Object-level Temporal Change Detection on Construction Sites with 3D Deep Learning Models

Jeffrey Rugg[1], Jingdao Chen[1], Mikias Gugssa[2], and Jun Wang[2]

[1]Computer Science and Engineering, Mississippi State University, Mississippi State, United States, Email: jhr250@msstate.edu, chenjingdao@cse.msstate.edu
[2]Civil and Environmental Engineering, Mississippi State University, Mississippi State, United States, Email: mgw285@msstate.edu, jwang@cee.msstate.edu

ABSTRACT

In the evolving landscape of construction progress monitoring, precise and automated change detection at the object level remains a pivotal challenge. This paper introduces a novel methodology that leverages machine learning for object-level temporal change detection in construction environments, focusing on semantic segmentation of 3D point clouds and highlighting changed objects across construction stages. Our approach builds upon the foundations of 3D laser scanning and Building Information Modeling (BIM), integrating advanced machine learning algorithms to interpret complex 3D datasets. We specifically address the challenge of distinguishing and tracking temporal changes of individual objects, a notable gap in existing methodologies which had only focused on changes at the point level or voxel level. This capability is critical for accurately differentiating between permanent and temporary elements in construction projects, thus enabling timely and data-driven progress monitoring and decision-making. In this research, we use a pre-trained PointNet semantic segmentation model to label points in a 3D point cloud of a construction site. Then, clustering algorithms are used to extract objects from the point cloud and match them across different construction stages. This research makes use of the Nothing Stands Still (NSS) dataset, which consists of 3D point clouds collected from real construction sites across multiple months, covering long temporal durations and significant additions and removal of different construction-related elements. Experimental results show that the proposed algorithm is successful in automatically extracting changed building elements over multiple construction stages.

INTRODUCTION

The advent of 3D scanning technologies has significantly impacted the construction industry by enhancing the monitoring and management of construction projects. Central to this progress is the integration of 3D laser scanning with Building Information Modeling (BIM) (Chen et al. 2019; Zeng et al. 2020), which has been extensively explored in scholarly research to improve construction progress tracking and change detection.

Change detection, in the context of construction progress tracking, is defined as the process of analyzing 3D space and identifying installed or changed building elements over time to measure physical progress (Meyer et al. 2022). Current methods in the literature for change detection are based on conducting laser scans of the construction environment and identifying changes through the 3D laser-scanned point clouds. Existing research has demonstrated automated change detection methods mostly using scan-to-BIM registration (Braun et al. 2015; Bosché et al. 2015), point-to-point comparison (Chen and Cho 2018), or voxel-to-voxel comparison (Meyer et al. 2022). However, an integration of change detection with semantic understanding of objects on a construction site remains a significant research gap. With the advent and rapid growth of machine learning and deep learning methods for automated recognition of construction entities (Chen et al. 2019; Chen et al. 2017), an integration of automated object detection with automated change detection could offer much value to construction progress monitoring.

This paper proposes a change detection algorithm based on the semantic segmentation of a point cloud using deep neural networks. The main contribution of this paper is a change detection algorithm that: (i) works at the object-level instead of a point-level and (ii) works without relying on existing as-planned BIM models. The following sections will describe the literature review, methodology, results, and discussion respectively.

LITERATURE REVIEW

In the area of change detection from point cloud data, Turkan et al. (2012) were among the early researchers in this domain, developing an innovative system that merged 3D object recognition with 4D scheduling information. Their system was capable of autonomously updating construction schedules by analyzing actual site conditions captured through 3D laser scanning. This integration promised increased efficiency and the potential for cost savings by enabling dynamic and responsive scheduling.

Bosché et al. (2015) introduced an approach that integrated Scan-to-BIM and Scan-vs-BIM systems, with a particular focus on cylindrical MEP components. Their method automated the comparison between as-built conditions and as-planned models, utilizing the Hough transform to identify deviations in MEP installations. This demonstrated the capability of such systems to recognize both the displacement and the completeness of MEP components, further streamlining the process of construction monitoring.

In the same vein, Chen and Cho (2018) presented a point-to-point comparison method for Scan-vs-BIM that was independent of the geometry of analyzed elements, allowing for broad application across various building components. By employing the RANSAC algorithm for alignment, their method proved efficient in deviation detection, indicating a reduction in the need for manual inspections and an advancement in automated construction monitoring.

