

Collection of 3D data useful for surveying and digitizing cultural heritage using UAV photogrammetry and lidar sensors: Iuvanum 3D dataset^{*}

Massimiliano Pepe^{1,*†}, Donato Palumbo^{1,†} and Ahmed Dewedar^{1,†}

¹ Department of Engineering and Geology, “G. d’Annunzio” University of Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy.

Abstract

The digitization of Cultural Heritage (CH) through 3D data acquisition has become a key tool in archaeological surveying and modeling, offering accurate documentation, preservation, and analysis of fragile artifacts and sites. In recent years, increasingly advanced technologies and methods have enabled high-quality digitization of cultural heritage, including Unmanned Aerial Vehicle (UAV) photogrammetry, Terrestrial Laser Scanning (TLS), and Simultaneous Localization and Mapping (SLAM). In this study, we present a multi-sensor dataset collected at the archaeological site of Iuvanum (Italy), focusing on the ancient theater. The survey combined UAV-based photogrammetry, SLAM, TLS, and Matterport Pro3 acquisition, supported by Ground Control Points (GCPs) measured with Global Navigation Satellite System (GNSS) to ensure precise georeferencing. The shared dataset comprises raw UAV images, GCP coordinates, and processed 3D models, all openly accessible on the Zenodo platform in accordance with FAIR principles (Findable, Accessible, Interoperable, and Reusable) to support replication and reuse. The creation and open sharing of such datasets are of growing importance and relevance today, as they significantly enhance research and education. For research, openly accessible datasets foster transparency, reproducibility, and the development of new methodologies in 3D reconstruction, semantic classification, and digital preservation. For education and training, they provide authentic materials for students and professionals to explore real-world archaeological contexts, experiment with digital tools, and engage with cultural heritage through immersive and data-driven learning experiences. Furthermore, these datasets represent a valuable resource for virtual reconstruction and visualization of ancient archaeological sites, particularly when integrated with AI-based models and algorithms that enable automated interpretation, completion, and restoration of incomplete structures.

Keywords

UAV, TLS, SLAM, Point cloud, Archaeological Site, Cultural heritage, Dataset, Zenodo

1. Introduction

The digitization of cultural heritage has become a fundamental strategy for preserving, analyzing, and disseminating the tangible evidence of our shared past. Among the most significant advancements in this field is the creation of high-fidelity 3D models, which allow for precise documentation and immersive visualization of historical sites, monuments, and artifacts [1]. By exploiting modern acquisition technologies (such as photogrammetry, terrestrial laser scanning, UAV-based imagery, and Global Navigation Satellite System (GNSS) sensors) researchers can now capture cultural assets with an unprecedented level of detail and realism. However, the true value of these digital reconstructions extends beyond their technical precision. Their long-term impact depends on accessibility, reusability, and interoperability. Without well-defined strategies for data sharing, standardization, and preservation, even the most advanced 3D models risk becoming obsolete or inaccessible.

^{*} VIPERC2025: 4th International Conference on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding, 3-4 December 2025

^{*} Corresponding author.

[†] These authors contributed equally.

✉ massimiliano.pepe@unich.it (M. Pepe); donato.palumbo@alumni.unich.it (D. Palumbo); ahmedkamalhamed.dewedar@phd.unich.it (A.K.H. Dewedar)

ORCID 0000-0003-2508-5066 (M. Pepe); 0009-0000-3023-8608 (D. Palumbo); 0000-0002-8521-0693 (A.K.H. Dewedar)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In this context, aligning the production and dissemination of 3D datasets with the FAIR principles (Findable, Accessible, Interoperable, and Reusable) is not only a technical necessity but also an ethical commitment within the digital heritage community.

Cultural heritage faces growing threats from natural hazards, climate change, and human-induced disasters. Its protection is essential not only for safeguarding physical artifacts and cultural values but also for supporting cultural tourism and transmitting heritage to future generations. Because archaeological excavation is inherently destructive, accurate and timely documentation plays a crucial role in ensuring that information about heritage assets endures. Heritage documentation thus involves both tangible and intangible dimensions, emphasizing data capture, monitoring, and long-term preservation. International organizations such as ICOMOS promote best practices and standards in this field, while the European Union has increasingly prioritized cultural heritage digitization as part of broader sustainability and cultural policy goals [2].

Since the 1950s, digital tools have progressively transformed archaeological practice—from early text-based data management and archiving to today’s advanced technologies, including 3D modeling, drone imaging, and GIS. This evolution has fostered interdisciplinary collaboration, integrating expertise from computer science, engineering, and visual arts. Over the past two decades, methods such as photogrammetry, laser scanning, and unmanned aerial vehicles (UAVs) have revolutionized heritage documentation, enabling the creation of detailed, accurate, and increasingly user-friendly 3D models of sites, monuments, and artifacts [3, 4, 5]. Portable solutions such as the Artec Space Spider, Artec Leo, and Creaform HandySCAN 3D allow for high-precision acquisition of small and medium-sized objects, achieving accuracies down to 0.05 mm. Within the ITSERR Resilience project, Diara [6] explored the use of the Artec Space Spider (a portable structured-light 3D scanner) for the semantic extraction of micrometric features from Near Eastern archaeological artifacts. More recently, innovations such as Simultaneous Localization and Mapping (SLAM)-enabled mobile LiDAR systems, UAV-mounted sensors, and even consumer-grade devices like the Apple iPad Pro LiDAR have made 3D capture faster, more flexible, and more affordable, broadening access beyond specialized laboratories [7, 8].

