POLITECNICO DI TORINO
Repository ISTITUZIONALE

Accessible 3D Gaussian Splatting from Consumer-Grade UAVs as a Framework for Rapid Documentation of Cultural Heritage

Original
Accessible 3D Gaussian Splatting from Consumer-Grade UAVs as a Framework for Rapid Documentation of Cultural Heritage / Martino, A., Chiabrando, F.. - In: HERITAGE. - ISSN 2571-9408. - Special Issue Architectural Heritage and Cultural Landscape - 9:8(2026). [10.3390/heritage9080328]

Availability:
This version is available at: 11583/3016119 since: 2026-09-29T22:25:31Z

Publisher:
MDPI

Published
DOI:10.3390/heritage9080328

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

15 November 2026

heritage

IMPACT FACTOR 2.6

CITESCORE 4.3

Essay

Accessible 3D Gaussian Splatting from Consumer-Grade UAVs as a Framework for Rapid Documentation of Cultural Heritage

Alessio Martino and Filiberto Chiabrando

Special Issue

Architectural Heritage and Cultural Landscape

Edited by

Dr. Silvia Beltramo, Dr. Filiberto Chiabrando and Prof. Dr. Monica Naretto

MDPI

https://doi.org/10.3390/heritage9080328

heritage

MDPI

Essay

Accessible 3D Gaussian Splatting from Consumer-Grade UAVs as a Framework for Rapid Documentation of Cultural Heritage

Alessio Martino * and Filiberto Chiabrando

Department of Architecture and Design, Polytechnic University of Turin, Viale Pier Andrea Mattioli 39, 10125 Turin, Italy; filiberto.chiabrando@polito.it
* Correspondence: alessio.martino@polito.it

Abstract

Cultural heritage faces accelerating threats from armed conflict and climate-induced disasters, generating emergency scenarios in which the window for documentation may be measured in hours. Traditional high-fidelity workflows such as Terrestrial Laser Scanning and survey-grade Structure-from-Motion campaign can be operationally incompatible with these constraints due to costly hardware and specialist-team requirements. Emergency image acquisition can be undertaken with consumer-grade equipment, while the same photographic dataset can subsequently support both an SfM-MVS mesh for metric and geometric analysis and a 3D Gaussian Splatting (3DGS) representation for view-dependent visual interpretation. This paper evaluates an accessible, entirely graphical workflow for producing these complementary outputs, with particular attention paid to the 3DGS stage and without requiring programming or command-line expertise. We develop a theoretical case drawing on Brandinian restoration theory and the international conservation framework, introducing the concept of informational amnesia, the irreversible loss of a site's documentary record, as a distinct and undertheorized category of heritage failure. We further argue that the democratization of 3DGS documentation cannot be measured by licensing cost alone: a tool distributed as open-source code but requiring command-line expertise presents an effective accessibility barrier functionally equivalent to a commercial paywall. True accessibility requires installable, graphical software operable by non-specialist practitioners in the field. We present a proof-of-concept application of this accessible pipeline, namely DJI Mini 5 Pro, RealityScan for photogrammetric reconstruction, and Lichtfeld Studio for 3DGS generation, to the Fontana d'Ercole at the Reggia di Venaria Reale (UNESCO World Heritage, Turin, Italy). Visual results show photorealistic, navigable 3DGS surrogates of architecturally complex heritage assets achievable by non-specialist operators within hours of acquisition and establish the empirical foundation for future quantitative evaluation.

Keywords: 3D Gaussian Splatting; AI-driven reconstruction; rapid documentation; cultural heritage; consumer UAV; accessibility; democratization

Check for updates

Academic Editor: Devrim Akca

Received: 17 June 2026
Revised: 1 August 2026
Accepted: 14 August 2026
Published: 19 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Cultural heritage is being lost at a rate that outpaces the discipline's capacity to document it. Armed conflicts, climate-induced disasters, and systemic neglect collectively threaten hundreds of sites [1], while sudden-onset events can reduce centuries of architectural fabric to rubble within minutes. The 2016 seismic sequence in Central Italy caused the

partial or total collapse of over 300 registered heritage structures within 72 h [2], including the near-total destruction of the Basilica di San Benedetto in Norcia.

This paper argues that the central challenge of heritage conservation in high-risk environments is no longer exclusively one of physical preservation but increasingly one of informational survival: the capacity to capture, before destruction, the evidence from which the cultural significance of a site can be transmitted across time. Where no high-fidelity record exists prior to a destructive event, the capacity for any future critical reconstruction is permanently severed. We term this condition informational amnesia, a form of heritage failure distinct from physical degradation alone, because it is the one loss for which no subsequent intervention can compensate.

Addressing informational amnesia under emergency conditions demands documentation tools capable of operating under severe constraints of time, physical access, and financial resources. The traditional techniques, such as Terrestrial Laser Scanning (TLS) and survey-grade Structure-from-Motion/Multi-View Stereo (SfM-MVS) pipelines, have come to be regarded as the professional benchmark for heritage recording, mainly because their metric accuracy and geometric completeness align with the documentary rigor that frameworks such as the Venice Charter and ICOMOS guidelines implicitly demand. Structure-from-Motion recovers camera poses and a sparse three-dimensional point cloud from an unordered set of images through feature matching and bundle adjustment and then performs dense depth-map estimation from those calibrated views to reconstruct textured polygonal surfaces. TLS instrumentation and comprehensive survey-grade SfM-MVS campaigns can still require substantial equipment, multi-person specialist teams, and, for large or geometrically complex sites, demanding processing times measured in days [3]. In an emergency scenario, where mandatory safety clearances routinely close access windows within hours, this combination of specialist-team dependence and high instrument cost is operationally incompatible with the pace of the crisis.

3D Gaussian Splatting (3DGS) [4] provides a complementary image-based representation within this documentation chain. Using the same source images and estimated camera poses as SfM-MVS, it optimizes a radiance field for generating photorealistic, navigable digital surrogates of heritage sites within hours of image acquisition, using consumer-grade hardware and freely available software. Its application to cultural heritage sites remains recent, and existing comparisons of 3DGS against NeRF and SfM-MVS are already diversifying across exterior architecture, immersive environments, threatened sites, and damaged objects [5,6]; the present paper extends this emerging literature to a UAV-acquired, exterior, sculpturally complex heritage site under an explicit accessibility framing. However, the transition from a technically capable method to a genuinely deployable emergency tool requires a distinction that the literature has not yet made explicit. The original reference implementation of 3DGS [4], while foundational and open-source, requires CUDA configuration, command-line operation, and Python dependency management that present an effective barrier to non-specialist operators, here understood as practitioners without formal geomatics training, Global Navigation Satellite System (GNSS) or total-station survey experience, or software-development skills such as CUDA configuration or Python scripting, though not as practitioners without any acquisition-related knowledge whatsoever. Open-source code is not the same as accessible software. A practitioner responding to a heritage emergency in Norcia or Amatrice cannot be expected to debug a CUDA installation. The democratization of documentation depends not on whether source code is publicly available but on whether a non-specialist in this narrower sense can move on standard hardware from installation to trained scene without software development expertise.

