Cross-domain urban and heritage point cloud benchmarks for semantic segmentation: insights from the 3M (Mapping Mappano Municipality) and ArCH^X datasets.

Francesca Matrone [1], Andrea Maria Lingua [1], Arnadi Murtiyoso [2], Madhavan Sridhar [3], Luis Javier Sanchez Aparicio [4 5], Ruben Santamaria Maestro [4], Pablo Sanz-Honrado [6], Hina Pande [3], Poonam Seth Tiwari [3], Shefali Agrawal [3]

[1] Dept. of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy – francesca.matrone@polito.it, andrea.lingua@polito.it
[2] Université de Strasbourg, INSA Strasbourg, CNRS, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000 Strasbourg, France - arnadi.murtiyoso@insa-strasbourg.fr
[3] Indian Institute of Remote Sensing, Dehradun - 248001, Uttarakhand, India - maddysridhar12345@gmail.com, hina@iirs.gov.in, poonam@iirs.gov.in, shefali_a@iirs.gov.in
[4] Dept. of Construction and Building Technology in Architecture (DCTA), Escuela Técnica Superior de Arquitectura de Madrid (ETSAM), Universidad Politécnica de Madrid, Av. Juan de Herrera,4,28040 Madrid, Spain - lj.sanchez@upm.es, ruben.santamaria.maestro@upm.es
[5] Dept. of Cartographic and Land Engineering. University of Salamanca, High Polytechnic School of Ávila, Hornos Caleros, 50, 05003, Ávila, Spain - luisj@usal.es
[6] Institute of Physical and Information Technologies Leonardo Torres Quevedo (ITEFI), CSIC, C/Serrano 144, 28006 Madrid, Spain - pablo.sanz@csic.es

Keywords: Semantic segmentation, point clouds, cultural heritage, architecture, urban environment, dataset.

Abstract

Semantic segmentation of 3D point clouds is a key enabler for intelligent urban analysis and cultural heritage documentation. However, current benchmarks remain uneven for different reasons: from the fragmented scenes of cityscape scenarios where a unique environment is not provided, to the difficulties in representing the great variety of architectural styles and lexicons in the cultural heritage (CH) domain. This study introduces two complementary datasets, the 3M (Mapping Mappano Municipality) dataset and the ArCH^X (Architectural Cultural Heritage eXpanded) dataset, meant to enhance cross-domain learning between urban and architectural heritage environments. The 3M dataset, derived from UAV photogrammetry and GNSS-based surveys over an 11.9 km² area near Turin (Italy), provides a highly accurate, semantically labelled urban point cloud (~900 M points) suitable for urban planning, automatic mapping, and infrastructure design. In parallel, the ArCH^X dataset extends the existing ArCH benchmark with new scenes from Portuguese masonry gates and Indian Nagara-style temples, reaching over 400 million points. These additions diversify architectural typologies and improve the generalisation capacity of AI models applied to heritage semantic segmentation tasks. Methodological challenges in harmonising class taxonomies, managing heterogeneous materials, and aligning cultural architectural semantics across datasets are eventually discussed. Collectively, the 3M and ArCH^X datasets establish a novel foundation for cross-domain benchmarking in 3D semantic segmentation, advancing data-driven research in digital twins, mapping, Historic Building Information Modelling (HBIM), and AI-assisted heritage documentation.

1. Introduction

1.1 Semantic segmentation of point clouds

Semantic segmentation of point clouds has emerged as a fundamental enabler in modern computer vision and spatial data analysis. By attributing semantic labels, semantic segmentation transforms raw, unstructured 3D data into rich, interpretable information essential for intelligent systems across multiple domains. One compelling motivation comes from the autonomous driving sector, where understanding the environment in fine detail is fundamental.

Unlike coarse object detection (e.g., bounding boxes), semantic segmentation provides dense, point-level understanding of the scene enabling a more precise navigation, object avoidance, and lane awareness (Guo et al., 2020). On the other side, in the cultural heritage (CH) domain, semantic segmentation enables the conversion of raw point clouds into intelligible, labelled data, identifying components like facades, structures, vegetation, or damaged areas. This annotated data supports e.g. the scan-to BIM/HBIM (Historic Building Information Modelling) process (Croce et al., 2023), preservation, conservation and maintenance activities as well as thematic mapping.

