DEEP LEARNING–BASED POINT CLOUD SEGMENTATION AND RELATIONSHIP GRAPH CONSTRUCTION FOR HYDROPOWER DAM DIGITAL TWINS

Alwyn Mathew, Soheila Kookalani, and Ioannis Brilakis
University of Cambridge, England, United Kingdom

Abstract

This paper proposes a deep learning pipeline for generating semantic meshes and relationship graphs from intensity-only laser point clouds of hydropower plants. It segments point clouds into object-level instances using supervised models and derives spatial relationships, including proximity and containment, between components. These are refined into a structured graph, providing a lightweight representation for digital twin applications. Replacing rule-based methods with learning-based segmentation enables automated graph construction for dam infrastructure despite irregular geometries, limited features, and the lack of standardised schemas.

Introduction

Hydropower dams are long-lived, safety-critical assets that underpin electricity supply, water storage, flood control, and regional resilience. Their operation and maintenance increasingly depend on reliable “as-is” information about structural components as well as embedded mechanical and hydraulic equipment. Digital twin (DT) approaches have therefore gained traction to link physical assets to continuously usable digital representations that support inspection planning, performance assessment, anomaly detection, and scenario analysis. Recent large-scale evidence also shows that the DT concept is used inconsistently across the built environment, motivating clearer operational definitions and explicit data requirements for DT implementations as reported by Abdelrahman et al. (2025); Kookalani et al. (2026).

A practical DT for civil infrastructure is not only a geometric model, but an integrated information system that connects geometry, semantics, and relationships to enable querying, automation, and analytics. Construction DT research emphasizes that robust DT pipelines require systematic data capture, semantic structuring, and interoperability-aware information integration across lifecycle workflows as described by Tuhaise et al. (2023).

In this context, terrestrial laser scanning and photogrammetry-derived point clouds have become key modalities for capturing high-fidelity “as-is” geometry. However, translating raw point clouds into usable DT-ready representations remains challenging due to occlusions, clutter, irregular shapes, scale, and the need for detailed semantic labelling as outlined by Yue et al. (2024).

Most scan to building information modelling (BIM) and point-cloud-to-semantic-model workflows have been developed around buildings and industrial facilities, where component taxonomies are mature and standardized schema are relatively well supported as reported in Alavi et al. (2024). For dam infrastructure, the semantic breadth of elements and the heterogeneity between structural and electromechanical subsystems create gaps in standardization and representation. Commonly used exchange standards were originally designed for buildings and do not fully cover dam-specific components and hierarchies, which limits direct “BIM-first” relationship modelling for DT creation. This mismatch motivates DT representations that can be inferred directly from sensed data, especially when a complete parametric BIM is infeasible or uneconomical to create.

A second practical limitation is that many real-world infrastructure point clouds, especially open datasets, do not include rich appearance cues (RGB), and instead provide geometry plus intensity only. This reduces the discriminative signal available to segmentation algorithms and amplifies the need for methods that can learn robust geometric and reflectance patterns from limited features, while also handling the severe class imbalance typical in infrastructure scenes. Domain gaps between synthetic data and real scans further complicate generalization according to Hu et al. (2024).

Finally, DT utility depends heavily on relationships. Even if components are segmented accurately, downstream tasks such as navigation, hierarchical organization, adjacency reasoning, and structured asset inventories require an explicit relationship graph. Prior work in the built environment has shown the value of enriching geometric reconstructions into graph-structured semantic representations, enabling machine-readable links between objects beyond pure geometry as noted by Werbrouck et al. (2020).

To address these challenges, this paper proposes an end-to-end, learning-based framework for deriving digital twin–ready representations directly from intensity-only dam point clouds. The method is evaluated on the Platanovrisi Dam dataset, covering both exterior dam surfaces and interior hydropower plant spaces. The main contributions of this work are summarised as follows:

1. Learning-based segmentation for large-scale dam point clouds: We develop a deep learning pipeline

Figure 1: Platanovrisi dataset: HPP ground level (left) and turbine floor (right).

tailored for non-planar, complex dam surfaces that enables semantic labeling of intensity-only point clouds.

