EPiC Series in Built Environment
Volume 7, 2026, Pages 803–812
Proceedings of Associated Schools of Construction 62nd Annual International Conference

Deep Learning Driven Framework for Construction Point-Cloud Processing

Caochenyu Zhou[1], Imran Jebarah[1], Junshan Liu[1]
[1]Auburn University

Point-cloud data have become central to digital construction workflows through technologies such as laser scanning and photogrammetry. However, current processing methods remain fragmented, time-consuming, and heavily dependent on expert supervision. This study aims to clarify how construction workflows handle point-cloud processing, compare traditional and deep-learning-based approaches for key processing tasks, and propose a more cohesive workflow. A focused literature-based data collection process (2023–2025) identifies common processing tasks, denoising, sampling, registration, semantic segmentation, and completion, along with representative deep-learning applications and their advantages. The study summarizes these findings in an evidence table that links processing tasks, construction use cases, and task-level benefits of deep learning. Building on this analysis, it presents a conceptual framework that connects preprocessing and semantic inference into an integrated deep-learning-driven pipeline for construction point-cloud processing. The paper concludes by outlining research priorities for task-specific benchmarking, standardized datasets, and integration with construction information systems to enable reproducible and scalable automation.

Keywords: point cloud, deep learning, construction automation, point cloud processing

Introduction

The digital transformation of the construction industry has accelerated the adoption of technologies such as Building Information Modeling (BIM), reality capture, and artificial intelligence (AI). Among these, point-cloud data, captured through terrestrial laser scanning (TLS) or photogrammetry, has become essential for as-built documentation, progress tracking, and digital twin development. By converting physical environments into precise 3D representations, point clouds enable improved automation, accuracy assessment, and data integration across construction workflows.

Despite their growing importance, point-cloud processing remains largely manual and fragmented. Tasks such as denoising, registration, segmentation, and reconstruction rely on multiple software platforms and expert supervision. This dependence increases processing time, introduces subjectivity, and limits scalability (Ye et al., 2025; Prieto, 2024). As data volumes expand, manual workflows struggle to maintain efficiency and consistency, underscoring the need for automated, learning-based solutions capable of handling large 3D datasets reliably. Deep Learning (DL) offers such potential by enabling neural networks to learn geometric and contextual features directly from point-cloud data. However, current applications in construction are scattered across isolated tasks, typically segmentation, registration, or detection, without an integrated framework connecting these processes into a coherent pipeline. This fragmentation defines a key research gap. The construction research community lacks a unified understanding of how DL can be systematically applied across all stages of point-cloud processing, from data acquisition and denoising to reconstruction and model integration. Similarly, consistent mappings between tasks, methods, datasets, and evaluation metrics remain underdeveloped, limiting reproducibility and large-scale adoption within construction practice.

To address these issues, this study proposes a deep learning–driven framework for automating point-cloud processing in construction. The framework identifies which stages of the pipeline are most suitable for deep learning, highlights their benefits, and examines how these methods improve accuracy and efficiency compared with manual approaches.

The research is guided by three questions:

- RQ1: In construction practice, what point-cloud processing tasks are performed?
- RQ2: For the same processing tasks, what advantages do deep-learning methods offer over traditional approaches?
- RQ3: How can we strengthen the overall cohesion of point-cloud processing, linking currently independent steps into an integrated pipeline with consistent representations and quality control?

By addressing these questions, the paper contributes a conceptual framework that synthesizes recent advances in 3D DL and aligns them with construction workflows. The remainder of this paper reviews relevant literature, describes the research methodology and proposed framework, compares manual and automated approaches, and concludes with key findings and directions for future research.

Background and Related Work

Manual and Classical Point-Cloud Processing in Construction

Manual and rule-based point-cloud processing in construction typically follows a sequential workflow involving denoising, sampling, registration, segmentation, and reconstruction (Ye et al., 2025). Denoising is used to remove stray points caused by dust, reflections, moving equipment, or scanner artefacts so that structural geometry can be analyzed reliably. Sampling or thinning is often applied to reduce file size and make downstream processing feasible on large construction scenes.