Meyer, Brunn, and Stilla (2022) contributed to this body of literature by focusing on high-resolution change detection within indoor construction environments, utilizing dense Terrestrial Laser Scanning (TLS) point clouds. They applied the Dempster-Shafer theory to model uncertainties and assess metric accuracy, demonstrating their method's efficacy in documenting construction progress and verifying BIM compliance according to specified Levels of Accuracy. This approach highlighted the critical role of high-quality scanning data in effective change detection processes.

A comprehensive review by Stilla and Xu (2023) on the use of 3D point clouds for change detection in urban environments further expanded on these methodologies. They highlighted the superiority of 3D data over 2D imaging for conducting detailed geometric and attribute analyses in a range of urban applications, from land use monitoring to infrastructure supervision. Their work identified current technological limitations and emphasized the need for continued research to fulfill the demand for automated change detection in urban settings.

Collectively, these studies underscore the ongoing transformation in construction project management through the integration of advanced technological solutions. Despite that, research gaps remain in associating semantic information about a point cloud scene with detected geometrical changes. This paper aims to address these gaps by utilizing deep neural networks to extract semantic labels from point cloud data and using that to perform object-level change detection.

METHODOLOGY

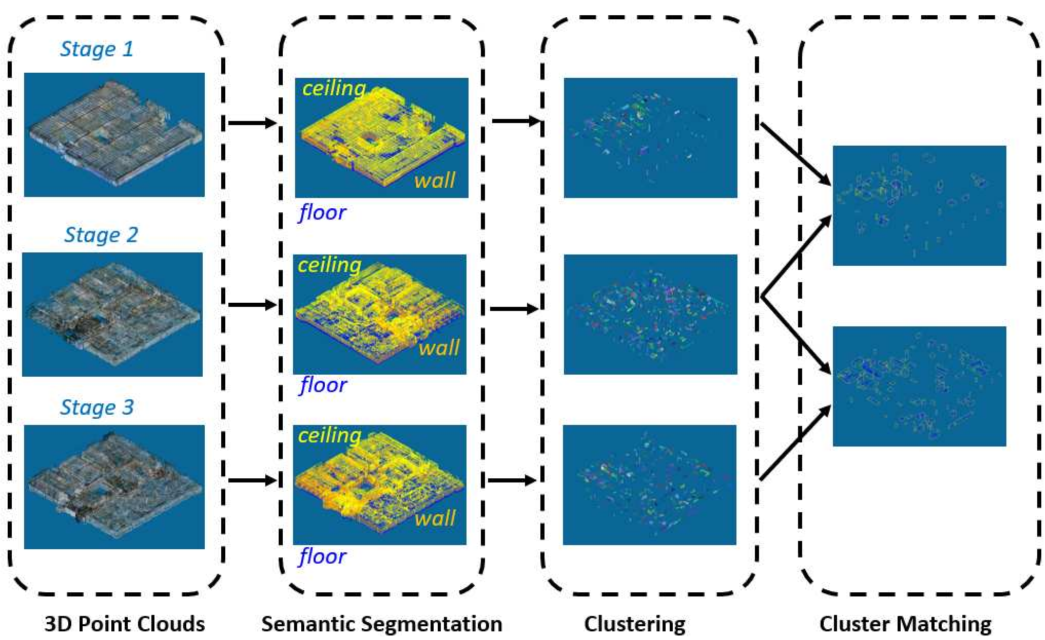


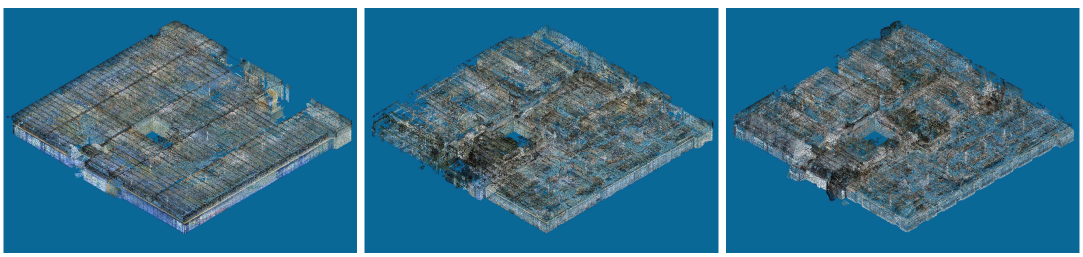
Fig. 1. Proposed pipeline for point cloud semantic segmentation, clustering and cluster matching for object-level change detection

Figure 1 shows the overall pipeline of the proposed methodology for object-level change detection. First, the point cloud is segmented at the point-level into classes such as ceiling, wall, floor, door etc. using a pre-trained neural network using the PointNet architecture. Next, Euclidean clustering is applied to organize the points into objects corresponding to individual building elements. Then, cluster matching is used to compare clusters between different construction stages and determine newly constructed entities. The following subsections will go into each step of the pipeline in more detail.