2. Semantic mesh generation and instance relationship extraction: The pipeline converts segmented point clouds into structured semantic meshes and identifies spatial relationships between individual components.
3. Relationship graph construction and connectivity modeling: Semantic predictions are transformed into a connected graph representation that captures adjacency and containment relationships among dam components.
4. End-to-end digital twin workflow from point cloud to UML (Unified Modeling Language): The connected graph can be exported into a generic format and used to generate UML class diagrams, providing a practical scaffold for digital twin modeling and downstream system integration.

Background

DTs in the built environment are commonly positioned as integrated digital representations that can be updated and used to support decision-making across the asset lifecycle. Large-scale bibliometric analyses highlight that the term digital twin is often used ambiguously in built-environment literature, but converges on core characteristics such as representational fidelity, purposeful synchronization and update mechanisms, and actionable use in analysis and operations as described by Abdelrahman et al. (2025). Construction-focused DT reviews in Tuhaise et al. (2023) further emphasize that DT value depends on connecting sensing, data integration, and analytics into an end-to-end workflow, rather than treating the DT as a standalone 3D model. For civil infrastructure, DT research is expanding beyond buildings into bridges, tunnels, and dams, where monitoring and risk assessment are central. For example, recent work in Tian et al. (2025) demonstrates DT approaches for high arch dams by fusing physics-based models with learning-based methods and uncertainty quantification to support real-time state estimation and failure risk assessment. These efforts underscore two recurring bottlenecks for infrastructure DTs: (1) obtaining up-to-date, spatially consistent as-is representations, and (2) structuring heterogeneous information into computable forms that support reasoning and automation. A key implication is that DT-ready representations must encode more than geometry. They must support asset semantics and connectivity or hierarchy, enabling queries such as "which elements belong to the turbine floor?", "what is adjacent to the intake tower?", or "which equipment is connected to which pipes?". Knowledge graph approaches in Tuhaise et al. (2023) explicitly target this integration challenge by using graph-based models to unify multi-source information for built assets, thereby improving cross-system interoperability and downstream analytics.

Laser scanning produces dense 3D point clouds that capture detailed surface geometry, making them attractive for modeling existing assets and enabling downstream DT workflows. However, scan-to-BIM pipelines typically require preprocessing, semantic segmentation and instance grouping, and object reconstruction and relationship modelling. Recent work of Patil and Kalantari (2025) emphasizes how segmentation accuracy can directly influence reconstruction outcomes and highlights practical hybrid workflows that combine deep learning segmentation with geometric fitting for downstream modelling. Across the Architecture, Engineering, and Construction (AEC) domain, deep learning has become the dominant approach for point cloud semantic segmentation because it reduces reliance on handcrafted geometric features and improves robustness across complex scenes as shown by Liu et al. (2025). A recent review in Yue et al. (2024) summarizes how point-cloud deep learning has been increasingly adopted for construction tasks, while also documenting persistent challenges: data scarcity, domain shift, scale variation, computational cost, and class imbalance, issues that are amplified in large infrastructure environments. Seminal architectures such as PointNet and PointNet++ as in Qi et al. (2017a) established the foundational concept of learning directly from unordered point sets, with hierarchical variants capturing local-to-global context for improved segmentation in complex geometries. In the construction and infrastructure domain, prior work reported in Yin et al. (2021) has demonstrated the effectiveness of deep learning-based segmentation for industrial scenes and transportation assets, including ResPointNet++ for industrial point clouds and hierarchical graph-based models for bridge component segmentation. Nevertheless, dam environments introduce additional complexity:

- Feature limitations: Many datasets are intensity-only (no color information), requiring models to separate classes using geometry and reflectance patterns rather than appearance cues.
- Heterogeneous scale and geometry: Dams mix large continuous surfaces with small mechanical compo-

nents and dense equipment layouts, increasing confusion between visually similar elements.