Emergency heritage-documentation research has demonstrated the value of multi-temporal UAV photogrammetry, integrated TLS-UAV survey, and 3D spatial records for

post-earthquake and post-fire assessment and decision-making [7,8]. Studies aimed at non-specialist adoption likewise show that operational accessibility depends on explicit acquisition protocols, task-appropriate workflow selection, and robust interfaces rather than on low equipment cost alone [9,10]. Finally, the London Charter and FAIR-oriented 3D-data frameworks require source documentation, provenance, paradata, and preservation strategies so that photorealistic outputs remain interpretable and reusable over time [11,12]. Together, these strands position the present study not as a replacement for photogrammetry but as an investigation of how a GUI-based 3DGS output can be added to an established image-based documentation chain.

This paper addresses that gap directly. We evaluate Lichtfeld Studio [13], a graphical 3DGS application available as a free downloadable installer for Windows, as the primary documentation tool in our pipeline. We further compare its outputs qualitatively with those of Jawset PostShot [14], a commercially licensed but similarly accessible GUI-based alternative, to provide practitioners with evidence on the quality trade-off between the free and paid tiers of accessible 3DGS software. Together, these tools define an accessible pipeline whose adoption barrier is defined by hardware investment and the time to learn a graphical interface, not by software licensing cost or development expertise.

This paper is structured around three research questions: can the complete acquisition-to-3DGS workflow be executed through installable, graphical software alone, without programming or command-line expertise; how does a 3DGS reconstruction compare, qualitatively, with a textured polygonal mesh for heritage surface documentation; and how do a free (Lichtfeld Studio) and a commercially licensed (Jawset PostShot) accessible 3DGS tool compare in output quality and processing time?

The specific contributions of this paper, developed in response to these questions, are as follows: a theoretical framework positioning emergency 3DGS documentation within conservation science, introducing informational amnesia as a first-order ethical category; an explicit argument for distinguishing open-source availability from operational accessibility in the evaluation of democratized documentation tools; and a proof-of-concept visual demonstration of the accessible pipeline applied to the Fontana d'Ercole at the Reggia di Venaria Reale (Turin, Italy), a UNESCO World Heritage Site [15] whose history of material loss and critical reconstruction makes it an ideal testbed for these questions. The study therefore evaluates 3DGS as an additional derivative of a photogrammetric acquisition, not as a replacement for SfM-MVS.

2. Theoretical Framework

2.1. Authenticity in Crisis: From Material Integrity to Informational Survival

The concept of authenticity has long served as the primary qualifying criterion for heritage value. Within the foundational framework of the Italian school of restoration, authenticity was inseparable from the physical integrity of the original material, the *consistenza materiale*, as established by Cesare Brandi's dialectic between the historical instance (*istanza storica*) and the esthetic instance (*istanza estetica*) [16]. The historical instance represents the work as a product of a specific human time and place, bearing the unmediated witness of history; the aesthetic instance refers to its formal presence and artistic value. Brandi argued that the material substrate is the irreplaceable vehicle of the historical instance, and that any restoration must protect it from erasure, codified in the principle of *riconoscibilità* (recognizability), which demands interventions to be clearly distinguishable from original fabric to prevent the creation of the "false historical evidence".

This material-centric paradigm was codified internationally by the Venice Charter [17], which defined the authentic monument as one whose value resides in its original fabric and historical layering. Yet the framework was designed to govern physical intervention,

such as restoration, consolidation, or reconstruction. It offers no doctrine for the situation in which the physical fabric is not at risk of being altered but of being destroyed entirely, and in which the question is not how to intervene on what exists but whether to document it before it ceases to exist.

The Nara Document on Authenticity [18] initiated the conceptual revision required to address this gap. By demonstrating that authenticity is not a universal physical constant but a culturally contingent value, one that can reside in craft continuity, spatial memory, and transmitted knowledge as much as in original stone, the Nara Document shifted the discipline's frame from the object to the process. This shift carries a critical implication for digital documentation: if authenticity can inhere in the transmission of knowledge about a site, then the act of recording that knowledge before physical loss becomes an ethical obligation with the same value of physical conservation itself.

It is within this theoretical space that informational amnesia acquires its full significance. We define it as the irreversible severance of the interpretive chain between a destroyed site and its future scholarly and social reception, the condition that obtains when a monument is lost without a record sufficient to sustain critical understanding of what it was. A well-documented ruin is richer in historical value than an undocumented intact building, because the record preserves the capacity for interpretation even when the object is gone. A monument destroyed without documentation is lost twice: once in its material substance, and once in its informational identity. The second loss is in a meaningful sense the more permanent, because it forecloses interpretive possibilities that physical reconstruction, however imperfect, could in principle reopen.

2.2. The Ethics of Emergency Documentation: Speed and Accessibility as Moral Variables

The theoretical shift from "preservation of the object" to "management of significance" [19] transforms digital documentation from secondary support to primary ethical act. In multi-hazard environments, the failure to document a site at risk before its destruction is no longer a technical oversight but a breach of intergenerational equity, the Faro Convention's principle that the present generation holds cultural heritage in trust for future societies [20]. This ethical reframing has two direct operational consequences that the literature has not yet made sufficiently explicit.

The first is the primacy of speed. Every heritage documentation scenario has what we term a "Golden Hour": the interval between the onset of a destructive event and the moment at which secondary collapses, mandatory safety clearances, or physical inaccessibility permanently close the window for survey. In practice, this window is frequently measured in hours, and it constrains field acquisition specifically: once a site becomes inaccessible, no further images can be captured, whereas the images already acquired remain a durable archival record that can be processed at any later time. Any documentation workflow that cannot be initiated and produce a usable archival output within this interval fails the ethical test, regardless of its technical quality, because it is acquisition, not processing, that an emergency permanently closes off. Processing speed becomes independently decisive in the narrower circumstance where a completed model is needed immediately, on-site or within hours, to support urgent structural or conservation decisions before further collapse; documentation in this dual sense serves both an archival function and, where required, an operational one. A photorealistic 3DGS surrogate captured in two hours is worth infinitely more than a millimetric TLS scan that could only have been performed the following week, after the building had collapsed. Accordingly, the ethical priority is to secure an adequate primary record during the available access window; the source images can then support both photogrammetric and 3DGS derivatives, while rapid processing is required only

when an immediate model is needed for operational assessment. Within the framework established here, speed is an ethical variable, not merely an operational convenience.