Point cloud data preprocessing

For our research, we utilized Building 1 from the NSS dataset (Sun et al. 2023). The dataset was collected using the Matterport Camera v1, which is a static tripod-based reality capture system that acquires 360^o fragments from multiple scan locations. This building is approximately 3600 sq meters and represented by three point clouds taken at three different stages of construction. The point clouds consist of between 6.1 to 6.7 million points for each stage and are shown in Figure 2. The point clouds between different construction stages were calibrated and pre-aligned in the NSS dataset using a combination of manual alignment and Iterative Closest Point (ICP).

After acquiring the dataset, each point cloud was preprocessed in Cloud Compare to produce a color PLY representation. The calibration matrices for each stage was used to apply a transformation matrix to the point cloud to ensure that the point clouds from all stages are aligned in the same global coordinate system. We then used the Open3D and Numpy libraries to read the PLY file into arrays to prepare them for processing.



(a) Building 1 - Stage 1 **(b)** Building 1 - Stage 2 **(c)** Building 1 - Stage 3

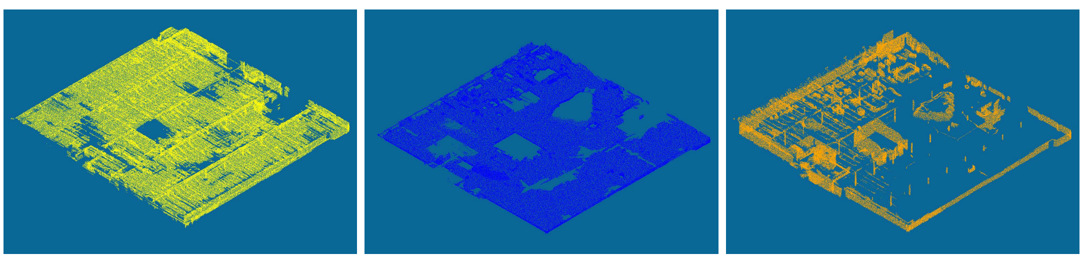
Fig. 2. 3D point clouds from three different construction stages of the same building site from the NSS dataset

Semantic segmentation and clustering

In order to ensure efficient processing by neural networks, the point cloud data for the entire building was broken up into 5m x 5m grids. The grid points were then normalized with respect to the grid coordinate system and processed by the PointNet model (Charles et al. 2017) for semantic segmentation. The PointNet model is pre-trained on the Stanford 3D Indoor Spaces (S3DIS) dataset (Armeni et al. 2016). This means that the PointNet segmentation model is trained to look for 13 classes of objects in indoor environments, though not necessarily for construction environments. For the NSS point clouds utilized in our research, the only segmentation categories that were sufficiently detected and classified were the ceiling, floor, and wall classes. The complete segmentation results for all classes and building stages are shown in Table 1. For illustration purposes, the individual point cloud segments are shown for Stage 1 in Figure 3. These results are typical for all the stages of construction for the building used in our research.

Table 1. Semantic segmentation results showing the number of points detected in each class for different construction stages

Class	Stage 1	Stage 2	Stage 3
clutter	25	307	16
board	0	0	0
bookcase	3	1	2
beam	0	0	0
chair	0	30	19
column	0	2	0
door	58	829	623
sofa	0	0	0
table	0	0	0
window	0	0	0
ceiling	456,379	391,793	353,588
floor	329,968	330,766	300,120
wall	209,722	276,714	343,428