In this framework, the preservation and study of cultural heritage increasingly depend on the ability to collect, integrate, and share high-quality digital data. The aim of this paper is therefore to present a collection of 3D datasets produced through diverse acquisition techniques and state-of-the-art instruments, applied across a range of cultural heritage contexts. These datasets are shared in accordance with the FAIR principles, ensuring their discoverability, openness, and usefulness for scholars, conservators, educators, and the wider public. Through this work, we aim to contribute to the expanding ecosystem of openly available 3D cultural heritage data, while addressing key challenges related to metadata standardization, file formats, licensing, and long-term digital preservation.

2. Value of the Data

The dataset, focused on the cultural heritage environment, enables the analysis of a structure of significant historical and architectural value, while also serving as a resource for both the scientific community and for the broader dissemination of knowledge about sites of particular historical, architectural, and archaeological interest. The dataset is useful for experts in geomatics, civil engineering, computer vision, archaeology, and architecture. For example, it supports tasks such as testing 3D reconstruction methods, processing point clouds, and creating virtual or augmented reality experiences. It also helps develop new methods and tools for building and classifying 3D point clouds, especially for use in cultural heritage projects. The 3D models derived from the dataset can support restoration planning by providing accurate measurements and visual references of structural conditions. These models can also be used in educational and museum settings to create interactive and immersive experiences for the public, enhancing understanding and engagement with cultural heritage. The value of shared datasets lies in their capacity to enable reuse, reduce redundant fieldwork, foster transparency, and ensure long-term preservation and accessibility. In the field of

cultural heritage, datasets extend beyond technical documentation, serving as vital tools for preservation, research, and public engagement. They safeguard fragile sites and artifacts, support restoration and conservation planning, enable interdisciplinary research, and foster innovation in 3D reconstruction and visualization. Moreover, they enhance accessibility and education through open dissemination and virtual experiences, bridging scholarly work and public understanding. Table 1 summarizes key contributions across different domains.

Table 1

Overview of recently introduced datasets

Dataset	Authors (Year)	Data Type	Method	Applications
ESPADA	Lopez-Campos & Martinez-Carranza (2021) [9]	RGB + depth image pairs	AirSim + photogrammetric	Training DNNs for single-image depth estimation, tested with ORB-SLAM
STPLS3D	Chen et al. (2022) [10]	Point clouds	UAV flight simulation	Large-scale 3D data training and evaluation
Temple of Hera	Pepe et al. (2022) [11]	Images + dense point clouds + GCPs	UAV photogrammetry	Cultural heritage documentation, 3D analysis
Historic Buildings dataset	Pellis et al. (2025) [12]	Images + point clouds	photogrammetry	Semantic segmentation, scan-to-BIM automation
Beam dataset	Janiszewski et al. (2024) [13]	Point clouds, meshes, textures	SfM/MVS (lab conditions)	Structural simulation, numerical modelling, ML training, validation

The added value of our dataset lies in its integration of UAV photogrammetry, Terrestrial Laser Scanning (TLS), SLAM, Matterport Pro3, and GNSS-NRTK (Network Real Time kinematic) within a single georeferenced framework. This multi-sensor approach combines the wide coverage and rapid acquisition of UAVs, the high-resolution detail of TLS, the operational flexibility of SLAM in complex environments, and the geodetic accuracy ensured by GNSS control.

3. Experimental Design, Materials and Methods

3.1 Method

The objective of this research is to produce and share high-quality 3D datasets of cultural heritage assets for the scientific community, in full alignment with FAIR principles. The adopted methodology involves several stages, from on-site data acquisition to model construction, quality evaluation, and open dissemination through international platforms. Data collection was conducted using a combination of laser scanner and digital photogrammetry. All surveys followed metric control protocols using calibrated targets and Ground Control Points (GCPs) to ensure the geometric accuracy of the captured data.

The raw data were processed following a multi-step workflow designed to ensure both accuracy and usability. The process began with the registration and merging of point clouds, aligning the individual scans into a coherent dataset. When necessary, noise reduction and topological optimization were applied to refine the models and enhance their structural consistency. Finally, the

completed models were exported in open and interoperable formats, such as (.las), to guarantee broad accessibility and compatibility with different platforms. Each 3D model was carefully assessed to ensure both geometric precision and visual fidelity. All datasets and models were georeferenced to the Universal Transverse Mercator (UTM) coordinate system (RDN2008/UTM zone 33N N - EPSG:6708). Using GNSS with NRTK, TLS and UAV datasets were directly registered to the control network, while the SLAM model required an additional transformation for alignment. In addition, a visual inspection was carried out to confirm the coherence of textures and the overall photorealistic quality of the renderings.