Registration aligns multiple scans into a common coordinate system using target matching, feature-based alignment, or iterative closest-point (ICP) algorithms, so that a unified view of the site can be obtained. Segmentation then partitions the registered point-cloud into regions or objects based on geometric rules such as region growing, Random Sample Consensus (RANSAC) plane extraction plane fitting, or curvature clustering, enabling identification of walls, slabs, columns, or equipment. Finally, reconstruction converts segmented point clouds into meshes or parametric BIM elements through surface fitting, triangulation, or primitive extraction to support documentation and modeling tasks. (Prieto, 2024). These classical methods are implemented across separate software tools (e.g., CloudCompare, Autodesk Recap, and Leica Cyclone) and require manual parameter tuning and repeated visual checks. As a result, processing large-scale or complex construction scenes remains time-consuming, subjective, and difficult to reproduce, motivating research into learning-based automation.

Deep Learning Techniques for Point-Cloud Processing

In this study, tasks refer to the main computational processes in point-cloud workflows, such as denoising, registration, segmentation, and reconstruction. Deep learning has advanced automation in these processes by enabling computers to learn geometric and spatial relationships directly from raw 3D data (Chen et al., 2023). To address the second research question, Table 1 compares traditional and deep-learning approaches across major construction point-cloud processing tasks. The table synthesizes findings reported across recent construction-focused studies and is intended as a comparative summary rather than original experimental results. The comparison highlights how DL improves task performance, reduces manual effort, and enhances scalability relative to classical algorithms. Instead of relying on manually set parameters, DL models analyze thousands of examples to identify shapes, align multiple scans, and fill missing information automatically (Wang et al., 2019).

In construction contexts, DL has been used to remove scanner noise, align as-built scans with BIM models, classify structural and MEP components, and reconstruct incomplete or occluded elements (Qi et al., 2017). These applications significantly reduce manual editing and improve the consistency of as-built modeling and progress tracking (Zhao et al., 2021). Recent neural network architectures for 3D data can directly process unordered point sets, capturing both local geometry and large-scale spatial context. Such models have achieved strong results in tasks like segmentation, object detection, and reconstruction, demonstrating that DL can generalize across complex construction environments without handcrafted rules (Park et al., 2022).

Table 1. Comparison of traditional and deep-learning approaches for construction point-cloud processing tasks (synthesized from Chen et al., 2022; Park, 2021; Mehranfar et al., 2024)

Task	Traditional Approach	Deep-Learning Approach	Key Advantages (DL)
Sampling	Uniform or random downsampling; voxel grids	Task-aware or attention-based sampling networks	Retains critical geometry; reduces redundancy; improves downstream accuracy
Denoising	Statistical/bilateral filtering; rule-based noise removal	Learned denoising networks (encoder–decoder, point-based)	Learns non-Gaussian noise; preserves thin edges and boundaries; less manual tuning
Registration	ICP; feature-based matching; target-based alignment	Deep feature–correspondence networks (e.g., Deep Closest Point, Predator)	More robust under clutter and low overlap; faster convergence; less initialization dependence
Segmentation	Region growing; geometric clustering; RANSAC plane extraction	Transformer or sparse 3D CNN segmentation models	Higher labeling accuracy; maintains geometry; reduces manual labeling effort
Completion	Poisson reconstruction; surface fitting; interpolation	Transformer-based completion networks (e.g., PoinTr)	Fills occlusions; preserves topology; enables mesh-ready geometry for Scan-to-BIM

These task-level comparisons clarify where DL offers practical advantages for construction workflows and provides the analytical basis for the integrated framework presented in the following section.