1.2 Objectives

In this framework, the rise of high-resolution uncrewed aerial vehicles (UAVs), lidar techniques and mobile mapping systems, has dramatically increased the volume and variety of 3D geospatial data. The resulting datasets are often heterogeneous in structure, noisy, and semantically inconsistent. As such, they pose critical challenges for accurate processing, seamless integration into digital twins or decision-support systems and analysis with AI algorithms. In fact, the possibility of building robust deep learning models on top of them surely benefits from high-quality labelled datasets that, currently, are still lacking in the CH sector and flawless urban environment. For this reason, the new 3M (Mapping Mappano Municipality) dataset and the extended version of the ArCH dataset, namely ArCH^x, are proposed in this paper.

2. Related works

Large-scale urban benchmarks have driven progress in semantic segmentation and scene understanding. Pioneering terrestrial datasets such as Semantic3D (Hackel et al., 2017) established billion-point, manually labelled benchmarks for outdoor scenes, enabling data-hungry deep learning at survey-grade resolution.

https://doi.org/10.5194/isprs-archives-XLVIII-1-W6-2025-131-2025 | © Author(s) 2025. CC BY 4.0 License.

Building on mobile and automotive lidar, SemanticKITTI (Behley et al., 2019) densely annotated all KITTI odometry sequences for full 360° lidar, defining single-scan and multi-scan tasks and catalyzing temporal methods.

While urban mobile laser scanning resources broaden diversity: on one side, Paris-Lille-3D (Roynard et al., 2018) provides high-quality ground truth over ~2 km in two French cities with 50 semantic object classes; on the other, Toronto-3D (Tan et al. 2020) adds North-American roadway scenes with ~78 M points and eight semantic classes. As regard to the Paris-Lille 3D, one of its strengths lies in its high quality manual annotation and diverse class set; however, it remains limited in its extension (road scenes or single urban scenes with several occlusions) and the point-density is constrained by the vehicle-based acquisition method.

More recently, the SensatUrban dataset (Hu et al. 2021) was introduced, covering approximately 7.6 km² across three UK cities with nearly three billion photogrammetric points and 13 semantic labels. Its major advantage is the large spatial coverage and very high sampling resolution, which draws it closer to the Mappano municipality dataset, here proposed. Whilst the latest WHU-Urban3D dataset (Han et al., 2024) addresses some gaps by providing richly annotated LiDAR point clouds for urban scenes, improving representativeness of varying city forms.

However, most of these urban and city landscape datasets usually comprise single scenes without a comprehensive view of a complete and continuous urban area. For these reasons, the 3M dataset aims to provide a new perspective on an urban dataset for autonomous driving, automatic cartography reconstruction, and telecommunication design, among others, offering a unique scenario that is interconnected and interrelated from the built-up area to the surrounding context.

On the other hand, the ArCH dataset (Matrone et al., 2020), the Images&PointClouds Cultural Heritage Dataset (Pellis et al., 2025) and the ZAHA benchmark for façade semantic segmentation (Wysocki et al., 2025) are well-known dataset in the framework of architectural heritage classification and semantic segmentation. However, they still lack a variety of architectural styles, lexicons, and geographical augmentation. Thus, an extension of the ArCH dataset was deemed necessary, and the updated ArCH^{X} (Architectural Cultural Heritage point clouds for classification and semantic segmentation - eXpanded) dataset is proposed.