The 3D models and their metadata were published on open-access platforms using Zenodo web platform which is an open-access repository developed by CERN and supported by the European Commission, enabling researchers to share, cite, and preserve datasets and research outputs. It provides persistent DOIs and ensures FAIR data principles compliance for long-term accessibility and reuse.

3.2. Case study

The case study focuses on the Theater located in the archaeological site of Iuvanum, Italy (Latitude: 41°59'50.78" N, Longitude: 14°14'57.83" E) as shown in Figure 1 (a). The site, inhabited since the Late Bronze Age, is home to two Samnite temples built around the 2nd century BC. The sacred altar, surrounded by polygonal walls, dominates the plain below and the forum, the nerve center of the more recent Roman municipality. The theater, also of Samnite design, is located beneath the temple area. The first seven rows of steps and the semicircular orchestra remain intact. The flooring, like the rest of the structure, was made with exquisitely crafted slabs of local stone (Figure 1b).

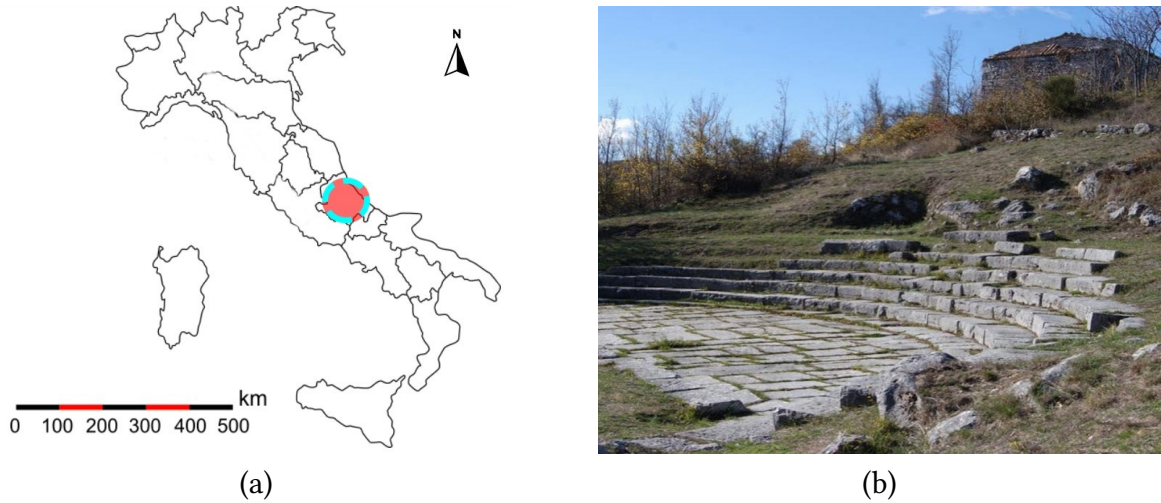


Figure 1: Study area: location of Iuvanum archaeological site (a) and panoramic image(b).

3.3. 3D dataset

3.3.1. Devices used for the experimentation

This study integrates multiple datasets acquired using diverse survey techniques and instruments to ensure accurate and comprehensive 3D documentation of the site. The following devices were employed:

- UAV photogrammetry;
- SLAM;
- Matterport Pro 3 scanner;
- GNSS;
- Faro TLS;
- Stonex X100 TLS.

Figure 2 illustrates the devices employed to collect the dataset, highlighting the range of technologies used, including terrestrial laser scanners, UAVs, SLAM-based mobile scanner, TLS Scanners, and GNSS, each contributing complementary capabilities to ensure comprehensive coverage, high accuracy, and detailed 3D documentation of the site.



Figure 2: Devices used for dataset acquisition

Figure 3 shows the datasets collected during the survey campaign and available at the following DOI: <https://doi.org/10.5281/zenodo.17431997>.

Name	Size	Download
GNSS Data.rar md5: dca1b9390172e8a6d78df380dc82ead	2.1 kB	Download
Model From UAV.rar md5: 2a8bf7079410871b613cab61a115117	1.3 GB	Download
SLAM Model.rar md5: cb4924c1face21cb5baf4322d862078	42.0 MB	Download
TLS-FARO Model.rar md5: 21f0ebf49dfe3408d773cb952fc89859	2.0 GB	Download
TLS-STONEX Model.rar md5: c46bbec87c4548654a9b96b88b2bd0	1.3 GB	Download
UAV_Images.rar md5: a718df72529c0b131526fbbdaec0f1a	6.4 GB	Download

Details

DOI
DOI: 10.5281/zenodo.17431997

Resource type
Conference paper

Publisher
Zenodo

Conference
Pattern Extraction and Recognition for Cultural Heritage Understanding (VIPERC 2025) (VIPERC) , Pescara, 3-4 December 2025

Rights

License
Creative Commons Attribution 4.0 International

Figure 3: Overview of the datasets available on Zenodo website.