Applications of Deep Learning in Construction Point-Cloud Workflows

Between 2019 and 2025, research on DL for construction point clouds has expanded across multiple workflow stages. Scholars have explored denoising (Rakotosaona et al., 2020), adaptive sampling (Li et al., 2019), registration (Wang & Solomon, 2019), and semantic segmentation (Zhao et al., 2021), confirming that neural networks outperform traditional methods in accuracy and robustness. Recent studies have also addressed point-cloud completion, where models reconstruct missing geometry to overcome occlusions common on construction sites. Transformer-based approaches such as Yue et al. (2025) and Wang et al. (2025) achieve high accuracy in recovering slender or hidden MEP elements, enhancing data continuity for downstream modeling. Within construction, DL models have been applied to as-built BIM generation (Park, 2021), progress tracking (Kavaliauskas, 2022), and site safety assessment (Valdebenito & Forcael et al., 2024), though most remain limited to isolated tasks rather than integrated, end-to-end workflows.

A growing number of studies emphasize the need for consistent evaluation metrics to compare results across models and tasks. Common measures include mean Intersection over Union (mIoU) for segmentation accuracy, mean Average Precision (mAP) for object detection, and geometric error metrics such as Chamfer Distance or Root Mean Square Error (RMSE) for reconstruction quality (Chen et al., 2022; Wang et al., 2025). Despite methodological progress, inconsistencies in dataset types, annotation quality, and metric definitions make cross-study benchmarking difficult. Furthermore, existing research faces persistent domain-shift challenges. Models trained on clean, synthetic datasets such as ShapeNet or ScanNet often underperform on real construction scans characterized by clutter, occlusion, and complex materials. Recent reviews (Chen et al., 2022; Han & Lee, 2024) stress that developing construction-specific datasets and standardized evaluation protocols is critical for meaningful adoption of DL in industry practice.

Identified Research Gaps

The reviewed studies reveal two major deficiencies. First, research remains fragmented, targeting single-point tasks such as segmentation or detection while overlooking end-to-end integration across the full processing pipeline. Second, reproducibility and benchmarking challenges persist due to inconsistent metrics and limited availability of labeled, domain-specific datasets. These shortcomings hinder both academic comparison and industrial implementation. Addressing these challenges requires a comprehensive conceptual framework that aligns DL architectures with specific processing tasks, supported by standardized data and evaluation protocols. Such an approach can guide future automation research and support a unified methodology for AI-driven point-cloud workflows in construction.

Methodology

As shown in Figure 1, we first conducted a scoped literature review to define the construction context, search keywords, and time window, then queried major scholarly databases. Next, we screen titles/abstracts/full texts to remove duplicates and exclude studies not situated in construction. For the retained corpus, we perform manual data extraction with a structured sheet. We then synthesize findings by processing order and purpose, summarize common patterns, and based on the extracted evidence and DL characteristics, propose an integrated deep-learning framework for point-cloud processing, explicitly stating inputs and outputs, processing stages, dataset traits, and evaluation standards.

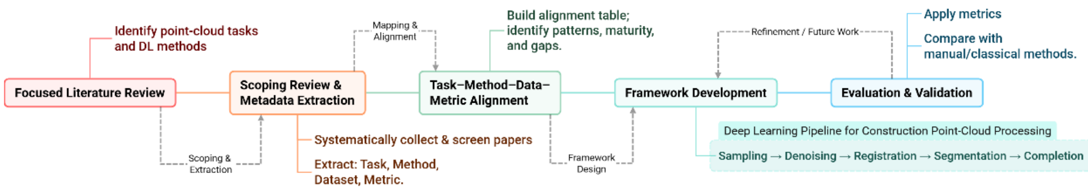

Figure 1. Overview of the research methodology and framework-development workflow. Developed by the authors.

Data Collection

The scoping review examined how DL has been applied to point-cloud processing in construction contexts. Searches were conducted in Scopus and Web of Science using Boolean queries that combined domain, data, and method terms for English-language articles published between 2023 and 2025. From the reviewed studies, we extracted relevant information and compiled a tabular summary, as shown in Table 2 detailing (i) the point-cloud processing performed, (ii) the purpose of employing DL for that processing (iii) the advantages of DL over manual or classical methods. The evidence summarized in Table 2 is extracted from peer-reviewed construction studies and reflects reported applications rather than proposed use cases.