3. Datasets

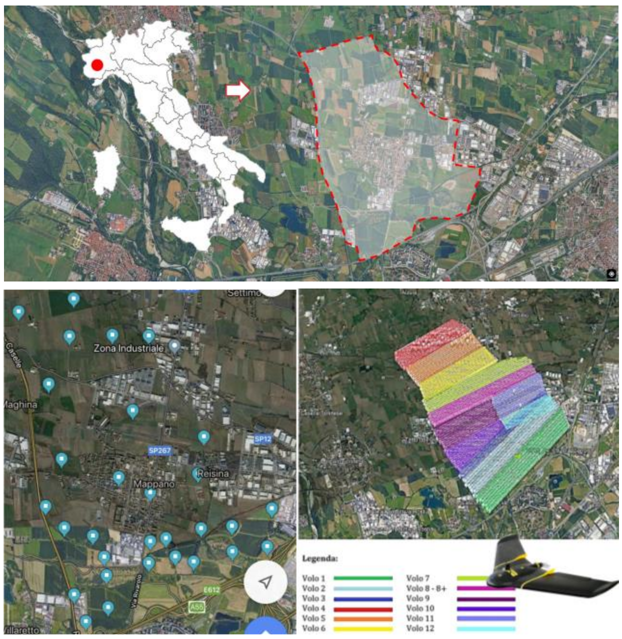
3.1 3M – Mapping Mappano Municipality

The Mappano municipality is located within the first surrounding of the metropolitan city of Turin, in the north-west part of Italy. Since it is the youngest municipality in Italy, officially formed in 2017, it needed a new and updated cartography. For this reason, a detailed UAV survey of the whole municipality was carried out.

The first phase involved conducting a 3D metric survey that combined traditional topographic techniques (GPS/GNSS) with advanced photogrammetric methods. The process began with the measurement of the topographic network to establish the framing coordinates. Subsequently, UAV-based aerial photogrammetry was employed at varying levels of detail and representation scales. Two different drones were used for this purpose: the fixed-wing *eBee PLUS* by SenseFly for wide-area territorial mapping, and the *DJI Phantom 4 RTK* for detailed urban object surveys. 30 Ground Control Points (GCPs) were located across the site (Figure 1a), and their positions were determined using a GNSS receiver.

A multi-flight drone strategy was adopted to ensure full coverage of the municipal area, totalling approximately 11.9 km². Image capture was performed using the *senseFly SODA* camera, which integrates a 13.8 × 8.8 mm RGB sensor with a 10.6 mm focal length and a 20-megapixel global-shutter configuration. This setup enabled a theoretical ground sampling distance of approximately 2.9 cm/pixel at a flight altitude of 120 meters.

The complete mapping of Mappano required a total of 13 flights (Figure 1b), resulting in the acquisition of 5646 frames at an average altitude of 120 meters, corresponding to a spatial resolution of 3.16 cm/pixel (Beani et al., 2023). During the survey, both nadiral and oblique images were captured, the latter being essential for documenting vertical features (e.g., façades) and enhancing the level of detail in 3D reconstruction of some specific areas of the city.



(a) (b)

Figure 1. Mappano municipality area (top), GCPs distribution (a); overview of the flights carried out (b).

The processing of the photogrammetric data involved Agisoft Metashape and Pix4D, to enable comparative evaluation of the results. Although the workflow followed a largely automated sequence, careful validation of the residual errors associated with the GCPs utilized in the process was carried out. This methodology yielded a mean positional error of 2.15 cm, with horizontal accuracies of approximately 1.52 cm (x) and 1.4 cm (y), and a vertical accuracy of 0.6 cm (z). These accuracy levels enabled the production of high-resolution cartographic outputs suitable for scales up to 1:1000.

The resulting dense point cloud consists of approximately 900 million points, and most of it has been labelled according to the standard LAS classes (Figure 2). Some remaining parts are currently under annotation, and the full dataset will be published on Zenodo.

Figure 2. Mappano's municipality point cloud with an example of the level of detail acquired with oblique images (top right) and the annotation according to the LAS classes.

In addition to the point clouds, shapefiles of the whole municipality have been generated for the public administration according to CityGML level of detail and, in the national framework, the Spatial Data Base Reference for the authorities (BDTRE), which is the geographical database of whole Piedmont Region, whose contents are structured according to the national "Technical rules for the definition of the content specifications of the geo topographic databases" (MD 2011/11/10) and INSPIRE compliant (Beani et al., 2023).

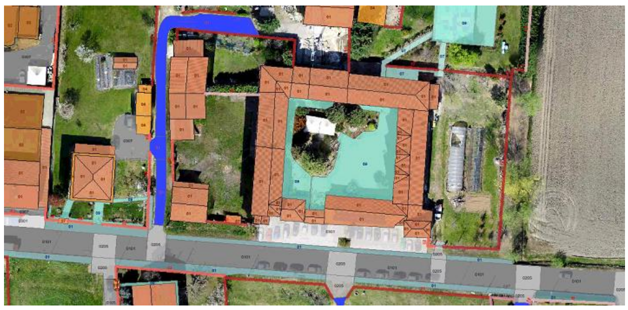
In this context, the shapefiles of the volumetric units (footprints), roofs, and transport networks (Figure 3) will be published in the 3M dataset along with the point clouds, to establish a common ground truth useful for potential automatic mapping purposes.