3.3.2. Dataset From GNSS-NRTK

The topographic survey of the Roman theatre at Iuvanum was carried out using GNSS technology in NRTK mode, with corrections provided by the regional permanent stations network of Abruzzo. The primary objective was the accurate determination of the coordinates of a set of GCPs, which served as reference markers for the subsequent photogrammetric processing, as well as of several characteristic points on the architectural structures of the theatre. In particular, a total of 63 points were surveyed. Observations were performed with an EMLID Reach RS2+ dual-frequency GNSS receiver, capable of tracking multiple constellations (GPS, GLONASS, Galileo, BeiDou). Connection

to the NTRIP server of the regional network enabled real-time reception of differential corrections, ensuring centimeter-level accuracy and direct integration into the official Italian geodetic reference frame. The GCPs were carefully selected and materialized in stable, well-distributed locations around the theater, while characteristic points were surveyed directly on architecturally significant elements of the monument. Each point includes geographic coordinates, positional accuracy metrics, baseline distances, and satellite visibility across multiple constellations (GPS, GLONASS, Galileo, BeiDou, QZSS). RTK corrections from the Abruzzo Region (Italy) ensured high positional reliability. These GCPs provided a unified geodetic framework for precise alignment and integration of UAV photogrammetry, TLS, SLAM, and Matterport datasets.

3.3.3. Dataset From UAV

For image acquisition, a DJI Mini 4 Pro UAV (DJI Company, Shenzhen, China) was employed. The DJI Mini 4 Pro is a compact quadcopter equipped with an advanced 1/1.3" CMOS sensor capable of capturing 48 MP still images and recording 4K HDR video, providing high-quality data suitable for photogrammetric processing. Prior to the aerial survey, GCPs made of durable wooden material, marked with black-and-white checkerboard or rhombus patterns, and sized for easy identification in the imagery, were uniformly distributed across the study area. Flight planning was conducted using the DJI Fly application, following a grid flight path with 80% forward overlap and 60% side lap. Additionally, 45° oblique images with 60% overlap were acquired to ensure full coverage of vertical and complex surfaces within the surveyed site. All collected images were consolidated into a single folder as shown in Figure 4 within the dataset directory to streamline their management and ensure compatibility with photogrammetry software.



Figure 4: The collected UAV images organized in the dataset folder.

This workflow ensures efficient processing, allowing Agisoft Metashape to generate a complete 3D model while preserving spatial accuracy and georeferencing. Indeed, Agisoft Metashape is a professional photogrammetry software designed to process digital images and generate high-precision spatial data. It relies on Structure-from-Motion (SfM) and Multi-View Stereo (MVS) algorithms to reconstruct 3D models, dense point clouds, textured meshes and orthomosaics. Figure 5 (a) shows the camera poses of the images while Figure 5 (b) shows the orthomosaic realized on the case study with the visualization of the GCPs.

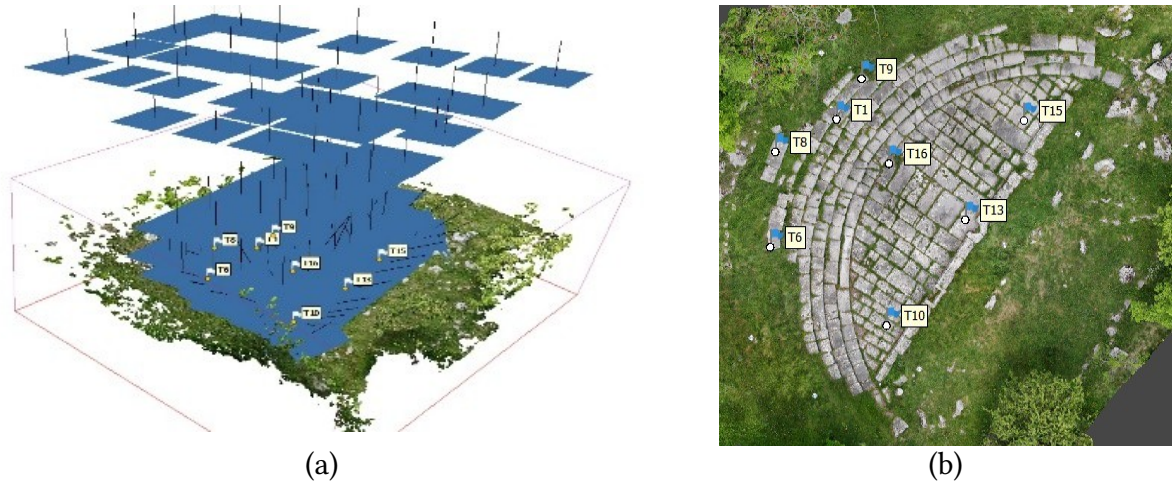


Figure 5: Processing of UAV images: camera poses (a) and orthomosaic (b).

To further enhance accessibility and dissemination, the georeferenced 3D model can also be published on Sketchfab, offering an interactive environment for visualization and annotation on the several part of the structures [14].