(a) Ceiling Class　**(b)** Floor Class　**(c)** Wall Class

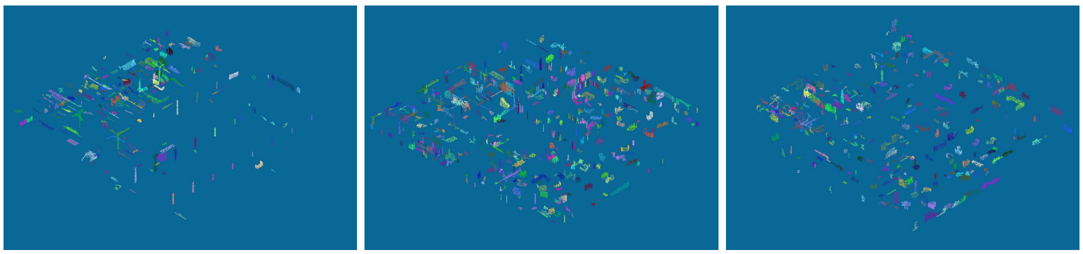
Fig. 3. PointNet semantic segmentation results for 3 of the most prominent classes in Building 1 - Stage 1

Once the semantic segmentation process was complete, the grid points were reassembled into a multidimensional array representing the entire building. This array was then processed by the Open3D DBSCAN function to perform Euclidean clustering. The clustering process aims to group adjoining points that have the same semantic labels and are within a close enough distance (0.15m) into distinct clusters. To eliminate very small clusters from our results, any clusters with a size less than 50 points were removed from the data set. Initial experimentation revealed the presence of several large clusters caused by undersegmentation that complicated the change detection analysis. To overcome this, an upper limit of 1000 points was added to the cluster filter.

As illustrated in Figure 4, the visualization of the clusters revealed that the class of wall objects offered the largest and best represented points. Therefore, the wall clusters

(a) Ceiling clusters **(b)** Floor clusters **(c)** Wall clusters

Fig. 4. Euclidean clustering results for each class of objects detected for Building 1 - Stage 1. Points are colored according to the cluster that they belong to.



(a) Stage 1 wall clusters **(b)** Stage 2 wall clusters **(c)** Stage 3 wall clusters

Fig. 5. Wall cluster results for Building 1 - All Stages. Points are colored according to the cluster that they belong to.

were utilized as the basis for comparing changes between the stages of construction. Figure 5 shows the results of Euclidean clustering for all three stages of construction.

Object-level change detection with bounding box matching

In order to detect which clusters changed from one stage to another, the coordinates for each cluster in the later stage were processed by an overlapping algorithm and compared to all the clusters found in the previous stage. If the bounding box for a cluster in the previous stage overlapped a cluster in the same semantic class at a later stage by more than 10 percent, then it was considered to be existing. Otherwise, the cluster was categorized as a new cluster for the later stage. To further analyze the changes, the comparisons were reversed to determine how many clusters existed in a previous stage but not in the next.

RESULTS

Evaluation metrics

Comparison of segmentation methods and effect on change detection

The results of the object-level change detection algorithm are shown in Figures 6 and 7. When comparing Stage 2 to Stage 1, 75 clusters were classified as existing in both stages, while 403 clusters were identified as new in Stage 2. When comparing Stage 3 to Stage 2, 127 clusters were classified as existing in both stages, while 195 clusters were identified as new in Stage 3. These results indicate that there were more changes between the first two stages, which is consistent with what is shown in the 3D point clouds presented in Figure 2.