Table 2. Evidence Table: Purpose and Advantages of Deep-Learning–Based Point-Cloud Processing for Construction Tasks, In the Processing Task column, the numeric codes correspond to the following operations: 1 = Sampling, 2 = Denoising, 3 = Registration, 4 = Semantic Segmentation, and 5 = Completion.

Construction Task	Processing Task	Purpose	Advantages
Earthwork Progress Monitoring	1, 2, 3, 4	Remove moving machinery from UAV-generated point clouds to calculate net excavation volumes	1. Drastically reduces manual effort and time 2. Maintains high accuracy in volume estimates
Scaffold Safety Inspection	1, 2, 3, 4	Detect deviations or modifications in scaffolding by comparing scan data to design specifications	1. Automates inspection, cutting time and labor compared to manual checks 2. Enhances site safety by quickly identifying potential hazards
Building Defect Inspection	1, 2, 3	Fuse multi-sensor (RGB-D) scans into a unified 3D model of building interiors. Automatically identify structural defects in the point cloud	1. Improves accuracy and consistency of defect detection over manual inspection 2. Enables early, objective identification of quality issues to improve construction outcomes
Bridge Geometry Inspection	1, 2, 3, 4, 5	Perform LiDAR–camera fusion SLAM to reconstruct large-scale bridge point clouds in real time.	1. Achieves much higher geometric accuracy than UAV photogrammetry 2. Significantly faster than traditional scanning
		Deep-learning	

		segmentation to extract bridge components for dimensional measurement	
Worker Safety Monitoring	1, 2, 3, 4	Detect and localize construction workers in LiDAR point clouds to avoid collisions and enforce safety	1. Higher detection precision and lower error than baseline models 2. Provides rich 3D location data of workers for real-time safety alerts.
Road Digital Twin Modeling	1, 2, 3, 4, 5	Segment road infrastructure in point clouds to build colored, 3D digital twin models	1. Very high segmentation accuracy on road assets 2. Fully automated pipeline makes twin generation scalable and cost-effective
Building Digital Twin (Connectivity)	1, 2, 3, 4, 5	Identify connectivity between building surfaces by building a surface-topology graph from the point cloud	1. Captures structural relationships and spatial context that are crucial for as-built models 2. Enables efficient automated construction of detailed building digital twins
Industrial Plant Digitization	1, 2, 3, 4	Segment complex plant components in point clouds to generate digital models for existing facilities	1. Using synthetic, simulation-based training data boosts segmentation accuracy 2. Overcomes lack of labeled data, enabling effective learning in data-scarce industrial settings
Construction Quality Control	1, 2, 3, 4	Use BIM-derived synthetic point clouds to train segmentation models that identify as-built building components in scan data	1. Synthetic data improves segmentation performance 2. Automates inspection of element dimensions and installation, reducing manual QA effort.

Synthesizing the extracted evidence on processing, purpose, and DL advantages, we distilled design requirements for an integrated framework that unifies preprocessing with semantic inference and modeling.

Proposed Framework

Our framework accepts raw scan point clouds as input irrespective of different formats and organizes processing into two modules: preprocessing and semantic inference and modeling, as illustrated in Figure 2, each task implemented using deep learning. Preprocessing comprises sampling, denoising, and registration. The objective is to increase signal-to-noise ratio, spatial uniformity, and global coherence of the raw data so that features are more learnable for subsequent stages: the sampling module normalizes density while preserving structure-critical geometry and reducing computational load; the denoising module suppresses sensor noise and outliers while retaining edges and thin elements; the registration module aligns multi-station scans into a common coordinate frame, reduces drift, and consolidates complementary views. Semantic inference and modeling include semantic segmentation and completion. Segmentation produces per-point and instance labels that enhance interpretability and enable element-level modeling, QA/QC, progress analysis, and clash reasoning. Completion operates on the segmented instances or regions to reconstruct occluded or inaccessible surfaces and to restore topology-consistent, part-aware geometry at a resolution compatible with the observed data; completed instances are fused back into the registered scene. The final output is a semantically labeled, globally registered, noise-reduced, density-normalized, and geometrically completed point cloud with consistent conventions suitable for automated 3D modeling and downstream construction analytics.