3.3.4. Dataset From SLAM

For terrestrial 3D data acquisition, an X120GO SLAM Scanner 3D (Stonex, Italy) was employed. The X120GO is a portable, high-speed mobile laser scanner based on technology, capable of capturing up to 320,000 points per second with a range of up to 100 m. Its 360° rotating LiDAR sensor allows efficient and accurate mapping of both indoor and outdoor environments without the need for GPS. The device integrates a high-resolution panoramic camera to complement point cloud data with colour information, enhancing visual detail and interpretation. Data collection was carried out by walking through the survey area along predefined paths to ensure complete coverage of all structures and surfaces. Post-processing and point cloud generation were performed using Stonex's dedicated software suite, enabling alignment, cleaning, and preparation of the dataset for further analysis. Figure 6 shows the path of turn during the survey of the Theater using SLAM device.

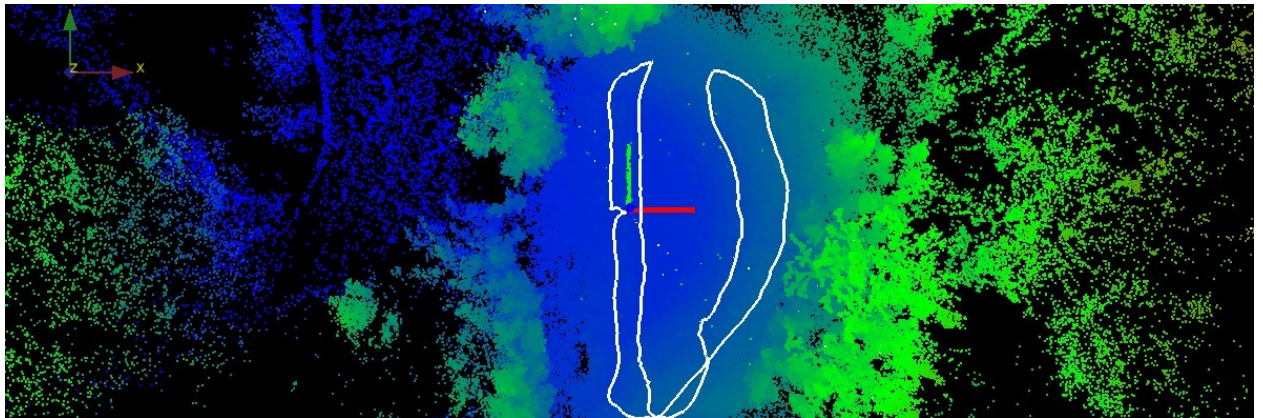


Figure 6: Survey path of the Roman Theatre using the SLAM device.

3.3.5. Dataset From TLS (FARO)

For terrestrial 3D data acquisition, a FARO Focus M70 Laser Scanner (FARO Technologies, USA) was employed. The Focus M70 is a compact and lightweight phase-based terrestrial laser scanner capable of capturing up to 976,000 points per second with a maximum range of 70 m. It provides millimetre-level accuracy and integrates a built-in HDR camera to generate colourized point clouds. Scanning was performed from multiple fixed positions to ensure full coverage of the site, with overlapping scans enabling accurate registration.

The acquired data were processed using FARO SCENE software, which supports point cloud alignment, noise filtering, and export for further analysis. The survey was conducted using a Faro TLS, which enabled the acquisition of high-density point clouds with sub-centimeter accuracy.

The scanning process is illustrated in Figure 7a, while the resulting 3D point cloud is shown in front view (Figure 7b) and top view (Figure 7c), providing a detailed representation of the site's geometry.



Figure 7: Scanning the site using a Faro TLS (a); 3D point cloud view in front view (b) and in top view (c).

3.3.6. Dataset From TLS (STONEX - X100)

For terrestrial 3D data acquisition, a Stonex X100 TLS (Stonex, Italy) was employed. The main features and differences of the two TLS considered are summarized in Table 2.

Table 2

Features and differences of the two terrestrial laser scanners (TLS) considered.

Feature	Stonex X100 TLS	FARO Focus M 70
Weight	~3.2 kg (light and easily portable)	~4.2 kg (still portable, but sturdier)
Range	Up to 120 m (claimed)	Up to 70 m reliable range
Accuracy @ 10 m	~6 mm	±3 mm
Scan Speed	320,000 pts/s	Up to 488,000 pts/s
Camera Integration	Insta 360 X5- HDR panoramic camera with a resolution of 72 MP (11,968 × 5,984)	High-resolution 165 MP HDR camera
Environmental Specs	Self-leveling ±5°, app control, lightweight	IP54, wide temperature tolerance, many sensors (GNSS, compass, barometer)

Scanning was conducted from multiple station points to ensure full coverage of the survey area. The captured data were processed using Stonex proprietary software for point cloud registration, alignment, noise reduction, and preparation for subsequent analysis.

Figure 8a captures the device in action while scanning, while Figure 8b shows the 3D point cloud model extracted from the Stonex device.



Figure 8: Acquisition by Stonex X100: picture of the device during the scans (a) and 3D points cloud (b).