- Domain shift and annotation scarcity: Methods trained on synthetic BIM-derived point clouds often underperform on real scans due to the domain gap. Domain adaptation methods such as DawNet explicitly address this for BIM-to-scan segmentation Hu et al. (2024).
- Class imbalance: Infrastructure point clouds typically have extreme imbalance, motivating loss reweighting, focal losses, and sampling strategies, as also discussed in the proposed pipeline.

Even with accurate semantic segmentation, DT value is limited if the output remains an isolated labelled point cloud. Relationship modelling provides the scaffolding for higher-level reasoning, data integration, and automation. In conventional BIM workflows, relationships are typically defined through schema-specific hierarchies and predefined object types. However, infrastructure DTs often face schema gaps for dams. Standardized BIM schemas may not fully capture dam-specific components and their relationships, motivating alternative graph-first representations inferred directly from data. Two complementary directions in the literature motivate relationship graph construction: (1) Semantic enrichment / scan-to-graph: The scan-to-graph concept integrates scan-to-BIM ideas with semantic network by representing objects and their relationships as a graph, enabling richer queries and greater interoperability than geometry alone as presented in Werbrouck et al. (2020). This approach is particularly relevant when full parametric BIM creation is not feasible, but a computable representation is still required. (2) Connectivity-aware modelling from incomplete data: Infrastructure scans are commonly incomplete due to occlusions and access constraints. Methods that exploit connectivity relationships to improve as-built modelling from incomplete point clouds demonstrate the importance of explicitly reasoning about adjacency and inter-element connections, rather than treating components independently as shown by Kim and Kim (2021). Graph-based asset as outlined by Tuhaise et al. (2023) representations also align naturally with knowledge graph integration approaches for DTs, where nodes or edges support linking geometry-derived entities to metadata, sensor streams, inspection records, and analytical outputs. For dam DTs, this suggests a pragmatic pathway: infer lightweight spatial relations from point cloud instances, progressively enrich those graphs with domain semantics as schema mature and expert knowledge becomes available.

Methodology

Constructing a geometric model of a dam involves several key stages. The point cloud is first segmented into individual components, such as walls, gates, and turbines, breaking down the complex structure into manageable elements. Spatial proximity analysis is then applied to these components to infer their interactions, resulting in a relationship graph that captures spatial and semantic connections. Surface reconstruction techniques are subsequently employed on the segmented point clouds to generate continuous meshes, producing a detailed and coherent geometric representation of the dam.

Semantic mesh

The proposed segmentation network follows the PointNet architecture according to Qi et al. (2017a) for point-wise semantic labeling. It begins with an input transformation module (T-Net) that learns a 3×3 affine matrix to align the raw point cloud, thereby improving invariance to rigid geometric transformations. Feature extraction is performed through a sequence of shared multilayer perceptrons implemented as one-dimensional convolution layers, progressively expanding the channel dimension from the input to 64, 64, 64, 128, and finally 1024. A second transformation network is applied after the early convolution layers to learn a 64×64 feature alignment matrix, enabling canonicalization in the learned feature space. Global context is captured by max pooling across all points to produce a 1024-dimensional global feature vector. For semantic segmentation, the global feature is concatenated with per-point local features to form a fused representation for each point. This combined feature is processed by a segmentation head composed of additional shared multilayer perceptrons with channel dimensions 1088, 512, 256, and 128, followed by a final classification layer that outputs per-point semantic labels for the target number of classes.

Several modifications are introduced relative to the original PointNet to improve robustness and representational power. The network incorporates flexible normalization and dropout strategies, which enhance stability and reduce overfitting, and it can process additional input features beyond spatial coordinates, such as intensity or surface normals, enabling richer point representations. Transformation modules are refined to better align features across points while supporting these enhancements, ensuring more effective learning of both local and global geometric structures. Furthermore, a hierarchical variant inspired by PointNet++ as presented in Qi et al. (2017b) is included to capture local geometric patterns at multiple scales. This extension employs set abstraction and feature propagation modules to encode multi-scale contextual information, significantly improving the network's ability to model complex spatial structures in large-scale infrastructure environments. Strategies for handling class imbalance and tile-based oversampling further enhance segmentation of underrepresented classes, and these are described in detail in the Discussion section. The segmentation method has shown improved accurately as shows in Figure 2 and Table 1.