In detail, within the *urban level*, there are multiple classes. Specifically, as regards the *Building Class* (EDIFC - 020102), the related attributes that have been inserted are:
- the building type (02010201 EDIFC_TY): generic '01', tower building '02', typical building '03', etc;
- the category of use (02010202 EDIFC_USE): residential '01', administrative '02, public service' 03 ', military' 04 ', etc;
- status (02010204 EDIFIC_STAT): under construction '01', ruined '02' or built '03'.

The *Roof level* is connected to the above-mentioned information, specifying the typology (pitch, terrace, canopy, shelter, flat roof…). Within the *Traffic/viability Level*, different classes are identified corresponding to the different types of traffic areas: vehicular (AC_VEI), cycle path (AC_CIC), pedestrian (AC_PED), mixed (AR_VMS), and so on.

(a)

(b)



(c)

Figure 3. Example of ground truth shapefiles. (a) Building class; (b) Roof typology; (c) Viability network with grey colour for the main roads, dark blue for the secondary roads, and light blue for the pedestrian zones.

3.2 ArCH^X

As regards to the ArCH^X dataset[1] (under publishing also on Zenodo), new scenes from Spanish- Portuguese and Indian architectures have been included. This extension will allow researchers to deal with different architectural styles and lexicons, improving the generalization ability of the AI models applied to the CH domain.

3.2.1 Portuguese gates

The Portuguese case studies focused on two masonry buildings, the San Antonio and San Francisco Master Gates (Figure 4),

[1] www.archdataset.polito.it

which were constructed between 1661 and 1676. Both gates are located within the Portuguese Trace Fortress of Almeida and originally played a key role in the stronghold's defensive system. These monumental gates feature a complex architectural design with multiple facades, roofs, mouldings, vaults, and arches. The dataset comprises only the X, Y, and Z coordinates and was acquired using a FARO Focus 120 TLS for both the exterior and interior areas. This method was chosen because certain parts of the structures were difficult to access and there were challenging lighting conditions (strongly shaded and illuminated areas) that limited the use of SfM photogrammetry (Sánchez-Aparicio et al., 2018; Sánchez-Aparicio et al., 2019).

Figure 4. Main façades of the Master Gates of the Almeida Fortress: San Francisco Gate (top) and San Antonio Gate (bottom).

The raw datasets were subsequently filtered to remove outliers and irrelevant areas using statistical outlier removal (SOR) and region-based trimming techniques (Figure 5).

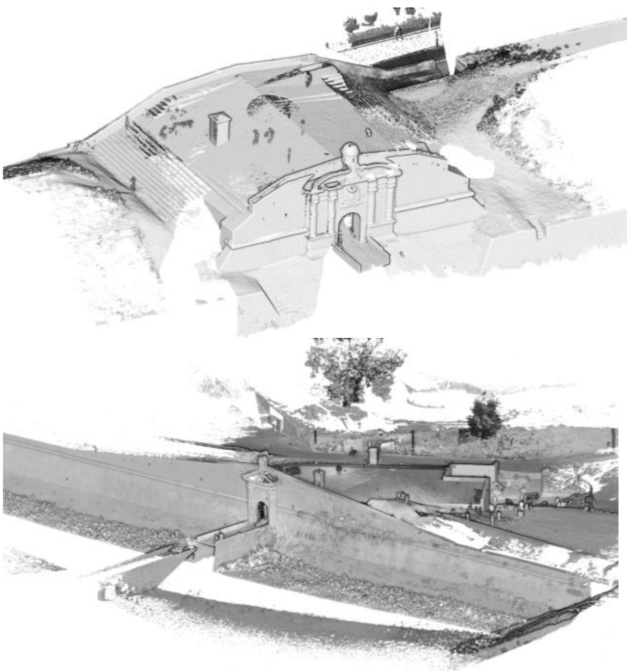


Figure 5. Portuguese 3D point clouds before labelling.

After pre-processing, the 3D point clouds were manually labelled according to the standard classification of ArCH dataset (Figure 6 and 7). The resulting 3D point clouds provide full coverage of both gates at a spatial resolution of 5 mm. There are 45,852,933 points for San Antonio and 48,477,883 points for San Francisco.