3.3.7. Dataset From Matterport

For image and spatial data acquisition, the Matterport Pro3 3D camera (Matterport Inc., Sunnyvale, USA) was employed. The Pro3 is a portable LiDAR-based device designed for rapid indoor and outdoor scanning, capable of capturing high-resolution panoramic imagery and generating precise 3D point clouds with an effective range of up to 100 meters. Its rotating LiDAR sensor, combined with a 20 MP RGB camera, enables the simultaneous acquisition of colorized point clouds and immersive spherical images, facilitating the creation of photorealistic digital twins. Data acquired with the Pro3 were automatically aligned in the Matterport Capture app, then exported for further integration with TLS, SLAM, and UAV photogrammetric datasets, contributing to a unified 3D reconstruction of the theatre. The Matterport device was employed on-site to perform the scanning (Figure 9a). The acquired data were subsequently processed to generate the model layout after 15 scans (Figure 9b), and the final 3D model was exported from the Matterport software, allowing for measurements and further data analysis (Figure 9c).

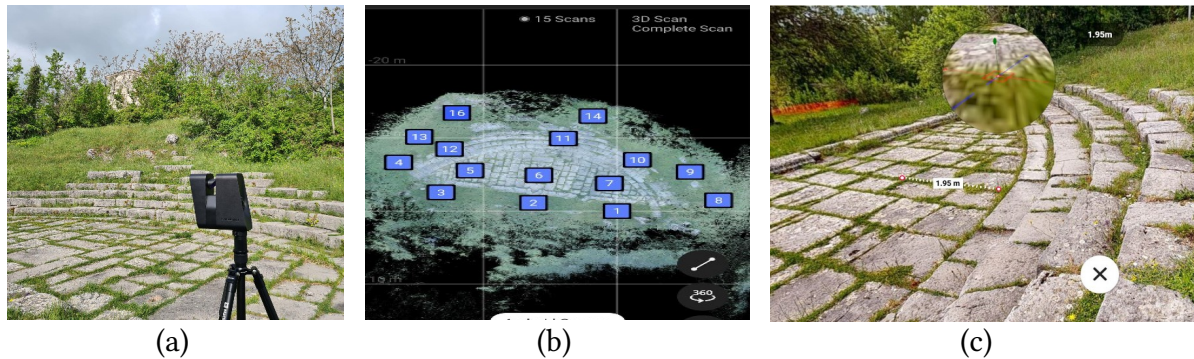


Figure 9: Data from Matterport: scanning on the site (a), layout output of the 15 scans (b) and model exported from Matterport software with measurements between two points (c).

The data collected by the Matterport Pro3 are processed in the Matterport Cloud platform, where scans are automatically aligned and converted into a navigable 3D environment. This system generates the virtual tour, linking panoramic viewpoints into an interactive walkthrough accessible via web browser or mobile application (project link: <https://my.matterport.com/show/?m=AYqN6gQToh1>).

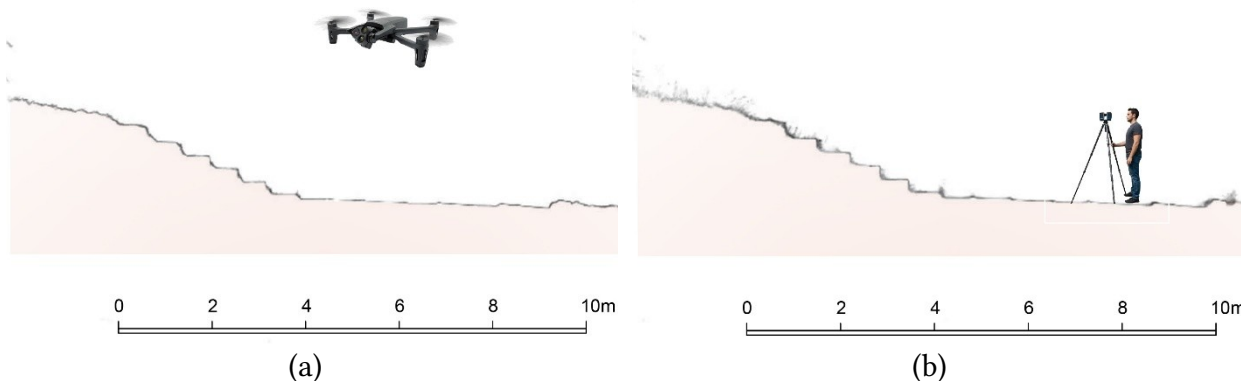
The digital twin is stored online within the user's account and can be shared publicly or embedded in external platforms. While the online viewer provides immersive access and basic measurement tools, downloading raw data is subject to restrictions. Matterport offers export packages such as the MatterPak Bundle (.XYZ point cloud, textured meshes in .OBJ/.MTL, and floor plan images) and the E57 export with full-resolution LiDAR data, both requiring an active subscription and additional fees.

