan-to-BIM to bring potential benefits to the procedure of as-built modeling. The NeRF-to-BIM framework was categorized into three steps: (1) Reconstruction of the 3D scenes using NeRF, (2) 3D semantic segmentation, and (3) Conversion to BIM. The goal of this study was to establish a toy example in a simple structural model to evaluate the feasibility of NeRF-to-BIM.

For generating point clouds with NeRF, Instant-NeRF and Luma AI were used. While the Luma AI generated fewer artifacts and smoother output than Instant-NeRF, Luma AI produced a model with a different scale. Compared to the laser scanning method, capturing time and expensive equipment costs can be reduced.

In the semantic segmentation step, fine-tuning of the PointNeXt algorithm was executed with a synthesized point cloud dataset generated from the 3D model. The performance increased significantly compared to the case where only pre-trained weights were applied. This result indicates that more dataset related to building and synthetic point cloud generated from BIM is the key to achieving higher performance. Comparing Instant-NeRF and Luma AI performance, Luma AI achieved higher performance than Instant-NeRF.

In the conversion step, a four-step process is proposed: (1) Detect and set up the floor and ceiling levels, (2) Set up the grid using the detected center point of columns, (3) Rotate the entire point clouds along the XY grid and identify the beam size, (4) Reconstruct the BIM elements. The level & grid system, as well as the detected structural elements, were translated to the BIM model. It should be noted that these conversion steps are specifically designed for building objects that feature columns and beams positioned along the grid lines, which intersect at a right angle. The NeRF-to-BIM process still involves manual intervention between three steps. To address the limitations and challenges of the Scan-to-BIM method, an automated end-to-end NeRF-to-BIM workflow is necessary. Therefore, we identify this as an area for future work to advance our study and to realize this approach's potential benefits.

Acknowledgements

I would like to express my very great appreciation for Cesar Portocarrero in the Department of Civil and Environmental Engineering at Stanford University.

References

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

Bassier, M., Yousefzadeh, M. & Vergauwen, M. (2020), 'Comparison of 2d and 3d wall reconstruction algorithms from point cloud data for as-built bim', *Journal of Information Technology in Construction* **25**, 173–192.

Bi, S., Yuan, C., Liu, C., Cheng, J., Wang, W. & Cai, Y. (2021), 'A survey of low-cost 3d laser scanning technology', *Applied Sciences* **11**(9).
URL: *https://www.mdpi.com/2076-3417/11/9/3938*

Cao, Y. & Scaioni, M. (2022), 'A pre-training method for 3d building point cloud semantic segmentation', *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2**, 219–226.

Chen, J., Kira, Z. & Cho, Y. K. (2019), 'Deep learning approach to point cloud scene understanding for automated scan to 3d reconstruction', *Journal of Computing in Civil Engineering* **33**(4), 04019027.

Condorelli, F., Rinaudo, F., Salvadore, F. & Tagliaventi, S. (2021), 'A comparison between 3d reconstruction using nerf neural networks and mvs algorithms on cultural heritage images', *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **XLIII-B2-2021**, 565–570.

Corydon, T. J., Kopp, S., Wehland, M., Braun, M., Schütte, A., Mayer, T., Hülsing, T., Oltmann, H., Schmitz, B., Hemmersbach, R. et al. (2016), 'Alterations of the cytoskeleton in human cells in space proved by life-cell imaging', *Scientific reports* **6**(1), 1–14.

Croce, V., Caroti, G., De Luca, L., Jacquot, K., Piemonte, A. & Véron, P. (2021), 'From the semantic point cloud to heritage-building information modeling: A semiautomatic approach exploiting machine learning', *Remote Sensing* **13**(3).
URL: *https://www.mdpi.com/2072-4292/13/3/461*

De Geyter, S., Vermandere, J., De Winter, H., Bassier, M. & Vergauwen, M. (2022), 'Point cloud validation: On the impact of laser scanning technologies on the semantic segmentation for bim modeling and evaluation', *Remote Sensing* **14**(3).
URL: *https://www.mdpi.com/2072-4292/14/3/582*

Fedorova, S., Tono, A., Nigam, M. S., Zhang, J., Ahmadnia, A., Bolognesi, C. & Michels, D. L. (2021), 'Synthetic 3d data generation pipeline for geometric deep learning in architecture', *arXiv preprint arXiv:2104.12564* .