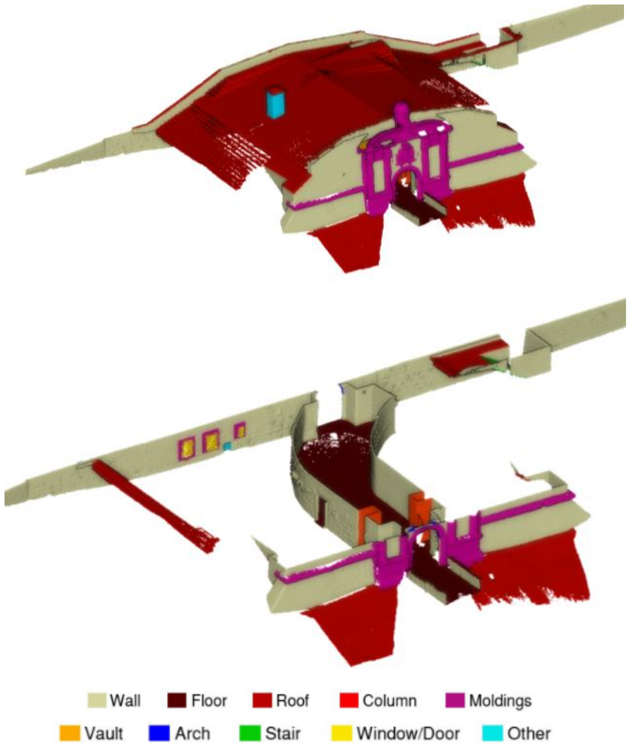


Figure 6. San Francisco Gate: outdoor and indoor section.

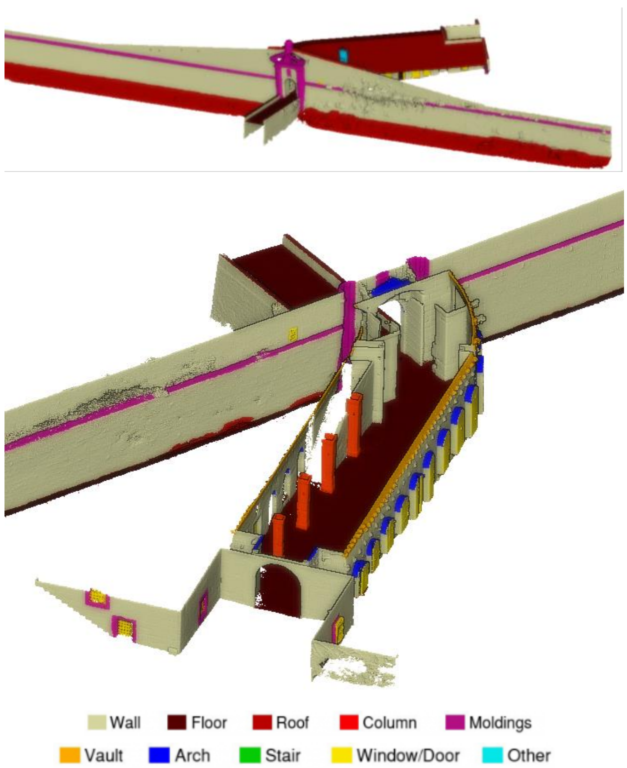


Figure 7. San Antonio Gate: outdoor and indoor section.

3.2.2 Indian temples

The ARCHx dataset also incorporates the 3DITA – 3D Indian Temple Architecture dataset (Sridhar et al., 2025). This dataset comprises point cloud data of Nagara-style temple architectures of India, commonly referred to as the North Indian temple style. India is home to a vast number of Hindu temples, and documenting these cultural heritage structures is crucial for their conservation and preservation. Advanced 3D documentation techniques play a vital role in integrating digital technologies with heritage studies, facilitating applications such as AR/VR-based heritage tourism and immersive storytelling (Basu et al., 2023). However, due to the large number of temples across the country, it is challenging to document each site individually. Therefore, a scientific and scalable approach is required for large-scale documentation.

With this motivation, the 3DITA dataset was developed as India's first point cloud heritage dataset featuring semantically segmented classes. It aims to advance research in automated documentation using AI/ML-based methods. The dataset is structured for training, validation, and testing, serving as a benchmark for evaluating various algorithms in the cultural heritage domain. By enabling consistent benchmarking, this dataset provides significant scientific input for the development of computational tools and applications in digital heritage documentation, conservation, and visualization.