4. Discussion

The integration of UAV photogrammetry, TLS, SLAM, and Matterport scanning, supported by GNSS-NRTK control, proved highly effective for documenting the Roman theatre at Iuvanum. Each technique contributed complementary strengths, and their combination overcame the limitations of single approaches, resulting in a spatially coherent and geometrically precise dataset. GNSS-NRTK ensured centimeter-level accuracy (RMS: 0.010–0.020 m), providing a robust geodetic framework for aligning heterogeneous data. The point cloud analysis highlighted clear differences in density, coverage, and application. TLS (FARO M70, Stonex X100) achieved the highest density (up to 120,549,859 points per scan), making it ideal for detailed documentation. UAV photogrammetry generated 79,813,940 points with broad coverage but lower local density. SLAM (Stonex X120GO) recorded 3,487,698 points along its path, offering mobility and speed but requiring drift correction. The Matterport Pro3, with a range accuracy of about 2 cm and adjustable density, provided a rapid and accessible solution for visualization, though with lower precision compared to TLS. The integration of these datasets produced a scalable and comprehensive digital record of the site. Beyond documentation, they support conservation planning, digital reconstruction, and machine learning applications in point cloud analysis.

The comparison showed that TLS with the FARO device delivered the highest-resolution and most reliable 3D reconstruction, confirming its suitability for detailed large-scale surveys. In contrast, SLAM produced sparse and low-quality results when based on a single site traversal, emphasizing its dependence on survey design and the need for multi-pass or complementary data. UAV photogrammetry proved to be a valuable alternative, offering high-resolution results and broad area coverage, though with lower geometric precision than FARO TLS. Its efficiency and practicality in challenging or resource-limited contexts make it an effective complement to TLS.

The datasets were georeferenced within a unified reference system and as a result, it is also possible to perform profiling along the theater stairs in order to better identify the qualities of the point cloud acquired with different sensors and techniques. In particular, the point clouds were imported into PointCab software, which enables the management of 3D point clouds and their conversion into raster representations for comparative analysis. Four-point clouds (from TLS Faro, TLS Stonex, SLAM, and UAV) were imported in E57 format and compared by extracting a common 10 cm thick section, allowing the generation of four directly comparable profiles. The analysis highlights the distinct strengths of each sensor. Faro and Stonex TLS provided high-density point clouds with sub-centimeter precision, clearly capturing structural details. SLAM produced a less dense cloud but still delivered reliable representations of the theater's main geometries. UAV data offered abundant points and effective surface description, but with higher noise that limited interpretability. Overall, the integration of these datasets produced a coherent and robust information model, where the limitations of each method were offset by the strengths of the others, resulting in comprehensive documentation of the Iuvanum theater. Figure 10 shows the details of section profiles for the four analyzed models.



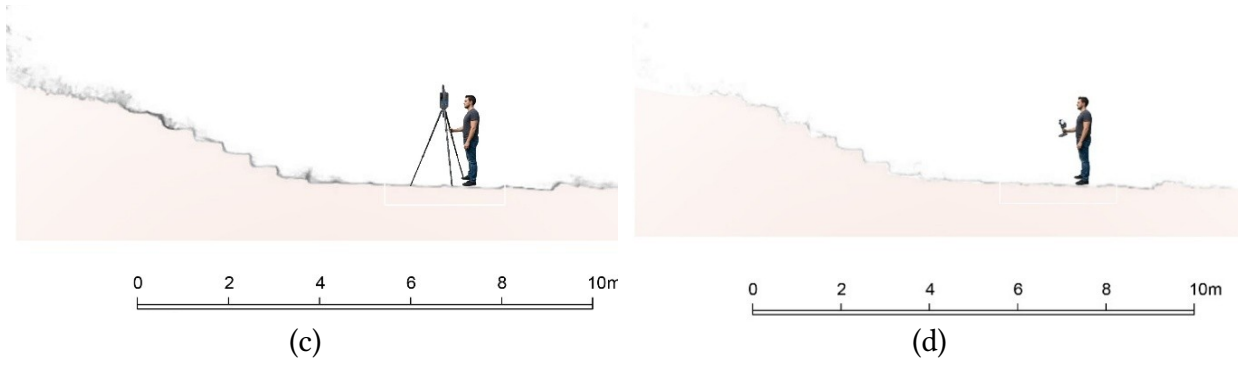


Figure 10: Section profiles from the four analyzed models: UAV model (a), TLS FARO model (b), TLS Stonex (c) and SLAM model (d).

The profile comparison confirms the complementary strengths of TLS, SLAM, and UAV photogrammetry for documenting the Iuvanum theater. TLS delivered the highest precision and structural detail, SLAM provided reliable geometries with efficient acquisition, and UAV data ensured extensive coverage despite higher noise. Integrated within a unified reference system, these datasets produced a coherent and robust model in which individual limitations were offset, enabling comprehensive and accurate documentation of the site.