Surface reconstruction is essential for converting segmented point clouds into continuous, high-fidelity geometric models suitable for digital twin applications. In this study, surface normals computed during segmentation provide local geometric context, guiding the reconstruc-

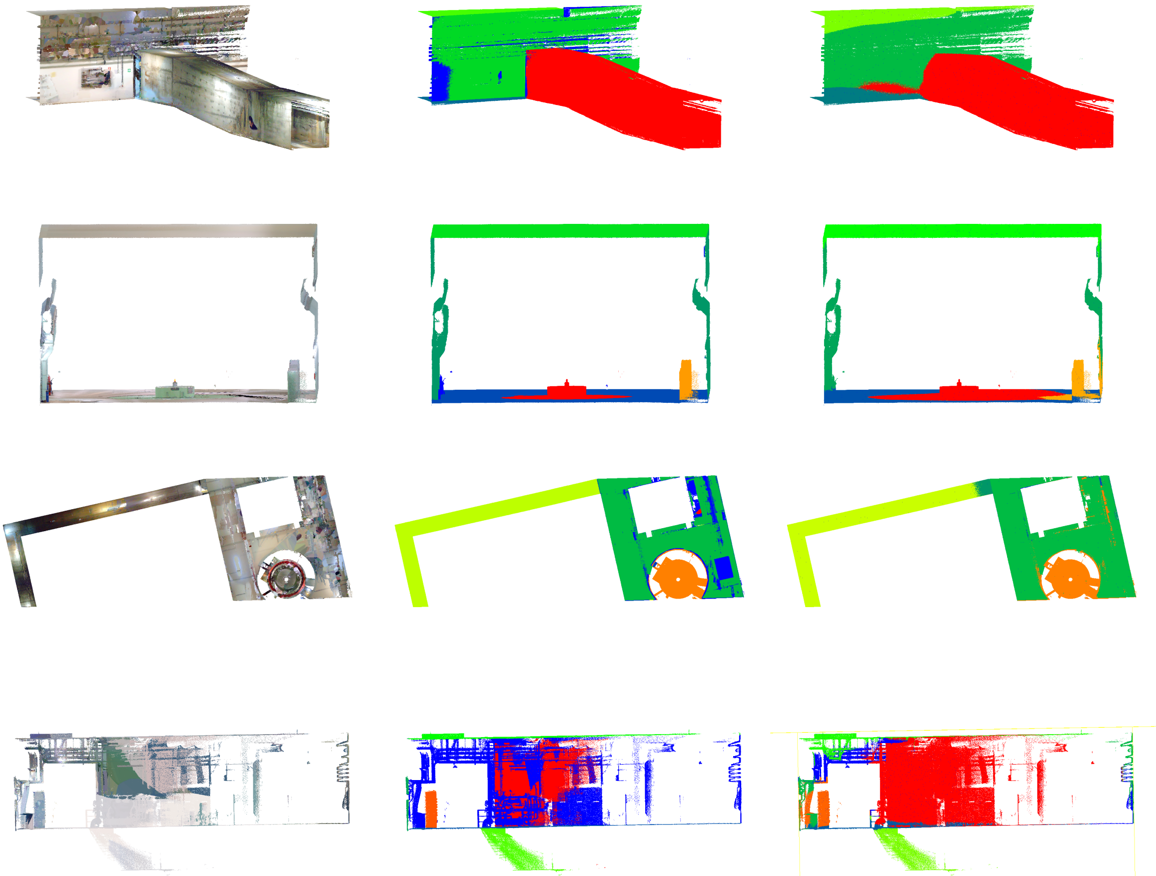

Figure 2: Platanovrisi dataset: color pointcloud (left), ground truth segmentation (middle) and predicted segmentation (right).