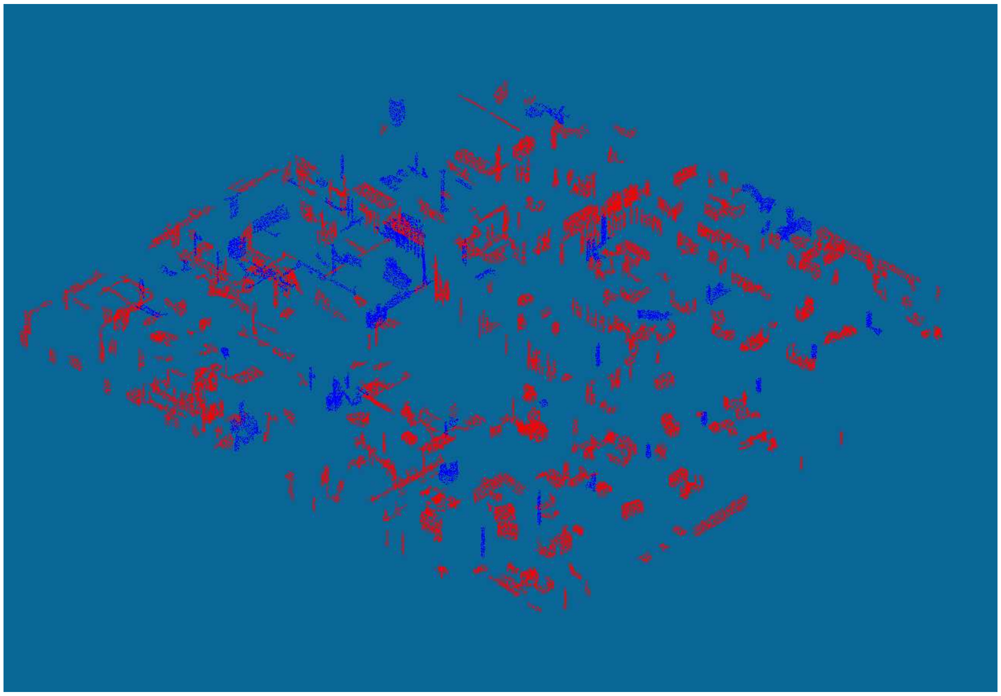


Fig. 6. Change detection results between Stage 1 and Stage 2. Blue clusters represent objects that existed in both stages. Red clusters represent new objects detected in Stage 2.

As for the reverse comparisons, the algorithm revealed that 147 clusters that existed in Stage 1 were not detected in Stage 2, and 346 clusters that existed in Stage 2 were not detected in Stage 3. Although the detection of missing clusters in later stages was not the primary focus of our research, such results could provide some useful information regarding the removal of temporary structures, building materials, and construction equipment as a project nears its completion.

Discussion

The main advantage of the proposed method is the ability to quickly perform semantic segmentation and change detection at the object level even with large buildings and construction sites. In addition, the computational resources required for this method are reasonable, since the PointNet model can be trained within 1 hour and the inference time is within a few seconds. Open standards, file formats (e.g. PLY), and open-source libraries (e.g. NumPy, PyTorch) are utilized in this method, making it easier to understand and optimize for specific applications. Furthermore, the accuracy of this approach will continue to show significant improvements as better deep neural networks are developed and utilized.

The limitation of the proposed method is that a better model may be needed to more accurately segment the point cloud data. Also, using a model trained with classes that are

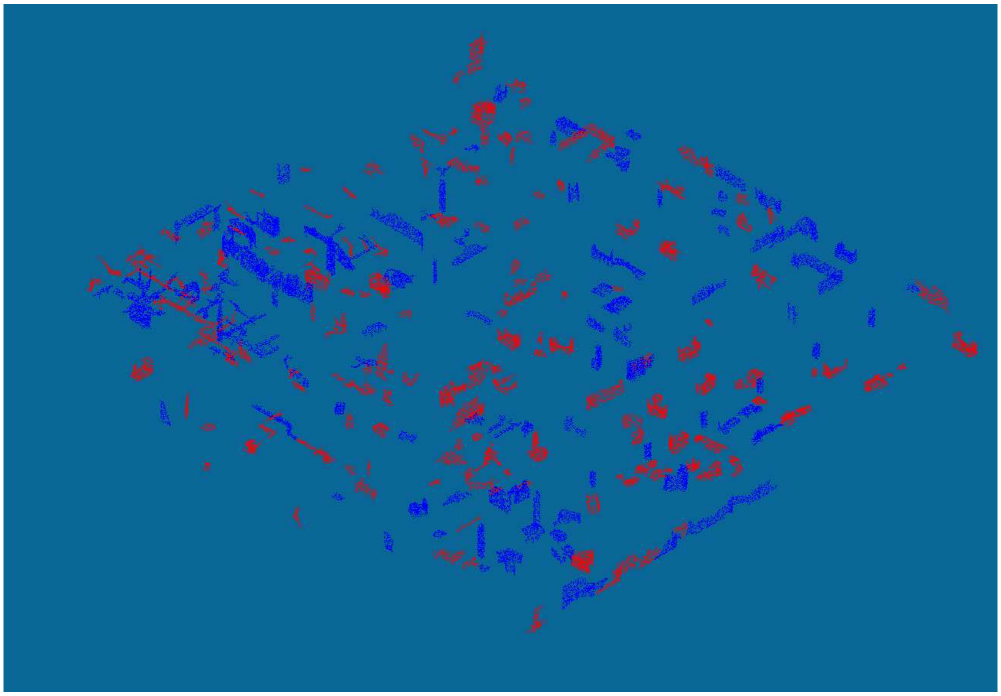


Fig. 7. Change detection results between Stage 2 and Stage 3. Blue clusters represent objects that existed in both stages. Red clusters represent new objects detected in Stage 3.