Fobiri, G. (2021), A scientometric review of reality capture applications in the built environment.

Fobiri, G., Musonda, I. & Muleya, F. (2022), 'Reality capture in construction project management: A review of opportunities and challenges', *Buildings* **12**(9), 1381.

Gao, B., Pan, Y., Li, C., Geng, S. & Zhao, H. (2022), 'Are we hungry for 3d lidar data for semantic segmentation? a survey of datasets and methods', *IEEE Transactions on Intelligent Transportation Systems* **23**(7), 6063–6081.

Gao, K., Gao, Y., He, H., Lu, D., Xu, L. & Li, J. (2022), 'Nerf: Neural radiance field in 3d vision, a comprehensive review'.

Golparvar-Fard, M., Bohn, J., Teizer, J., Savarese, S. & Peña-Mora, F. (2011), 'Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques', *Automation in Construction* **20**(8), 1143–1155.
URL: *https://www.sciencedirect.com/science/article/pii/S0926580511000707*

Golparvar-Fard, M., Peña-Mora, F. & Savarese, S. (2011), 'Integrated sequential as-built and as-planned representation with d 4 ar tools in support of decision-making tasks in the aec/fm industry', *Journal of Construction Engineering and Management* **137**(12), 1099–1116.

Grilli, E., Menna, F. & Remondino, F. (2017), 'A review of point clouds segmentation and classification algorithms', *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **42**, 339.

Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. (2017), 'Higher-order interactions stabilize dynamics in competitive network models', *Nature* **548**(7666), 210–213.

Hackel, T., Wegner, J. D. & Schindler, K. (2016), 'Fast semantic segmentation of 3d point clouds with strongly varying density', *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences* **3**, 177–184.

Jung, J., Hong, S., Yoon, S., Kim, J. & Heo, J. (2016), 'Automated 3d wireframe modeling of indoor structures from point clouds using constrained least-squares adjustment for as-built bim', *Journal of Computing in Civil Engineering* **30**(4), 04015074.

Liu, H., Zhang, J., Zhu, J. & Hoi, S. C. (2017), Deepfacade: A deep learning approach to facade parsing, IJCAI.

Ma, J. W., Czerniawski, T. & Leite, F. (2020), 'Semantic segmentation of point clouds of building interiors with deep learning: Augmenting training datasets with synthetic bim-based point clouds', *Automation in Construction* **113**, 103144.
URL: *https://www.sciencedirect.com/science/article/pii/S0926580519311884*

Matthews, J., Love, P. E., Heinemann, S., Chandler, R., Rumsey, C. & Olatunj, O. (2015), 'Real time progress management: Re-engineering processes for cloud-based bim in construction', *Automation in Construction* **58**, 38–47.
URL: *https://www.sciencedirect.com/science/article/pii/S0926580515001478*

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R. & Ng, R. (2020), 'Nerf: Representing scenes as neural radiance fields for view synthesis', *CoRR* **abs/2003.08934**.
URL: *https://arxiv.org/abs/2003.08934*

Moon, D., Chung, S., Kwon, S., Seo, J. & Shin, J. (2019), 'Comparison and utilization of point cloud generated from photogrammetry and laser scanning: 3d world model for smart heavy equipment planning', *Automation in Construction* **98**, 322–331.
URL: *https://www.sciencedirect.com/science/article/pii/S0926580517305125*

Müller, T., Evans, A., Schied, C. & Keller, A. (2022), 'Instant neural graphics primitives with a multiresolution hash encoding', *arXiv preprint arXiv:2201.05989* .

Murtiyoso, A., Pellis, E., Grussenmeyer, P., Landes, T. & Masiero, A. (2022), 'Towards semantic photogrammetry: Generating semantically rich point clouds from architectural close-range photogrammetry', *Sensors* **22**(3), 966.