Figure 3: Semantic mesh generated from sematic pointcloud of the Platanovrisi dataset.

tion algorithm to accurately capture the orientation and topology of each component. We employ the Ball Pivoting Algorithm (BPA) Bernardini et al. (1999) to generate triangular meshes, where a virtual ball rolls and pivots across points to form triangles, adapting its radius to handle both smooth and intricate surfaces. This approach is particularly effective for hydropower plants, which combine curved, flat, and angular geometries such as turbine blades, pipes, and structural reinforcements. Challenges such as noise, outliers, occlusions, and irregular geometries are addressed through careful parameter tuning and post-processing techniques, including smoothing, hole filling, and mesh decimation. The resulting semantic meshes provide a coherent and accurate representation that serves as the foundation for visualization, simulation, and downstream tasks, integrating seamlessly with the segmentation and relationship graph pipeline to model complex hydropower infrastructure as shown in Figure 3.

Instance relationships

In traditional BIM workflows, relationship graphs are usually derived from well-defined object hierarchies and standardized schemas. However, for dam infrastructure, the creation of complete BIM models is limited because commonly used exchange formats, such as Industry Foundation

Classes (IFC), were originally developed around building and industrial facility taxonomies Alavi et al. (2024); Zhang et al. (2024) and do not yet provide comprehensive coverage of the heterogeneous structural and electromechanical component hierarchies typical of dam infrastructure.

To address this gap, this study proposes constructing a relationship graph directly from segmented point cloud data. Unlike conventional BIM graphs, this representation is lightweight and inferred, relying on spatial and semantic cues rather than predefined schemas. The graph is generated by establishing "connected" relationships based on spatial adjacency between segmented components. Each node represents a single component, while edges indicate spatial or semantic associations.

A k-dimensional tree (KD-Tree) is employed to efficiently index points and perform radius based searches within the point cloud. For each pair of segments, if any point from one segment lies within a radius of 0.5 m of a point in the other, an adjacency edge is created between their corresponding graph nodes. The KD-Tree is particularly effective in handling the geometric complexity of hydropower plants, which include irregularly shaped components such as turbines, generators, pipes, and reservoirs. Manual identification of such relationships is impractical due to this complexity, making automated inference highly valuable.

Beyond adjacency, components are also associated with higher-level structural abstractions, such as the dam body, ground floor, or turbine floor. If a component is spatially contained within one of these broader regions, a hierarchical "has" relationship is inferred (e.g., the ground floor has a pipe). This introduces an organized semantic structure, though it is limited to spatial and containment logic and does not capture functional or operational dependencies. Capturing these deeper relationships would require domain-specific schemas and expert knowledge, which are currently unavailable for dam infrastructure.

The basic relationships extracted from the point cloud are exported to a generic format, such as puml. This puml file serves as the input to PlantUML, to automatically generate a draft UML diagram. While the generated UML is not complete, it provides an initial framework that can be manually refined, extended, and annotated to include more detailed relationships. Figure 4 shows a preliminary example of the resulting graph, illustrating inferred connectivity and hierarchical groupings among segmented components.

Table 1: Segmentation evaluation result on test split of Platanovrisi dataset

Method	Precision	Recall	F1	IoU
RANSAC	0.5-0.6	0.4-0.5	2 0.5	0.40.4-0.6
Our	0.6813	0.9271	0.7393	0.6253

Dataset

The Platanovrisi Dam is a concrete facility located roughly 7.5 km southeast of the Thesauros Dam in northeastern Greece, has been operational since 1999 (Figure 1). It forms a reservoir capable of storing around 90 million cubic meters of water. Together with the 175-meter-tall earthen Thesauros Dam, among the tallest in Europe, Platanovrisi functions as part of a coordinated hydropower system that maximizes efficiency through water reuse and nighttime pumping between the reservoirs. The combined infrastructure supports a total generation capacity of 500 MW under a Greek-Bulgarian water management agreement. The dataset captures key structural elements of the dam, including the spillway, auxiliary spillway, downstream face, toe still basin, crest, and an intake tower. Interior scans encompass two levels of the hydropower plant: the ground-floor turbine gallery, providing maintenance access, and the lower turbine floor, which houses critical operational components such as the turbine well, piping, tanks, pumps, and control panels.