The second consequence is the primacy of accessibility. A documentation tool is only an emergency tool if it can be deployed by the people who are present in the immediate aftermath of a crisis: local professionals, first responders, community members, and researchers without specialist geomatics training. A workflow that requires a team of surveyors, expensive instruments, and a software license available only to subscribers is not an immediate tool for the vast majority of at-risk heritage; it is a tool for the fraction of that heritage that already has institutional support. True emergency documentation must be executable by a single non-specialist operator equipped with affordable hardware and software that does not require development expertise to install and operate.

It is within this dual framework of speed and accessibility as ethical variables that the case for accessible 3DGS software, rather than raw open-source code, becomes not merely a practical preference but a theoretically grounded methodological choice.

3. The Proposed Accessible 3DGS Documentation Pipeline

3.1. Consumer-Grade UAVs: Accessible Aerial Survey

Within the European regulatory framework (EASA) established by Regulation EU 2019/947, unmanned aircraft with a maximum take-off mass below 250 g are classified in the Open Category A1, which is subject to fewer operational restrictions than heavier platforms, including permission to fly near people in many urban environments [21]. This does not constitute a blanket exemption for operation at or near UNESCO World Heritage properties: the regulation allows Member States to designate geographical zones restricting UAS operation for reasons of safety, security, privacy, or environmental protection, and individual authorization can still be required within such zones, including at heritage sites, irrespective of aircraft weight class. For a practitioner responding to a heritage emergency in a restricted zone, verifying local geozone restrictions before flight remains a necessary step, though the A1 weight-class exemption still substantially reduces the general regulatory burden relative to heavier platforms.

The metric performance of current consumer platforms is well-established. Murtiyoso and Grussenmeyer [22] demonstrated that consumer UAV SfM documentation of Javanese temple achieved Root Mean Square Error (RMSE) values of 1.2 cm against ground control points, with dense matching quality determined primarily by image overlap rather than sensor cost. For the documentation tasks most critical in emergency conservation, current sub-€1000 UAVs operating represent a technically and operationally adequate solution [23].

The present study uses the DJI Mini 5 Pro (SZ DJI Technology Co., Ltd., Shenzen, China), a sub-250 g platform acquiring imagery at 12 MP with a CMOS sensor of 13.2 × 8.8 mm, combined with RealityScan 2.0.1 [24] for SfM reconstruction and camera pose estimation. RealityScan is distributed free of charge under a proprietary license, for individuals and organizations with annual gross revenue below 1 million USD, while above this threshold a commercial subscription is required, making the complete acquisition-to-reconstruction pipeline available at zero software cost for the large majority of individual and small-team practitioners relevant to this paper's non-specialist emergency-response scenario, and without any command-line configuration. Critically, RealityScan handles the entire photogrammetric pipeline internally, namely feature matching, bundle adjustment, and camera pose computation, and exports the resulting pose data in a format directly importable by Lichtfeld Studio, bypassing the need for COLMAP [25] or any other standalone SfM tool. This is not merely a practical convenience; it means that the full pipeline from raw images to trained 3DGS scene operates entirely through graphical applications, with no terminal interaction required at any stage.

3.2. 3D Gaussian Splatting: From Polygonal Meshes to Radiance Fields

The dominant paradigm for three-dimensional heritage documentation and dissemination has been dense point clouds [26] and the subsequent product of a polygonal mesh: a discrete surface approximation in which a continuous architectural surface is decomposed into triangular faces each carrying a color value baked statically from source photography [27]. While well-suited to metric analysis and Historic Building Information Modeling (HBIM) integration [28], this representation, relying on a single texture from source photography, cannot reproduce surface appearance: the view-dependent interaction between light, observer angle, and material microstructure that defines how a heritage surface actually looks and feels. This limitation is characteristic of the conventional, single-texture photogrammetric mesh evaluated in this paper rather than of polygonal representation as such; view-dependent or image-based rendering techniques exist that reproduce some of this behavior on mesh geometry by blending multiple projected textures according to viewing angle [29], though these remain far less common in accessible heritage-documentation software than the static-texture workflow compared here.

Neural Radiance Fields (NeRF) [30] resolved this limitation by encoding a scene implicitly within a neural network capable of predicting view-dependent color from any position, achieving photorealistic novel-view synthesis. Applied to heritage contexts, NeRF demonstrated a compelling capacity to reproduce the atmospheric depth and material character of architectural spaces [31]. Its critical limitation was computational: training of large or high-resolution datasets required several hours, and real-time rendering is not achievable on standard hardware, constraints incompatible with emergency deployment. Subsequent neural surface-reconstruction methods such as Neuralangelo [32] improved geometric fidelity specifically on textureless or reflective surfaces by combining multi-resolution hash-grid encodings with neural surface rendering, addressing one of NeRF's persistent weaknesses for materials such as polished stone or water that recur throughout the case study presented in Section 4; this refinement, however, did not resolve NeRF's underlying incompatibility with rapid, hardware-light emergency deployment.

3DGS achieves the radiometric fidelity of NeRF at a fraction of the computational cost. Rather than an implicit neural representation, 3DGS encodes a scene as a probabilistic field of anisotropic Gaussian primitives, each parameterized by its mean position, covariance matrix to produce freely oriented anisotropic ellipsoids, opacity, and color encoded via Spherical Harmonics (SH) coefficients. The SH encoding is the key to radiometric fidelity: each Gaussian stores a compact representation of how its emitted radiance varies with viewing angle, reproducing the glint of light on wet stone, the differential shadow within carved relief, or the subsurface scatter of aged marble as the observer moves. For heritage documentation, such appearance information can be evidentially valuable, but it should not be conflated with the explicit, metrically testable geometry supplied by a photogrammetric surface model.

Training begins from a sparse point cloud and calibrated camera poses, supplied in this study by RealityScan, as described in Section 3.1, iteratively optimizing all Gaussian parameters through differentiable rasterization and photometric loss minimization, with adaptive density control densifying under-reconstructed regions and pruning degenerate primitives. The procedure typically converges within 7000 to 30,000 GPU iterations, translating to training times of under one hour for bounded architectural scenes on modern consumer hardware. This speed places 3DGS firmly within the Golden Hour operational window established in Section 2.2. In this confined experiment, 3DGS training required longer processing time with respect to the RealityScan SfM-MVS output. These timings do not support a general speed advantage for 3DGS; they show that both derivatives can be produced rapidly after a shared acquisition. Accordingly, the preferred documentation

strategy is dual-output: preserve the source images and calibration/pose data; generate a metrically controlled point cloud and mesh for geometric analysis; and generate a 3DGS model for interactive, view-dependent visual interpretation. The representations are not interchangeable. This convergence is already visible in integrated software environments such as MipMap and PIX4Dmatic, which make mesh-based and Gaussian Splatting products available within the same photogrammetric project [33,34].