5. Conclusions

This paper presents a multi-sensor dataset of the Roman theatre at Iuvanum, generated through UAV photogrammetry, TLS, SLAM, and Matterport scanning, all georeferenced with GNSS-NRTK. The integration of these techniques ensured complementary strengths, producing a spatially coherent, geometrically accurate, and visually detailed 3D model with centimeter-level precision. The open-access dataset, shared in interoperable formats and aligned with FAIR principles, supports archaeological documentation, conservation planning, digital reconstruction, and machine learning research. Remaining challenges include optimizing point cloud density, computational efficiency, and standardized metadata, but the Iuvanum dataset provides a valuable benchmark for advancing cultural heritage digitization and preservation. This approach is in line with Niccolucci & Richards (2021) [15] and Vincent et al. (2022) [16] for the creation of open and well-documented datasets which are now a key element to ensure compliance with the FAIR principles and to foster a more transparent, collaborative, and sustainable science of cultural heritage. The datasets acquired on the archaeological context of Iuvanum are not only documentation tools, but also constitute a versatile database for multiple applications: from multi-temporal modeling and analysis to the generation of HBIM models and automatic classification using machine learning techniques. Sharing these datasets with the scientific community promotes interdisciplinary approaches and open science practices, ensuring new possibilities for research, conservation, and enhancement of cultural heritage.

Acknowledgements

The study presented in this article was funded by the Project Lip3D: Living forever the Past through a 3Digital world. Digital Europe Programme (DIGITAL), Cloud Data - AI 05 Cultheritage. Data Space for Cultural Heritage.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] P. Klapa, M. Żygadło, M. Pepe, 3D heritage reconstruction through HBIM and multi-source data fusion: geometric change analysis across decades, *Applied Sciences* 15 (16) (2025) 8929.
- [2] P. Svensson, D. T. Goldberg (Eds.), *Between Humanities and the Digital*, The MIT Press, Cambridge, MA, USA; London, UK, 2015. doi:10.7551/mitpress/9465.001.0001.
- [3] D. Moullou, D. Mavromati, Topographic and photogrammetric recording of the Acropolis of Athens, in: *Proceedings of the XXI International CIPA Symposium*, Athens, Greece, 1–6 October 2007, pp. 515–520.
- [4] D. Moullou, R. Vital, S. Sylaiou, L. Ragia, Digital tools for data acquisition and heritage management in archaeology and their impact on archaeological practices, *Heritage* 7 (2024) 107–121.
- [5] M. Pirker, S. Haun, J. Schneider, Evaluating drone-based photogrammetry for morphologic mapping of a hydraulic model with a mobile bed, in: *Proceedings of the EGU General Assembly 2025*, Vienna, Austria, 27 April–2 May 2025, EGU25-20898. doi:10.5194/egusphere-egu25-20898.
- [6] F. Diara, Data extraction from 3D scanning: post-processing filtering for analytic and informative models of small archaeological finds, *Archeologia e Calcolatori* 35 (2024) 225–234.
- [7] P. Chen, X. Zhao, L. Zeng, L. Liu, S. Liu, L. Sun, et al., A review of research on SLAM technology based on the fusion of lidar and vision, *Sensors* 25 (2025) 1447.
- [8] L. Es Sebar, L. Lombardo, P. Buscaglia, T. Cavaleri, A. Lo Giudice, A. Re, 3D multispectral imaging for cultural heritage preservation: the case study of a wooden sculpture of the Museo Egizio di Torino, *Heritage* 6 (2023) 2783–2795.
- [9] R. Lopez-Campos, J. Martinez-Carranza, Espada: Extended synthetic and photogrammetric aerial-image dataset, *IEEE Robotics and Automation Letters* 6 (2021) 7981–7988.
- [10] M. Chen, Q. Hu, Z. Yu, H. Thomas, A. Feng, Y. Hou, et al., STPLS3D: A large-scale synthetic and real aerial photogrammetry 3D point cloud dataset, *arXiv preprint arXiv:2203.09065* (2022).
- [11] M. Pepe, V. S. Alfio, D. Costantino, D. Scaringi, Data for 3D reconstruction and point cloud classification using machine learning in cultural heritage environment, *Data in Brief* 42 (2022) 108250.
- [12] E. Pellis, A. Masiero, M. Betti, G. Tucci, P. Grussenmeyer, A photogrammetric image-point dataset for the semantic segmentation of heritage buildings, *Data in Brief* 60 (2025) 111661.
- [13] M. Janiszewski, S. Elmas, A. A. Markou, J. Jaaranen, G. H. Filz, 3D dataset of a twisted bending-active beam element digitized using structure-from-motion photogrammetry, *Data in Brief* 54 (2024) 110254.
- [14] D. Palumbo, O. Menozzi, A. K. H. Dewedar, M. Pepe, From documentary analysis of historical texts to 3D modeling: A multidisciplinary approach to reconstructing archaeological sites, in: *2025 6th International Conference in Electronic Engineering & Information Technology (EEITE)*, 2025, pp. 1–7.
- [15] F. Niccolucci, J. L. Richards, The need for FAIR data in archaeology, *Internet Archaeology* 58 (2021).
- [16] M. L. Vincent, V. M. López-Menchero Bendicho, J. Sastre de Diego, FAIR data and open science in cultural heritage: challenges and opportunities, *Journal on Computing and Cultural Heritage* 15 (3) (2022) 1–17.