The framework is modular and method-agnostic. Each stage exposes a standard interface that accepts and emits point clouds, allowing components to be replaced with state-of-the-art methods selected by task–method fit, objective, data characteristics, robustness evidence, and computational budget. Because every processing task uses and produces point-cloud data, the pipeline avoids format conversions, preserves geometric fidelity, simplifies orchestration, enables consistent quality control and auditing, and supports fair benchmarking and end-to-end retraining when modules improve. In the next section, we instantiate the framework with representative examples and report task-appropriate metrics to demonstrate how module choices trade accuracy, efficiency, and reproducibility.

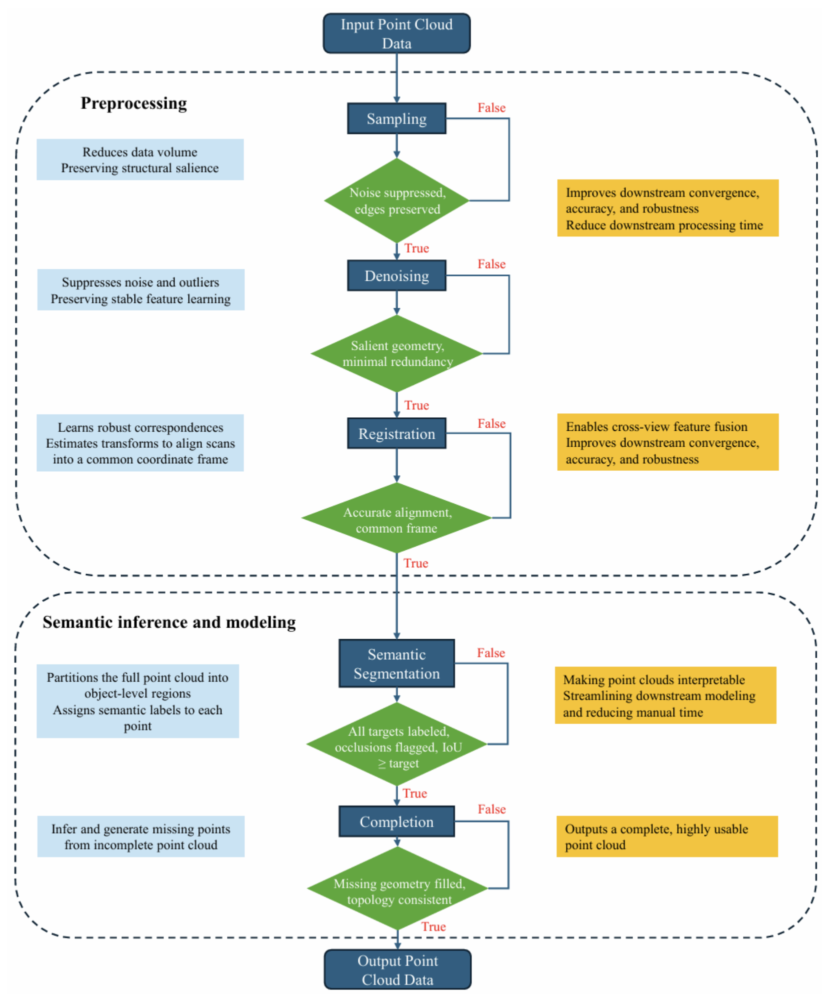

Figure 2. Proposed deep-learning-based point-cloud processing framework developed by the authors for automating point-cloud workflows in construction.

Framework Demonstration: Methods and Metrics

Sampling: The authors select a point-based differentiable network (SampleNet) that learns task-aware point selection and outputs a k-point subset that preserves structure while reducing computation. Metrics are one-sided Chamfer Distance, which measures coverage error from the full cloud to the subset, F-score at a fixed distance threshold, which balances precision and recall of coverage, and a downstream delta such as change in mean IoU or registration recall, which reflects impact on subsequent tasks.