3.3. From Reference Implementation to Accessible Software: The Accessibility Gap

The original reference implementation of 3DGS [4] is distributed as open-source code and remains the foundational methodological reference against which all derivative tools are benchmarked. However, its deployment requires CUDA toolkit installation, Python environment configuration, management of multiple library dependencies, and execution via command line, a sequence that presupposes software development competence that cannot be assumed in non-specialist heritage practitioners, emergency responders, or community-based documentation initiatives.

This gap between open availability and operational accessibility is the central practical limitation of the current 3DGS ecosystem for heritage applications. It is insufficient to establish that a powerful tool exists and that its code is public; the democratization argument requires the tool to be operable by the people who need it most, in the conditions under which they need to use it. Professionals responding to a post-seismic emergency cannot be expected to resolve CUDA version conflicts. The original implementation is, in this sense, not an emergency documentation tool but an invaluable research one.

The methodological developments since Kerbl et al. [4] have begun to address this gap. Mip-Splatting [35] and 2D Gaussian Splatting [36] have refined the algorithm's geometric and scale-dependent accuracy, while Scaffold-GS [37] has improved scene-structure management for large outdoor environments. More directly relevant to the accessibility question, graphical desktop applications have emerged that encapsulate the 3DGS pipeline behind a GUI-based interface, removing the command-line barrier entirely.

Lichtfeld Studio is the primary tool evaluated in this study. Its source code is publicly available on GitHub [13], satisfying the open-source criterion; crucially, it is also distributed as a compiled Windows installer, eliminating the need for manual compilation or dependency management. The latest stable release is available for a voluntary donation, while previous stable releases are available free of charge. Lichtfeld Studio accepts camera pose data and images directly from RealityScan's export, requiring no intermediate processing step. A practitioner can download, install, and initiate a training run within minutes of first contact with the software, with no prerequisite beyond a Windows PC with a compatible GPU and a completed RealityScan reconstruction. This combination of open-source transparency, installable accessibility, and compatibility with a freely available SfM tool defines the operational profile that genuine democratization requires.

Jawset PostShot [14] provides a commercially licensed counterpart in this accessible tier. At a yearly subscription price of less than 500 €, PostShot offers a polished Windows GUI that handles pose estimation internally, supports real-time training preview, enabling the operator to monitor reconstruction quality and export a partial result if time constraints require it, and provides a range of export options suitable for archival and visualization workflows. For this study, PostShot model reconstruction was tested both from the RealityScan camera pose estimation and from the internal pose estimation to better understand the feasibility of a 3DGS pipeline based on a single software. Its inclusion in this study allows for a direct qualitative comparison between the free and paid tiers of accessible 3DGS software, providing practitioners with evidence on whether the commercial licensing cost is justified by output quality for heritage documentation.

Known limitations shared by any SfM reconstruction software include sensitivity to the quality of the pose estimation and difficulty representing optically complex materials including water, reflective stone, and architectural glass. 3DGS reconstructions carry an additional, distinct sensitivity: because each Gaussian's appearance is optimized against the specific camera viewpoints used during training, rendering quality degrades when a viewer subsequently navigates to positions or angles substantially different from the acquisition geometry, and reliable display depends on the camera orientations and positions used during capture remaining broadly similar to the viewing positions from which the scene will later be observed. These limitations are relevant to the case study presented in the following section and are addressed in the Discussion.

4. Case Study: Fontana d'Ercole, Reggia di Venaria Reale

4.1. Site, Heritage Significance, and the Problem of Material Loss

The Reggia di Venaria Reale, located approximately 10 km north of Turin in Piedmont, is one of the most extensive Baroque royal complexes in Europe, inscribed on the UNESCO World Heritage List in 1997 as part of the "Residences of the Royal House of Savoy" [9]. Initiated by Duke Carlo Emanuele II of Savoy under the direction of Amedeo di Castellamonte from 1659, and substantially restructured by Filippo Juvarra between 1716 and 1728, the complex represents a defining ensemble of Baroque court architecture, landscape design, and hydraulic engineering.

The Fontana d'Ercole occupies a focal position between the Flower Garden and the Central Alley. Conceived by Amedeo Castellamonte from 1669 to 1672 as a theatrical apparatus of water and sculpture staging the dynastic mythology of the House of Savoy through the allegory of Hercules, the original fountain comprised a monumental sculptural group, an elaborate hydraulic system, and a spatial enclosure designed to animate the garden axis as a political spectacle (Figure 1). Following the decline of the Savoy court and successive military occupations from the late 18th century, the fountain was progressively dismantled: the hydraulic system decommissioned, the sculptural elements dispersed, and the architectural fabric was subjected to two centuries of unmanaged decay. The recent restoration, completed in 2022, adopted a strategy of critical reconstruction [38]: surviving fragments were used as generative seeds for a contemporary architectural intervention that acknowledges, rather than conceals, the historical lacunae. The result is an object in which original masonry, fragmentary carved stone, and contemporary elements coexist, with their respective historical statuses physically legible to the informed observer.

This condition makes the Fontana d'Ercole an exemplary testbed for the ethical dimensions of 3DGS documentation developed in Section 2.2. If the Gaussian representation reproduces the photorealistic surface with sufficient continuity that original and reconstructed fabric become visually indistinguishable in the digital record, this constitutes precisely the "false historical evidence" that Brandi's recognizability principle was designed to prevent. Conversely, if the radiometric fidelity of 3DGS preserves the material differentiation that is visible in the physical object, the chromatic and textural distinction between aged original stone and contemporary intervention, then the 3DGS representation adds an appearance-based evidentiary layer that complements the explicit geometric record of the photogrammetric mesh.

The site also presents a representative set of technical documentation challenges: complex three-dimensional geometry combining horizontal water basin surfaces, vertical carved stone, and architectural enclosure; variable surface optical properties including rough weathered limestone, smooth contemporary masonry, and specular water; and an open-air context subject to variable solar illumination. These conditions are typical

of outdoor Mediterranean heritage sites and give the results a degree of generalizability beyond the specific case.