Niemeyer, J., Rottensteiner, F., Sörgel, U. & Heipke, C. (2016), 'Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives 41 (2016)* **41**, 655–662.

Pei, S., Yang, R., Liu, Y., Xu, W. & Zhang, G. (2022), 'Research on 3d reconstruction technology of large-scale substation equipment based on nerf', *IET Science, Measurement & Technology* **n/a**(n/a).
URL: *https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/smt2.12131*

Perez-Perez, Y., Golparvar-Fard, M. & El-Rayes, K. (2021), 'Scan2bim-net: deep learning method for segmentation of point clouds for scan-to-bim', *Journal of Construction Engineering and Management* **147**(9), 04021107.

Qi, C. R., Su, H., Mo, K. & Guibas, L. J. (2017), Pointnet: Deep learning on point sets for 3d classification and segmentation, *in* 'Proceedings of the IEEE conference on computer vision and pattern recognition', pp. 652–660.

Qi, C. R., Yi, L., Su, H. & Guibas, L. J. (2017), 'Pointnet++: Deep hierarchical feature learning on point sets in a metric space', *Advances in neural information processing systems* **30**.

Qian, G., Li, Y., Peng, H., Mai, J., Hammoud, H. A. A. K., Elhoseiny, M. & Ghanem, B. (2022), 'Pointnext: Revisiting pointnet++ with improved training and scaling strategies', *arXiv preprint arXiv:2206.04670* .

Qu, T., Coco, J., Rönnäng, M. & Sun, W. (2014), Challenges and trends of implementation of 3d point cloud technologies in building information modeling (bim): case studies, *in* 'Computing in Civil and Building Engineering (2014)', pp. 809–816.

Qu, T. & Sun, W. (2015), 'Usage of 3d point cloud data in bim (building information modelling): Current applications and challenges', *Journal of Civil Engineering and Architecture* **9**(11), 1269–1278.

Ramakrishnan, S. K., Gokaslan, A., Wijmans, E., Maksymets, O., Clegg, A., Turner, J., Undersander, E., Galuba, W., Westbury, A., Chang, A. X., Savva, M., Zhao, Y. & Batra, D. (2021), 'Habitat-matterport 3d dataset (HM3D): 1000 large-scale 3d environments for embodied AI', *CoRR* **abs/2109.08238**.
URL: *https://arxiv.org/abs/2109.08238*

Reiser, C., Peng, S., Liao, Y. & Geiger, A. (2021), Kilonerf: Speeding up neural radiance fields with thousands of tiny mlps, *in* 'Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)', pp. 14335–14345.

Straub, J., Whelan, T., Ma, L., Chen, Y., Wijmans, E., Green, S., Engel, J., Mur-Artal, R., Ren, C., Verma, S., Clarkson, A., Yan, M., Budge, B., Yan, Y., Pan, X., Yon, J., Zou, Y., Leon, K., Carter, N. & Newcombe, R. (2019), 'The replica dataset: A digital replica of indoor spaces'.

Tserng, H.-P., Ho, S.-P. & Jan, S.-H. (2014), 'Developing bim-assisted as-built schedule management system for general contractors', *Journal of Civil Engineering and Management* **20**(1), 47–58.

Wang, T. & Xiong, Z. (2021), 'Methods of as-is bim reconstruction using point cloud data for existing buildings', *IOP Conference Series: Earth and Environmental Science* **676**(1), 012041.
URL: *https://dx.doi.org/10.1088/1755-1315/676/1/012041*

Weinmann, M., Schmidt, A., Mallet, C., Hinz, S., Rottensteiner, F. & Jutzi, B. (2015), 'Contextual classification of point cloud data by exploiting individual 3d neigbourhoods', *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences II-3 (2015), Nr. W4* **2**(W4), 271–278.

Xu, Q., Sun, X., Wu, C.-Y., Wang, P. & Neumann, U. (2020), Grid-gcn for fast and scalable point cloud learning, *in* 'Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition', pp. 5661–5670.