Denoising: The authors select a point-based CNN, PointCleanNet, that predicts per-point displacements and suppresses outliers to move noisy scans toward the underlying surface while preserving edges. Metrics are Chamfer Distance, which quantifies geometric deviation from a clean reference; normal angular error, which assesses surface detail retention; and point-to-mesh distance when available, which approximates surface accuracy. The authors also report downstream impact as changes in registration recall and segmentation mIoU.

Registration: For registration, the authors use a transformer-based correspondence network (Predator) that learns cross-frame features, predicts overlap and matches, and estimates a rigid SE (3) pose robustly under low overlap and noise. Metrics are translation error and rotation error, which quantify pose accuracy; registration recall at fixed tolerances, which measures success rate; point-to-plane RMSE after alignment, which measures geometric fit; and per-pair runtime, which indicates deploy ability.

Semantic Segmentation: The authors use a sparse 3D CNN (MinkowskiNet), which operates on voxelized point clouds with sparse convolutions to capture multi-scale context efficiently and produce per-point semantic labels at building scale. Metrics are mIoU, the mean Intersection over Union that measures per-class overlap between prediction and ground truth; Overall Accuracy, the fraction of correctly labeled points; mean Class Accuracy, the average per-class recall; and throughput in points per second, which indicates deploy ability.

Completion: The authors use a transformer-based network (PoinTr) that encodes partial scans, aggregates global context, and decodes a coarse-to-fine dense point set to recover occluded surfaces while preserving part topology. Metrics are Chamfer Distance and Earth Mover's Distance, which measure geometric fidelity; F-score at a fixed distance threshold, which balances precision and recall of surface hits; normal angular error, which reflects surface quality; and point-to-mesh distance when mesh ground truth is available, which approximates surface accuracy. The authors also report downstream impact as changes in segmentation mIoU and registration recall after inserting the completed cloud into the pipeline.

Training data: While each stage requires task-specific datasets, the unified framework enables a single multi-task construction dataset that supports all core processing steps. A shared schema and consistent coordinates allow joint training, feature reuse, coherent evaluation, and reduced labeling effort, improving transfer across projects and stabilizing end-to-end performance.

Discussion and Future Work

This study proposes a conceptual, deep-learning-driven pipeline for construction point-cloud processing based on a focused review of recent studies. Rather than presenting a deployed system, the contribution is an evidence-informed organization of core tasks, denoising, sampling, registration, semantic segmentation, and completion, into a coherent workflow that better matches construction needs such as Scan-to-BIM, progress tracking, and safety-related analysis. Across the literature, three recurring themes motivate the framework:

1. Efficiency and transferability improve when models reduce per-project re-tuning via pretraining and reduced-label/weak supervision.
2. Maintaining geometry–semantics integrity is critical in cluttered, partially observed scans where thin or occluded components are easily lost.
3. Scalability and repeatability suffer when pipelines are fragmented across tools and representations, suggesting the need for more consistent intermediate outputs and quality-control checkpoints.

At the same time, the framework remains conceptual and has not yet been validated as an end-to-end system. The reviewed studies vary widely in datasets, task definitions, and evaluation metrics, which limits direct cross-study comparison, and integrated pipelines may increase training complexity and computational cost. Validation is therefore the next priority. Future work should implement a minimal prototype and evaluate it across multiple projects, with models trained on some sites and tested on unseen sites. Evaluation should include task-level measures such as registration error, semantic segmentation performance reported with IoU or F1, and completion quality reported with Chamfer distance under different levels of noise, point density, and occlusion. It should also include workflow-level measures that reflect construction outcomes, including Scan-to-BIM element accuracy and completeness, progress tracking correctness, and practical deployment indicators such as runtime and labeling effort. Finally, ablation studies and comparisons against traditional workflows and single-task DL baselines should be conducted to quantify the added value of the integrated pipeline in realistic construction settings.

Generative AI Disclosure. Generative AI tools were used solely for language refinement and editing assistance. All technical content, analysis, and framework development were conducted by the authors.