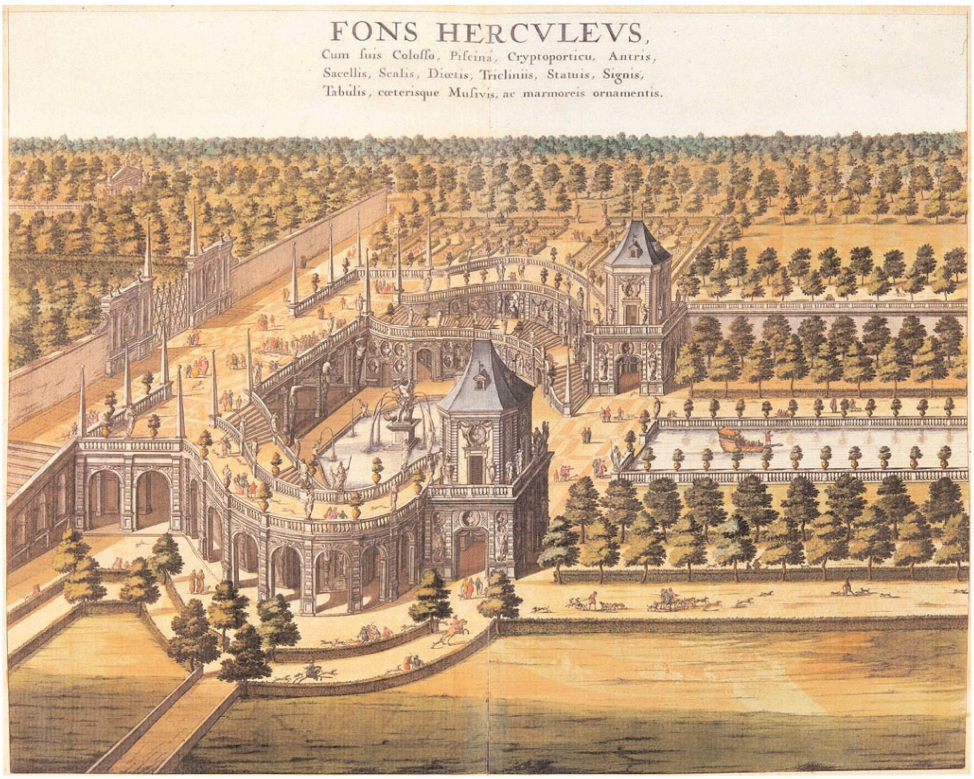


Figure 1. Hercules Fountain original theatrical apparatus of water and sculpture staging the dynastic mythology of the House of Savoy through the allegory of Hercules, designed by architect Amedeo di Castellamonte from 1669 to 1672; inscription published in Theatrum Sabaudiae [39] in 1682.

4.2. Acquisition and Processing

The acquisition was conducted under ordinary, non-emergency site-access conditions; it is intended to test the technical and procedural feasibility of the pipeline, specifically equipment requirements, workflow steps, and time budget, rather than to reproduce the physical hazards, restricted access, or psychological pressure of an actual post-disaster deployment. The documentation campaign was conducted using the DJI Mini 5 Pro UAV equipped with a 12 MP CMOS sensor (13.2 × 8.8 mm). With a take-off weight of less than 250 g, the drone is agile to pilot and therefore suited for close-range acquisitions of complex architecture. The manual flight plan combined cross-grid oblique stripes with radial acquisition around the main element of the Hercules' sculpture, conducted respectively at an elevation of around 10 and 5 m from the terrain, yielding a mean GSD (Ground Sample Distance, the linear ground distance represented by one image pixel) of 5.6 mm/pix. Forward and lateral image overlap were of approximately 80%, the gimbal pitch angle ranged between 30° for the oblique cross-grid pass and 0° for the radial pass around the sculpture, flight speed was maintained at approximately 1–2 m/s, and exposure, ISO, shutter speed, and aperture were set to automatic. Total field acquisition time, from the first to the last image captured, was of about 15 min. A total of 282 images were acquired under clear to partly cloudy daylight conditions, with direct solar illumination, high brightness and pronounced cast shadow. For the scope of this paper, only qualitative results will be compared. Further geometrical and quantitative analyses will be developed in future publications, in which the relative accuracies on control points from 3DGS models will be compared to the TLS ground truth.

The complete image dataset was processed in RealityScan 2.0.1 [24], which performed the camera alignment, including feature detection and matching internally, producing 1.9 million tie points and camera poses without requiring any command-line input in about 3 min. The camera alignment of 282 calibrated poses with a mean reprojection error of 0.64 px was exported directly from the RealityScan interface as the initial configuration for the 3DGS training, and it includes the images, along with three .txt files:

- Cameras.txt describing the camera model as a simplified single-parameter radial model using the default Brown3 distortion model, and internal calibration parameters: focal length (f), principal point coordinates (cx, cy), and radial distortion coefficient (k).
- Images.txt contains the exterior orientation of each image, described by the quaternion for the rotation (Qw, Qx, Qy and Qz) and the three components for the translation (Tx, Ty and Tz), plus all 2D feature observations in that image, described with the ID of the 3D point and the two pixel coordinates on that image (x and y).
- Points3D.txt contains the reconstructed sparse 3D points with their coordinates (X, Y and Z), their color components (R, G and B), the reprojection error in pixels, and their image track, which is the list of image measurements that observe the same reconstructed 3D point.

The MVS pipeline on RealityScan was carried out until the generation of a textured mesh model consisting of 30.6 M triangles, for the product to be visually analyzed in comparison to the 3DGS model, with the parameters and time of processing for each reconstruction step summed up in Tables 1 and 2.

Both the SfM and 3DGS processing described in this section were carried out on a laptop equipped with a NVIDIA GeForce RTX 4070 GPU and 8GB of dedicated vRAM. An exact one-to-one equivalence between the two software's quality tiers could not be established, and the comparison that follows should be read with that caveat. It must be noted that the generation of the final model on RealityScan takes a fraction of the time with respect to the one generated with these settings on commercial SfM software such as Agisoft Metashape Professional [40], most likely caused by the different optimization in rapid GPU-based reconstruction and to the various parametrizations that each step requires on the two solutions. A visual comparison among the textured models processed in each software will be provided in the following section.

Table 1. Photogrammetric reconstruction parameters in RealityScan 2.0.1 (SfM-MVS camera alignment and textured mesh generation).