Indian temple architecture can be broadly categorized into Nagara (North Indian style), Vesara (Central Indian style), and Dravida (South Indian style). These classifications, however, are often overlapping due to India's long historical evolution and the diverse influences of various dynasties and regional traditions. Among these styles, the Nagara style has been selected for inclusion in the 3DITA dataset owing to its antiquity, architectural richness, and current vulnerability to aging and environmental degradation. Nagara-style temples are predominantly constructed using locally available stone materials, characterized by strong structural forms and intricate carvings. However, these structures face increasing risks from anthropogenic factors, natural weathering, and disasters. These conditions underscore the need for systematic digital documentation for conservation and research purposes.

Accordingly, the 3DITA dataset focuses on Nagara-style temples located in Uttarakhand, India. The dataset was generated using close-range photogrammetry (CRP) and terrestrial laser scanning (TLS) techniques. It comprises over 325 million points collected from 47 temples across 13 sites. Among these, 46 structures were documented using CRP with 22,370 in-situ photographs, while one structure was recorded through TLS, capturing approximately 13 million points.

The raw data were processed using standard photogrammetry and lidar workflows, followed by integration and refinement in CloudCompare. Manual segmentation was carried out to define semantic classes representing major architectural components. In the absence of predefined IFC (Industry Foundation Classes) categories for heritage temple structures, segmentation was performed based on visual interpretation and architectural features, grouped into five primary classes:

1. **Walls:** Stone boundary structures of the temple premises
2. **Mandapa:** Roofed hall or porch supported by pillars
3. **Shikhara:** Curvilinear spire or tower above the sanctum
4. **Garbhagriha:** Inner sanctum housing the deity
5. **Amalaka:** Circular or ribbed stone element crowning the Shikhara

These five classes were selected to initiate structured research in automated segmentation of Indian temple architectures, with no cultural or religious bias intended. Figure 8 illustrates a visual sample representation of the dataset, displaying the true-color point cloud alongside the manually segmented point cloud with five architectural classes. This 3DITA dataset is publicly available for the research community to explore at https://zenodo.org/records/17295322.

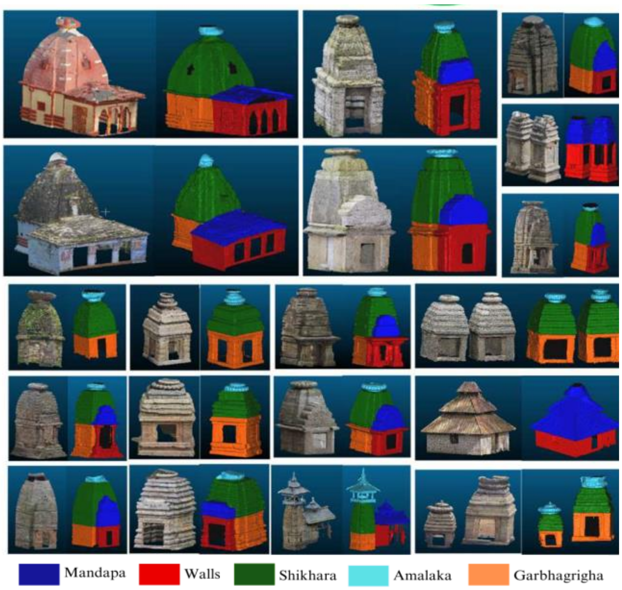


Figure 8. Semantically labelled Indian Nagara Architecture dataset.

3.3 Challenges

This section highlights the primary challenges encountered in creating and harmonizing various architectures and cross-domain datasets, especially for the CH field. A key challenge lies in standardizing semantic classes across diverse architectural traditions. When combining multiple datasets, the inherent inconsistencies in data acquisition methods, metadata structure, and the granularity of classification schemes pose significant difficulties. Each dataset may have been created under different conditions, with distinct documentation objectives and classification hierarchies.

For instance, the ArCH dataset includes detailed architectural classes such as windows/doors, stairs, mouldings, vaults, and arches, which can be considered as Level 2 or fine-grained classifications. In contrast, the 3DITA dataset adopts broader Level 1 classes relevant to Nagara-style Indian temple architecture, which do not typically feature elements such as windows or arches as major architectural components. Hence, direct alignment or merging of class labels between datasets becomes challenging.