References

Chen, L., Xu, Y., & Kim, J. (2022). Deep learning applications for construction 3D data: A systematic review. *Automation in Construction*. DOI: 10.1061/(ASCE)CP.1943-5487.0001010

Valdebenito, R., & Forcael, E. (2025). Integrating Artificial Intelligence and BIM in Construction: Systematic Review and Quantitative Comparative Analysis. *Applied Sciences*. https://doi.org/10.3390/app152312470

Wang, Y., Zhou, P., Geng, G. *et al.* End-to-end point cloud registration with transformer. *Artif Intell Rev* **58**, 26 (2025). https://doi.org/10.1007/s10462-024-10985-y

Ye Z, Antwi-Afari MF, Tezel A, Manu P (2025), "Building information modeling (BIM) in project management: a bibliometric and science mapping review". *Engineering, Construction and Architectural Management*, Vol. 32 No. 5 pp. 3078–3103, doi: https://doi.org/10.1108/ECAM-04-2023-0355

Mehranfar, M., Braun, A., & Borrmann, A. (2024). From dense point clouds to semantic digital models: End-to-end AI-based automation procedure for Manhattan-world structures. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2024.105392

Prieto, S. A., Mengiste, E. T., Menon, U., & Soto, B. G. de. (2024). Current State and Trends of Point Cloud Segmentation in Construction Research. *International Symposium on Automation and Robotics* *in Construction (ISARC) Proceedings, 2024 Proceedings of the 41st ISARC, Lille, France, 972–979.* https://doi.org/10.22260/ISARC2024/0126

Charles, R. Q., Su, H., Kaichun, M., & 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. https://doi.org/10.1109/CVPR.2017.16

Rakotosaona, M.-J., Barbera, V. L., Guerrero, P., Mitra, N. J., & Ovsjanikov, M. (2019). *PointCleanNet: Learning to Denoise and Remove Outliers from Dense Point Clouds* (No. arXiv:1901.01060). arXiv. https://doi.org/10.48550/arXiv.1901.01060

Rashdi, R., Martínez-Sánchez, J., Arias, P., Qiu, Z., Rashdi, R., Martínez-Sánchez, J., Arias, P., & Qiu, Z. (2022). Scanning Technologies to Building Information Modelling: A Review. *Infrastructures*, *7*(4). https://doi.org/10.3390/infrastructures7040049

Wang, Y., & Solomon, J. M. (2019). *Deep Closest Point: Learning Representations for Point Cloud Registration* (No. arXiv:1905.03304). arXiv. https://doi.org/10.48550/arXiv.1905.03304

Chen, G., Wang, M., Yang, Y., Yu, K., Yuan, L., & Yue, Y. (2023). *PointGPT: Auto-regressively Generative Pre-training from Point Clouds* (No. arXiv:2305.11487). arXiv. https://doi.org/10.48550/arXiv.2305.11487

Kavaliauskas, P., Fernandez, J. B., McGuinness, K., Jurelionis, A., Kavaliauskas, P., Fernandez, J. B., McGuinness, K., & Jurelionis, A. (2022). Automation of Construction Progress Monitoring by Integrating 3D Point Cloud Data with an IFC-Based BIM Model. *Buildings*, *12*(10). https://doi.org/10.3390/buildings12101754

Li, R., Li, X., Fu, C.-W., Cohen-Or, D., & Heng, P.-A. (2019). *PU-GAN: A Point Cloud Upsampling Adversarial Network* (No. arXiv:1907.10844). arXiv. https://doi.org/10.48550/arXiv.1907.10844

Park, J., & Cho, Y. K. (2022). Point Cloud Information Modeling: Deep Learning–Based Automated Information Modeling Framework for Point Cloud Data. *Journal of Construction Engineering and Management*, *148*(2), 04021191. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002227

Engel, N., Belagiannis, V., & Dietmayer, K. (2021). Point Transformer. *IEEE Access*, *9*, 134826–134840. https://doi.org/10.1109/ACCESS.2021.3116304

Yue, H., Wang, Q., Yan, Y., & Huang, G. (2025). Deep learning-based point cloud completion for MEP components. *Automation in Construction*, *175*(106218). https://doi.org/10.1016/j.autcon.2025.106218