Step	Parameters	Processing Time
Camera Alignment	Alignment Mode: High Feature Detection quality: High Max Features per mpx: 8000 Max features per image: 80,000 Detector Sensitivity: Medium Image Downscale factor: 1	Feature detection: 30 s Registration: 2 m 30 s
3D Model	Quality Level: Normal Image Downscale Factor for Depth Maps: 2 Coloring method: Multiband Coloring Style: Visibility-based Downscale Images before Coloring: 2 Unwrapping Style: Maximal texture count	Depth-map computation: 2 m 10 s Meshing: 5 m 10 s Post-processing: 1 m 15 s Coloring: 45 s Unwrapping: 10 s
Texture	Texturing Style: Visibility-based Downscale Images before Texturing: 1 Texture Resolution: 8192 × 8192 px Texel Size: 0.0168 m/px	Texturing: 1 m 30 s
Total Processing Time		14 min

Table 2. Photogrammetric reconstruction parameters in Agisoft Metashape Professional 2.3.1 (SfM-MVS camera alignment and textured mesh generation).

Step	Parameters	Processing Time
Camera Alignment	Accuracy: High Key Point Limit: 80,000 Key Point per mpx: 1000 Tie Point Limit: 8000	Matching: 15 m Alignment: 4 m 30 s
Dense Point Cloud	Depth Map Generation Quality: High Point Cloud Source: Depth Maps	Depth Map: 7 m 50 s Point Cloud: 43 m 40 s
3D Model	Source Data: Depth Maps Quality: High Face count: High	Meshing: 14 m 20 s
Texture	Texture Resolution: 8192 × 8192 px	Texturing: 15 m 20 s
Total Processing Time		100 min

The exported pose data and images were then imported into Lichtfeld Studio 0.5.2 [13] via its graphical interface. Training was initiated using default parameters, without modification, in keeping with the non-specialist operational scenario central to the paper's argument. No command-line interaction was required at any stage, from RealityScan processing through Lichtfeld Studio training to scene export. This is stated as an explicit empirical finding: the complete pipeline from raw drone images to photorealistic 3DGS surrogate was executed entirely through graphical Windows applications.

To evaluate processing performance across configurations, the image dataset and RealityScan-derived camera poses were independently processed in Jawset PostShot 1.1.0 [14] using its default parameters, with the main difference regarding the training strategies. LichtFeld-Studio prioritizes MCMC (Markov Chain Monte Carlo), in which more randomized sampling of the scene is used, resulting in less fine detail but being more tolerant with artifacts, imperfect image quality, weaker pose accuracy, moving elements, or noisy captures. Jawset Postshot, on the other hand, recommends the Splat3 profile, able to reconstruct fine detail in both foreground and background, and to use higher-resolution source images.

A second round of processing was then conducted in both tools using a matched set of training parameters, enabling a controlled comparison that isolates the effect of the software implementation from the effect of differing defaults. Finally, two additional PostShot reconstructions were performed using the software's built-in pose estimation rather than the RealityScan-derived poses, testing the feasibility of a fully self-contained, single-software image-to-3DGS workflow. In all configurations, the iteration limit was set to 30,000, the Spherical Harmonics degree to 3, and images were automatically subsampled to a maximum resolution of 3480 px. The remaining training parameters and processing times for each configuration are summarized in Table 3. The default configurations (LFS_D, PS_D) reflect each tool's own out-of-the-box settings, left deliberately unmodified to represent the non-specialist operational scenario central to this paper, rather than values selected or recommended by the authors; for this reason, PostShot's default maximum Gaussian count (3,000,000) differs from Lichtfeld Studio's default maximum Gaussian count (1,000,000) in Table 2, since each figure is the respective software's own shipped ceiling. The matched configurations (LFS_M, PS_M) instead equalize the densification strategy and maximum Gaussian count at 3,000,000 for both tools, specifically to isolate the effect of the software implementation from the effect of the differing configurations. In general, a higher maximum Gaussian count and a higher iteration limit, up to a converging value, allow finer detail to be resolved and reduce the risk of under-reconstructed regions, at the

cost of longer training time, while the Spherical Harmonics degree controls the complexity of view-dependent color that can be represented [4].

Table 3. Gaussian Splatting on LichtFeld-Studio 0.5.2 (LFS) and Jawset Postshot 1.1.0 (PS). Underlining denotes the default configurations of the parameters to be modified before the comparison. All configurations: 30,000 iterations, SH degree 3, and images subsampled to 3480 px.

ID	Tool	Strategy	Max Gaussian	Pose Estimation	Training Time
LFS_D	LichtFeld Studio	MCMC	1,000,000	RealityScan 3 min	30 min
PS_D	Jawset PostShot	Splat3	3,000,000	RealityScan 3 min	40 min
LFS_M	LichtFeld Studio	MCMC	3,000,000	RealityScan 3 min	54 min
PS_M	Jawset PostShot	MCMC	3,000,000	RealityScan 3 min	38 min
PS_D_Pose	Jawset PostShot	Splat3	3,000,000	13 min	40 min
PS_M_Pose	Jawset PostShot	MCMC	3,000,000	13 min	42 min

The default configurations reveal an immediate divergence between the two tools. Lichtfeld Studio (LFS_D) completed training in 30 min, using the MCMC densification strategy with a Gaussian count ceiling of 1,000,000. PostShot (PS_D) required 40 min, operating under the Splat3 strategy with a ceiling of 3,000,000 Gaussians. The difference in training time is therefore attributable to the combined effect of the densification strategy and the threefold difference in maximum Gaussian count rather than to software efficiency alone. Under matched parameters, MCMC strategy and 3,000,000 Gaussians, Lichtfeld Studio (LFS_M) required 54 min, while PostShot (PS_M) completed in 38 min, suggesting that PostShot's implementation of the MCMC strategy is computationally more efficient at equivalent scene complexity. The two PostShot configurations incorporating internal pose estimation (PS_D_Pose and PS_M_Pose) confirm that the pose estimation step adds approximately 13 min to the total pipeline time, with training durations comparable to their externally initialized counterparts. Combining field acquisition of 15 min with RealityScan processing (14 min) and 3DGS training varying from 30 to 54 min depending on the configuration, the complete pipeline from image capture to a trained, navigable 3DGS scene required well under two hours in total, comfortably within the "Golden Hour" window defined in Section 2.2, and also in the narrow circumstance of urgent structural or conservation decisions.

4.3. Visual Results and Demonstrated Potential

The following figures present a structured series of visual comparisons designed to evaluate the accessible 3DGS pipeline across four distinct dimensions: (1) the validity of RealityScan as an accessible SfM foundation relative to the commercial industry standard; (2) the complementary appearance information provided by a 3DGS render and a conventional static-texture mesh generated from the same image set; (3) the effect of software choice and training configuration on 3DGS output quality; and (4) the capacity of the 3DGS surrogate to preserve the material stratification required by the principle of *riconoscibilità*. Each comparison is offered as visual evidence rather than a quantitative benchmark; the formalization of these observations through standardized image quality metrics is identified as the primary direction of future work. Viewpoints across corresponding figure pairs in the SfM-MVS comparisons were matched by manually aligning the virtual camera in each software's viewer to a common reference position and orientation, render resolution was kept consistent within each comparison, and no color grading, exposure correction, or tone-mapping was applied to any output beyond each tool's default display settings. The 3DGS model comparisons, on the other hand, have been generated on the external

SuperSplat editor [41], in which the reconstructions are perfectly overlapping since they are generated from the same initialization point cloud and camera poses.