The dataset comprises approximately 369 million valid points distributed across 33 spatial tiles and annotated with 10 semantic classes: floor, wall, ceiling, stairs, crane, tunnels, electric distributor, turbine well, penstock, and water tank. The class distribution is severely imbalanced, as shown in Table 2, with wall and tunnels together accounting for over 55% of all points while minority classes such as crane and water tank each represent less than 1%. The dataset was partitioned into training, validation, and test splits using an 8:1:1 ratio at the tile level to prevent spatial leakage between splits.

Table 2: Distribution of dominant classes in the Platanovrisi dataset

Classes	#points (in M)	Percentage
Wall	105	28.8
Tunnels	98	27
Floor	73	20.2
Ceiling	49	13.4
Penstock	13	3.7
Electric distributor	9	2.6
Stairs	7	2
Turbine well	3	1.1
Crane	2	0.6
Watertank	2	0.6

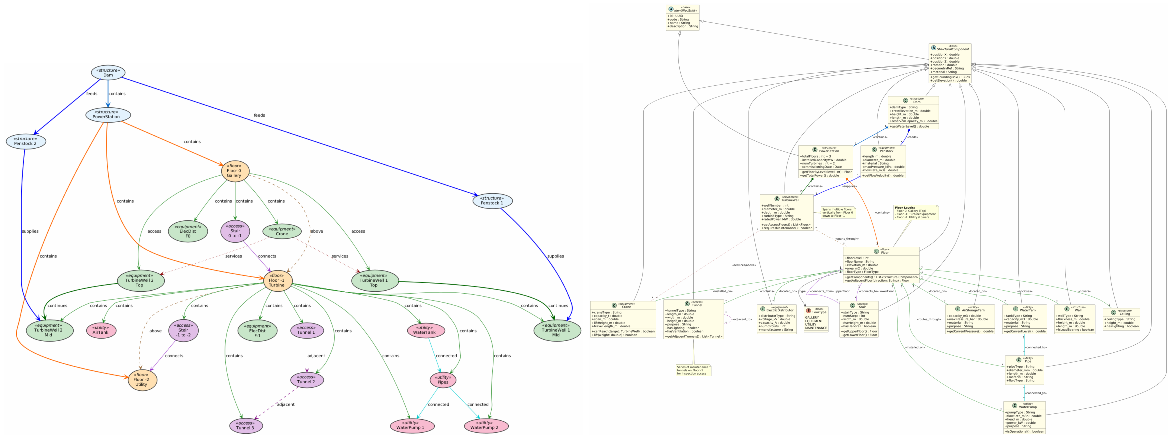

Figure 4: Relationship graph extracted from semantic output (left) and UML from relationship graph (right).

Discussion

The integration of semantic mesh generation and relationship graph construction relies critically on accurate deep learning–based segmentation, which forms the foundation of the proposed dam digital twin pipeline. Unlike conventional building structures, dams exhibit highly irregular geometries such as curved spillways, inclined faces, and embedded tunnels that challenge both traditional scan-to-BIM workflows and primitive-fitting approaches such as RANSAC as presented in Mathew et al. (2025) as seen in segmentation evaluation Table 1. Learning-based segmentation is therefore essential for capturing these complex, non-planar surfaces and heterogeneous mechanical components directly from point cloud data.

These geometric difficulties, together with scarce public datasets and inconsistent component definitions across projects, further complicate standardization and modeling. Additional obstacles include incomplete scans, missing hierarchical information, and limited prior knowledge of functional interactions between components. By leveraging deep neural networks to produce semantic predictions that drive surface reconstruction and relationship inference, the proposed pipeline generates compact yet structured representations that can support spatial queries and component-level data integration. This output provides an initial foundation for digital twin applications tailored to dam infrastructure, where the geometric and semantic complexity poses greater challenges than in typical building environments.