Another critical issue is the visual similarity among classes in the 3DITA dataset. Since most Indian temples are constructed from locally available stone, the surface texture and color across different architectural components (e.g., walls, mandapa, shikhara) appear visually uniform in the point cloud. This makes automated segmentation more difficult compared to datasets like ArCH, where color attributes aid machine learning models in distinguishing materials and features.

While preliminary harmonization efforts have been made, for example, aligning semantically similar classes such as mandapa (in 3DITA) with roof (in ArCH), and walls across both datasets, significant differences remain. These challenges underscore the need for robust standardization frameworks to ensure interoperability across global heritage datasets. The present study thus provides a foundation for future work aimed at harmonizing diverse heritage architectures through unified classification and ontology development.

Other challenges arose during the semantic segmentation of the Portuguese fortress gates' construction systems, primarily due to the geometric complexity of certain areas of their 3D point clouds. For example, extensive biological colonization on the masonry surfaces altered the geometric factors related to planarity and verticality, resulting in local misclassifications. Similarly, severe mortar loss between ashlar joints introduces irregularities that distort the walls' geometric continuity. These phenomena directly impact the reliability of the geometric features and consequently affect the overall classification performance.

Additionally, some difficulties arose when aligning certain construction system classes with the standard taxonomy defined in the $ArCH^X$ dataset. For example, the main façades of the fortress gates feature inclined surfaces that would logically correspond to the "slope" category. However, since this class is not included in the $ArCH^X$ standard classification, these elements were geometrically reassigned to the "roof" category to maintain taxonomic consistency (represented in dark red in Figure 6 and 7).

In summary, a few specific challenges can be identified from the integration of these new datasets into $ArCH^x$:

- Challenges related to architectural style: as the dataset expands beyond its original data, the problem of heterogeneity becomes more evident. Due to the organic nature of heritage objects, this is a very difficult point to generalize in an AI context. While some classes may be transferred to another style (e.g. mandapa and roof), they may still possess different intrinsics that remain important for interpretation.
- Challenges related to building materials: similarly to architectural styles, building materials in heritage objects depend strongly on local culture and geography. This poses another challenge since different materials may show different patterns in colorimetric information, thus rendering data standardization more difficult.
- Challenges related to technical issues: in many cases, the existing datasets consist of diverse point clouds taken for different purposes and with different conditions. This also presents an additional problem to the harmonization process of the existing datasets.

4. Discussions and conclusions

Benchmarking algorithms for semantic segmentation both in the urban and CH domains plays a crucial role in advancing computational methods for monitoring and managing the city, as well as documenting, preserving, and conserving heritage structures.

From the urban point of view, the possibility of having a unique dataset for an entire municipality is certainly challenging for data acquisition, processing, labelling and management. Still, at the same time, it allows new ways of exploiting this kind of data, e.g. for telecommunication purposes, gaming, virtual reality and monitoring objectives. The opportunity of having also a ground truth in the form of shapefiles supports all the automatic mapping tasks and activities, to improve these approaches.

On the other side, for the architectural heritage point clouds, unifying and benchmarking large-scale datasets that represent diverse architectural typologies pose both technical and conceptual challenges. Integrating such datasets yields several research benefits, most notably, it enables the development of generalized segmentation algorithms with improved robustness and accuracy. These advancements, in turn, facilitate scalable solutions for automated heritage documentation, aiding heritage management, digital archiving, and immersive visualization.

Nevertheless, it is important to emphasize that AI/ML algorithms rely heavily on the quality of training data. As the adage “garbage in, garbage out” suggests, the effectiveness of these algorithms depends on providing reliable, well-structured, and semantically meaningful datasets. Therefore, researchers must carefully curate and preprocess datasets before integration to ensure data consistency and contextual accuracy.

The integration of datasets like 3DITA and ArCH into a new extended solution represents an initial step toward domain-level generalization within CH. Despite existing variability among heritage structures, owing to differences in material, geometry, and regional architectural traditions, this study establishes a baseline framework for collaborative dataset development and benchmarking. Future work will focus on addressing these challenges by refining class harmonization, improving segmentation accuracy, and expanding the dataset coverage to include a wider range of global heritage typologies.