Before evaluating 3DGS output quality, it is necessary to establish that the SfM reconstruction underpinning the entire pipeline, produced by RealityScan, is a scientifically credible photogrammetric product. This is not self-evident on its face, since Agisoft Metashape Professional remains the industry standard for architectural photogrammetry. Before its 2024 rebranding and revenue-based free tier, however, RealityScan's underlying engine, RealityCapture, was already used as a commercial photogrammetric and surveying software with a substantial independent accuracy-validation literature, including studies on UAV survey accuracy [42] and, specifically in a cultural heritage context, UAV-and-TLS fusion for historic building documentation [43]. The evidence gap this section addresses is therefore narrower than a general absence of validation: it concerns the benchmarking of the current free consumer tier's export pipeline into GUI-based 3DGS software, specifically within emergency heritage-documentation contexts, which remains comparatively under-studied.

The comparison in Figures 2 and 3 shows that RealityScan handles the complete photogrammetric pipeline of camera alignment, meshing, and texturization, with a combined processing time of 14 min, generating a mesh with a higher polygon count than the one processed in 100 min in Metashape. Polygon and tie-point counts are reported here as descriptive processing outcomes rather than as quality indicators: a denser mesh or point cloud does not by itself indicate a higher-quality reconstruction, as the broadly comparable visual result Metashape achieves despite its sparser output, discussed below, illustrates. The case for RealityScan as a credible zero-cost SfM foundation within a rapid-response documentation pipeline instead rests on the visual comparison in Figure 3. At the texture quality level observable in Figure 3, the difference between the two outputs on the architectural and sculptural elements is not visually decisive for heritage archival purposes, suggesting that the commercial licensing cost of Metashape does not directly translate into a proportionate quality gain for this acquisition configuration. The shared failure reconstruction on the water basin is not attributable to either tool's algorithm but to the photogrammetric incompatibility of specular and transparent surfaces with passive image-based reconstruction.

The following figures show how the textured mesh generated from RealityScan bakes illumination into a static texture at the time of reconstruction; as a result, the surface appearance does not change as the observer moves, producing a flat, view-independent rendering in which the water surface is partially absent and the detail on the architectural stone and on the masonry reads as uniformly lit regardless of viewing angle (Figure 4, left). The 3DGS representation encodes view-dependent color through Spherical Harmonics, reproducing the chromatic modulation of the water surface, the luminance gradient across the marble of the Hercules statue as the observer's position shifts, and the differential shadow depth within the carved bas-relief on the pedestal. This difference is not esthetic but scientific: for a site where surface appearance encodes material condition, such as the presence of moisture, the depth of relief, and the chromatic contrast between original marble and reconstructed masonry, the 3DGS surrogate preserves radiometric evidence that the mesh structurally cannot capture (Figure 4, right).

Figure 5 compares the two Lichtfeld-Studio configurations at the same viewpoint, isolating the effect of increasing the maximum Gaussian count from 1M (LFS_D) to 3M (LFS_M) under the same MCMC densification strategy. In both renders, the Hercules statue is convincingly resolved: the surface modeling of the musculature, the carved detail of the bas-relief panel on the pedestal face, and the water reflection beneath the plinth are all legible with a degree of visual fidelity that no textured mesh at equivalent processing time could replicate. The primary observable difference between the two configurations

lies in the precision of fine geometric transitions: LFS_M allocates the additional Gaussian budget to the boundaries between raised and recessed surfaces, producing marginally sharper edge definition on the bas-relief and on the molded profiles of the pedestal. The improvement, while real, is subtle at normal viewing distances and unlikely to be decisive for emergency archival purposes, where the priority is scene completeness rather than sub-centimeter surface precision.

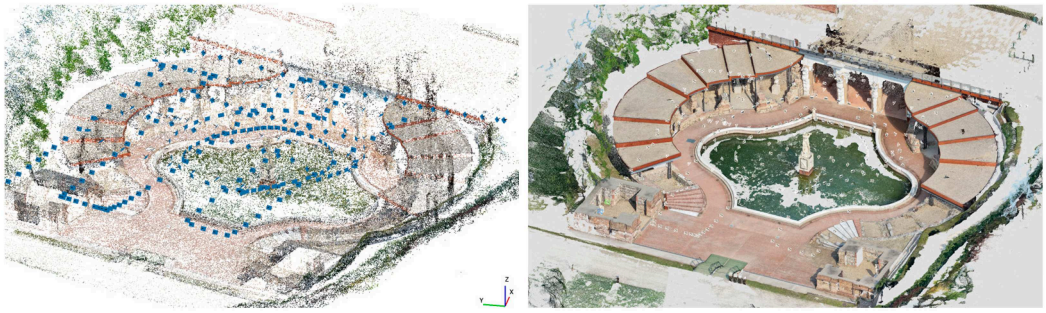


Figure 2. Sparse point cloud reconstructions of the Fontana d'Ercole generated from the same DJI Mini 5 Pro image dataset: Agisoft Metashape Professional (**left**) and RealityScan (**right**). Metashape produced 312k points in 20 min; RealityScan produced 1.9M points in 3 min.