more relevant to the construction process may be better suited for a commercial building jobsite. There is also a need to account for possible variances in the scanning process itself that was not directly addressed in this study. This is because some degree of new or missing clusters could result based on environmental factors (occlusion, missing data etc.) even if no actual changes took place. This study only considers a single building in the NSS dataset and in the future, a broader study may be needed to fully validate the method. Also, more tuning may be needed to determine the best hyperparameters for each process (e.g. clustering thresholding and bounding box overlap threshold).

CONCLUSION

In conclusion, this paper demonstrated a pipeline for object-level construction change detection using deep-learning based semantic segmentation, euclidean clustering, and bounding box matching. Although the semantic segmentation and clustering processes need some refinement to provide more accurate results, our research shows the value of this approach and its potential as an aid in construction management. Overall, our work addresses crucial challenges identified in prior research, including the need for heightened automation and accuracy in object-level change detection, while maintaining efficient processing of large-scale data. Our research paves the way for future developments in the industry by applying machine learning techniques to construction monitoring, contributing significantly to the field's advancement. In future work, we will investigate more advanced neural networks such as Point Transformers that are capable of finely distinguishing complex construction elements such as temporary structures, construction equipment, and Mechanical, Electrical, and Plumbing (MEP) components. In addition, we plan to perform quantitative evaluation of the change detection accuracy by comparing the algorithm's output to manually annotated ground truth data.

Acknowledgements

The work reported herein was supported by the National Science Foundation (NSF) (Award #IIS-2153101). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., and Savarese, S. (2016). "3d semantic parsing of large-scale indoor spaces." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1534–1543.

Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., and Haas, R. (2015). "The value of integrating scan-to-bim and scan-vs-bim techniques for construction monitoring using laser scanning and bim: The case of cylindrical mep components." *Automation in Construction*, 49, 201–213 30th ISARC Special Issue.

Braun, A., Tuttas, S., Borrmann, A., and Stilla, U. (2015). "A concept for automated construction progress monitoring using bim-based geometric constraints and photogrammetric point clouds." *ITcon Vol. 20, Special issue ECPPM 2014 - 10th European Conference on Product and Process Modelling*.

Charles, R. Q., Su, H., Kaichun, M., and Guibas, L. J. (2017). "Pointnet: Deep learning on point sets for 3d classification and segmentation." *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 77–85.

Chen, J. and Cho, Y. (2018). "Point-to-point comparison method for automated scan-vs-bim deviation detection." *17th International Conference on Computing in Civil and Building Engineering* (June).

Chen, J., Fang, Y., Cho, Y. K., and Kim, C. (2017). "Principal axes descriptor for automated construction-equipment classification from point clouds." *Journal of Computing in Civil Engineering*, 31(2), 04016058.

Chen, J., Kira, Z., and 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.

Meyer, T., Brunn, A., and Stilla, U. (2022). "Change detection for indoor construction progress monitoring based on bim, point clouds and uncertainties." *Automation in Construction*, 141, 104442.

Stilla, U. and Xu, Y. (2023). "Change detection of urban objects using 3d point clouds: A review." *ISPRS Journal of Photogrammetry and Remote Sensing*, 197, 228–255.

Sun, T., Hao, Y., Huang, S., Savarese, S., Schindler, K., Pollefeys, M., and Armeni, I. (2023). "Nothing stands still: A spatiotemporal benchmark on 3d point cloud registration under large geometric and temporal change." *arXiv preprint arXiv:2311.09346*.

Turkan, Y., Bosche, F., Haas, C. T., and Haas, R. (2012). "Automated progress tracking using 4d schedule and 3d sensing technologies." *Automation in Construction*, 22, 414–421 Planning Future Cities-Selected papers from the 2010 eCAADe Conference.

Zeng, S., Chen, J., and Cho, Y. K. (2020). "User exemplar-based building element retrieval from raw point clouds using deep point-level features." *Automation in Construction*, 114, 103159.