Semantic segmentation of large-scale infrastructure point clouds, however, often suffers from severe class imbalance as shown in Table 2. Structural components such as walls and tunnels dominate the point clouds, while smaller elements such as pipes, turbines, or auxiliary equipment contain far fewer points. To mitigate this issue, two complementary strategies were employed during training: loss reweighting and tile-based oversampling. Class loss balancing adjusts the contribution of each category to the training objective in proportion to its prevalence, penalizing misclassifications of underrepresented classes more heavily. Various weighting schemes can be applied, including balanced weighting based on inverse class frequency, simple inverse-frequency weighting, and normalized custom weights. In addition, focal loss is used to down-weight easy examples and focus optimization on hard-to-classify points, while label-smoothed negative log-likelihood further improves generalization by preventing overconfident predictions.

At the data level, tile-based oversampling ensures rare classes are sufficiently represented during training. Point clouds are divided into spatial tiles, and tiles containing rare classes are sampled more frequently based on the rarity of the least common component within the tile, with a maximum oversampling limit to prevent excessive duplication. This increases exposure to underrepresented geometries without artificially inflating individual points. Together, loss reweighting and tile oversampling act at complementary stages of the pipeline, aiming to improve segmentation accuracy for minority classes in complex hydropower environments. Training robust models for dam infrastructure remains challenging due to the limited availability of high-quality, publicly accessible datasets. The Platanovrisi dataset used in this study, while adhering to high-quality standards, still suffers from incomplete scans and inconsistent color mapping. As a result, the model was trained exclusively on laser intensity data, avoiding reliance on potentially unreliable color information to ensure robust generalization. A key limitation of the current study is that the pipeline has been evaluated on a single site (Platanovrisi Dam). While the architectural and geometric diversity within this dataset provides a meaningful testbed, the generalizability of the trained model to other dam structures with different typologies, materials, or scanning conditions has not yet been assessed. Evaluating transferability across multiple dam sites and identifying domain adaptation needs are planned for the follow up journal extension.

The segmentation model receives 4 input channels per point (XYZ coordinates and laser intensity) and was trained with a batch size of 5 using the Adam optimizer with a learning rate of 0.0008, weight decay of 1e-4, and cosine annealing with a 5 epoch warmup to a minimum learning rate of 1e-6. Group normalization was used for stability with small batches, and gradient clipping with a maximum norm of 1.0 was applied. Training ran for up to 150 epochs with early stopping (patience of 15 epochs). To address the severe class imbalance, balanced inverse frequency weighting was combined with focal loss (alpha = 0.75, gamma = 2.0). As reported in Table 1, the model segments dominant structural categories such as floor, wall, and ceiling with reasonable accuracy. However, minority classes including crane, turbine well, and water tank remain challenging due to their extremely low representation in the dataset (each below 1% of total points, as shown in Table 2). This reflects the inherent difficulty of dam point cloud segmentation compared to building environments, where class distributions tend to be more balanced and component geometries more standardized.

The quality of the relationship graph depends directly on the accuracy of the upstream segmentation. In the proposed pipeline, adjacency (''connected'') edges are inferred using a KD-Tree radius search: for each segment pair, if any point from one segment falls within a radius of 0.5 m of a point in the other, an adjacency edge is created between their corresponding nodes. Containment (''has'') relationships are assigned when a component lies spatially within a higher level region such as the ground floor or turbine floor. Segmentation errors, particularly over-segmentation of large surfaces such as walls or confusion between adjacent classes, can introduce spurious edges or miss valid connections. For instance, if a wall segment is split into multiple fragments, each fragment generates its own node and edges, inflating the graph with redundant adjacency links. Conversely, if a small component such as a pipe is absorbed into a neighbouring wall segment, the corresponding node and its relationships are lost entirely. Since no ground truth relationship graph currently exists for the Platanovrisi dataset, quantitative evaluation of graph quality (e.g., edge level precision and recall) is not feasible at this stage. Constructing such a reference through expert annotation and conducting sensitivity analysis of the distance threshold are planned for the follow up journal paper.