Acknowledgements

Part of this study was carried out within the FAIR - Future Artificial Intelligence Research and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.3 – D.D. 1555 11/10/2022, PE00000013). This manuscript reflects only the authors’ views and opinions, neither the European Union nor the European Commission can be considered responsible for them.
The authors also acknowledge the Mappano Municipality, which allowed the publication of the dataset.

References

Basu, A., Paul, S., Ghosh, S., Das, S., Chanda, B., Bhagvati, C., Snasel, V., 2023. Digital Restoration of Cultural Heritage With Data-Driven Computing: A Survey. *IEEE Access* 11, 53939–53977. https://doi.org/10.1109/ACCESS.2023.3280639.

Beani, E., Colucci, E., La Riccia, L., Lingua, A. M., Maschio, P. F., Matrone, F., ... & Traore, A., 2023. The 3D metric survey for the digital cartographic production to support the knowledge of the new municipality of Mappano. *Post Un-Lock: From Territorial Vulnerabilities to Local Resilience* (pp. 153-168). Cham: Springer International Publishing.

Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C. and Gall, J., 2019. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In Proceedings of the *IEEE/CVF international conference on computer vision*, pp. 9297-9307.

Croce, V., Caroti, G., Piemonte, A., De Luca, L. and Véron, P., 2023. H-BIM and artificial intelligence: classification of architectural heritage for semi-automatic scan-to-BIM reconstruction. *Sensors,* 23(5), p.2497. doi.org/10.3390/s23052497

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L. and Bennamoun, M., 2020. Deep learning for 3D point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(12), pp.4338-4364. doi.org/10.48550/arXiv.1912.12033

Griffiths, D. and Boehm, J., 2019. SynthCity: A large scale synthetic point cloud. arXiv preprint arXiv:1907.04758.

Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K. and Pollefeys, M., 2017. Semantic3d. net: A new large-scale point cloud classification benchmark. arXiv preprint arXiv:1704.03847.

Han, X., Liu, C., Zhou, Y., Tan, K., Dong, Z., & Yang, B., 2024. WHU-Urban3D: An urban scene LiDAR point cloud dataset for semantic instance segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *209*, pp. 500-513.

Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E.S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A. and Landes, T., 2020. A benchmark for large-scale heritage point cloud semantic segmentation. *Int. arch. photogramm. remote sens. spat. inf. sci.*., 43, pp.1419-1426. doi.org/10.5194/isprs-archives-XLIII-B2-2020-1419-2020.

Roynard, X., Deschaud, J. E., & Goulette, F., 2018. Paris-Lille-3D: A point cloud dataset for urban scene segmentation and classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2027-2030.

Sánchez-Aparicio, L. J., Del Pozo, S., Ramos, L. F., Arce, A., & Fernandes, F. M., 2018. Heritage site preservation with combined radiometric and geometric analysis of TLS data. *Automation in Construction*, 85, pp.24–39. doi.org/10.1016/J.AUTCON.2017.09.023

Sánchez-Aparicio, L. J., Rodriguez-Gonzalvez, P., Gonzalez-Aguilera, D., & Ramos, L. F., 2019. Evaluating The Structural Integrity Of The Saint Antonio Barrel Vault In The Fortress Of Almeida By Combining Laser Scanner And Limit Analysis. . *Int. arch. photogramm. remote sens. spat. inf. sci.*, XLII-2-W9(2/W9), pp.679–684. doi.org/10.5194/ISPRS-ARCHIVES-XLII-2-W9-679-2019.

Sridhar, M., Paygude, A., Pande, H., Tiwari, P.S., 2025. 3DITA – A 3D Benchmark Dataset for Nagara-Style Indian Temple Architecture: India's First Point Cloud Dataset for Semantic Segmentation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*. XLVIII (48, 1435–1441. https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1435-2025, 2025.

Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., ... & Li, J., 2020. Toronto-3D: A large-scale mobile LiDAR dataset for semantic segmentation of urban roadways. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 202-203.

Wysocki, O., Tan, Y., Froech, T., Xia, Y., Wysocki, M., Hoegner, L., ... & Holst, C., 2025, February. Zaha: Introducing the level of facade generalization and the large-scale point cloud facade semantic segmentation benchmark dataset. In *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 7648-7658). IEEE.