Conclusion

This paper presented an end-to-end pipeline for constructing relationship graphs from intensity-only laser point clouds of hydropower dams by integrating deep learning based segmentation, surface mesh reconstruction, and instance-level relationship inference. Evaluated on the Platanovrisi Dam dataset across exterior structures and interior hydropower plant spaces, the framework demonstrates the feasibility of producing structured geometric and relational models directly from laser scans without relying on complete parametric BIM schemas. Dam infrastructure poses greater challenges than typical building environments due to irregular geometries, heterogeneous component types, the absence of standardized schemas, and severe class imbalance in point cloud data. Remaining challenges include limited training data, poor performance on minority classes, the lack of a ground truth relationship graph for quantitative evaluation, and the current restriction to spatial relationships. Future work includes multi-site evaluation, construction of an expert annotated reference graph, improved handling of class imbalance, and richer domain informed semantics toward scalable digital twins for hydropower infrastructure.

Acknowledgments

This work is supported by the D-HYDROFLEX project under the European Union's Horizon Programme (Grant Agreement No. 101122357), BUILDSPACE project under the Horizon Europe (Grant Agreement No. 101082575) and M-TABS under Innovate UK (Grant Agreement No. 10132366).

References

Abdelrahman, M., Macatulad, E., Lei, B., Quintana, M., Miller, C., and Biljecki, F. (2025). What is a digital twin anyway? deriving the definition for the built environment from over 15,000 scientific publications. Building and Environment, 274:112748.

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024). Integrated Building Intelligence. Springer.

Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., and Taubin, G. (1999). The ball-pivoting algorithm for surface reconstruction. IEEE Transactions on Visualization and Computer Graphics, 5(4):349–359.

Hu, D., Gan, V. J., and Zhai, R. (2024). Automated bim-to-scan point cloud semantic segmentation using a domain adaptation network with hybrid attention and whitening (dawnet). Automation in Construction, 164:105473.

Kim, H. and Kim, C. (2021). 3d as-built modeling from incomplete point clouds using connectivity relations. Automation in Construction, 130:103855.

Kookalani, S., Green, S., Luo, P., Alavi, H., Parn, E., Sun, Z., and Brilakis, I. (2026). Mapping digital twin applications in infrastructure and the built environment across research types, methods, sectors, phases, and scales. Automation in Construction, 182:106778.

Liu, H., Su, H., Kookalani, S., Cao, M., Lei, Y., and Wu, W. (2025). Attention-enhanced spatiotemporal deep learning-based automatic warning model with uncertainty estimation in dam safety. Structural Health Monitoring, page 14759217251358537.

Mathew, A., Alavi, H., Wang, G., Su, Y., Che, W., and Brilakis, I. (2025). Geometric model of hydropower plants: a foundational step towards creating a hydropower digital twin. In EC3 Conference 2025, volume 6, pages 0–0. European Council on Computing in Construction.

Patil, J. and Kalantari, M. (2025). Automatic scan-to-bim—the impact of semantic segmentation accuracy. Buildings, 15(7):1126.

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

Qi, C. R., Yi, L., Su, H., and Guibas, L. J. (2017b). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. arXiv preprint arXiv:1706.02413.

Tian, J., Chen, C., Zhang, L., Chen, J., Huang, H., and Zhang, P. (2025). A digital twin method for real-time analysis of structural deformation and failure for high arch dams. Engineering Structures, 345:121426.

Tuhaise, V. V., Tah, J. H. M., and Abanda, F. H. (2023). Technologies for digital twin applications in construction. Automation in Construction, 152:104931.

Werbrouck, J., Pauwels, P., Bonduel, M., Beetz, J., and Bekers, W. (2020). Scan-to-graph: Semantic enrichment of existing building geometry. Automation in Construction, 119:103286.

Yin, C., Wang, B., Gan, V. J., Wang, M., and Cheng, J. C. (2021). Automated semantic segmentation of industrial point clouds using respointnet++. Automation in Construction, 130:103874.

Yue, H., Wang, Q., Zhao, H., Zeng, N., and Tan, Y. (2024). Deep learning applications for point clouds in the construction industry. Automation in Construction, 168:105769.

Zhang, S., Zhang, S., Wang, C., Zhu, G., Liu, H., and Wang, X. (2024). Extended ifc-based information exchange for construction management of roller-compacted concrete dam. Automation in Construction, 163:105427.