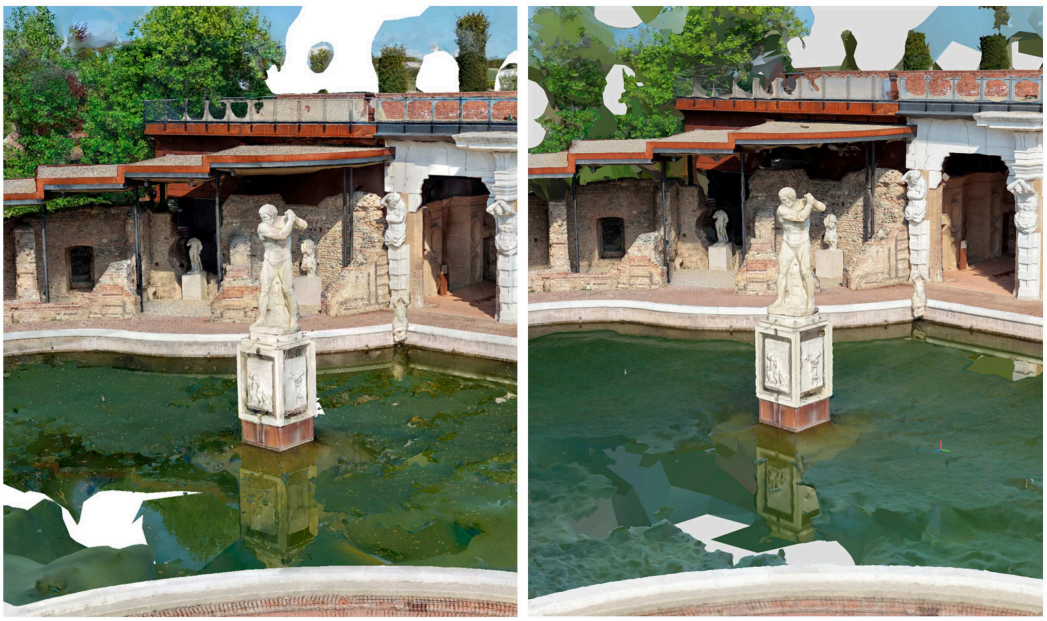


Figure 3. Detailed textured mesh reconstructions of the Fontana d'Ercole, Agisoft Metashape Professional (**left**) and RealityScan (**right**), rendered from the same viewpoint.

Figure 6 presents the equivalent comparison within Jawset PostShot, contrasting the default Splat3 strategy (PS_D) with the MCMC strategy used for the matched-parameter evaluation (PS_M), both at a maximum of 3M Gaussians. The two renders are the most visually similar pair in the entire comparison set: the transition between densification strategies within PostShot, at equivalent Gaussian count and iteration limit, produces differences that are less perceptible than the Gaussian count increase observed in Figure 5. Both PostShot configurations render the Hercules statue with a slightly warmer color temperature than the Lichtfeld Studio outputs. PS_M shows a subtle improvement in pedestal edge definition over PS_D, consistent with the MCMC strategy's stochastic relocation mechanism providing more efficient Gaussian coverage of the fine molded profiles.



Figure 4. Comparison of the Fontana d'Ercole: RealityScan textured mesh (**left**) and Lichtfeld Studio 3DGS model in the default configuration LFS_D (**right**), rendered from identical viewpoints: a general overview of the scene (**top**), a close-up on the aspect of different materials and surfaces (**middle**), and a nadiral perspective, substantially different from the acquired images (**bottom**).



Figure 5. Close-range 3DGS renders of the central Hercules statue and pedestal of the Fontana d'Ercole, processed in Lichtfeld Studio: default configuration LFS_D ((**left**), MCMC strategy, maximum 1M Gaussians, 30 min) and matched configuration LFS_M ((**right**), MCMC strategy, maximum 3M Gaussians, 54 min).



Figure 6. Close-range 3DGS renders of the central Hercules statue and pedestal of the Fontana d'Ercole, processed in Jawset PostShot: default configuration PS_D ((**left**), Splat3 strategy, maximum 3M Gaussians, 40 min) and matched configuration PS_M ((**right**), MCMC strategy, maximum 3M Gaussians, 38 min).

Figure 7 isolates the effect of camera alignment source on the quality of the PostShot 3DGS reconstruction by holding all training parameters constant and varying only the pose

estimation method. The difference in sparse point count between the two approaches is substantial: RealityScan produces 1.9M sparse points in 3 min, against PostShot's internal alignment, which generates 97.3k points in 13 min. The consequences for the final 3DGS output are directly observable. The model initialized from RealityScan poses produces a spatially coherent scene in which the Hercules statue is sharply defined, the arcaded portico reads with clear depth recession, and the column shafts, Atlante figures, and interior mosaic floor are all individually legible. The model initialized from PostShot's internal alignment shows a visually apparent quality degradation across the scene. The degradation observed here reflects some combination of factors at the pose-estimation stage, including calibration accuracy, feature-matching precision, and the spatial distribution of points across the scene, rather than point count in isolation; PostShot's internal alignment parameters at this stage are not exposed by the software and so could not be examined directly here, a constraint noted as a limitation in Section 5.3. This finding has a direct implication for the accessible pipeline proposed in this paper: PostShot's single-application image-to-3DGS workflow, while operationally simpler, produces a reconstruction of notably lower quality than the two-step approach in which RealityScan handles pose estimation before PostShot or Lichtfeld Studio performs the Gaussian optimization. For emergency documentation scenarios where archival quality is the priority, the additional step of external pose estimation via RealityScan is justified both by the superior output quality illustrated here and by the shorter combined processing time.



Figure 7. Effect of camera alignment source on PostShot 3DGS reconstruction quality. (**Left**): external RealityScan camera alignment (1.9M sparse points, 3 min pose estimation). (**Right**): PostShot internal camera alignment (97.3k sparse points, 13 min pose estimation). Both models use identical 3DGS training parameters: MCMC strategy, 3M maximum Gaussians, 30k iterations.

Figure 8 extends the LFS_M versus PS_M comparison to a materially more complex area of the Fontana d'Ercole ensemble, one in which the documentary challenge of recognizability is most explicitly present. The scene encompasses four distinct material and temporal layers within a single frame: the modern brick pavement in the foreground; a gravel infill surface transitioning to the restored brick-and-stone stepped element at the center; the original historic rubble masonry wall of mixed limestone and flint in the background; and the contemporary corten steel canopy structure at the upper register. Despite the parameters being identical across both configurations, the two implementations diverge in their handling of the flat and low-relief surfaces that dominate this area of the scene. Lichtfeld Studio's LFS_M produces a spatially consistent reconstruction across the full scene: the foreground brick pavement is resolved with clear mortar joint definition, and the central stone and mortar wall reads as a coherent, texturally homogeneous surface. PostShot's PS_M, by contrast, introduces several Gaussian floaters on the central wall section, misplaced primitives that have survived the MCMC pruning process and appear

as semi-transparent, luminance-inconsistent patches overlaid on the masonry texture, as well as a less defined reconstruction of the modern brick pavement on the foreground, in which the individual elements are not distinguishable. These issues do not affect the overall legibility of the scene but represent localized failures of the densification algorithm that would require manual correction before the model could be used as a precise material record of that surface. These localized floater artifacts may plausibly be reduced through denser image acquisition or video-based capture of this specific area, though this has not been tested in the present study and should be verified experimentally under comparable conditions before being treated as a general mitigation strategy. Both configurations exhibit mild Gaussian softness on the sculptural figures, attributable to the high depth discontinuity between the thin statue geometry and the background masonry, and reduced surface detail in the deep shadow zone beneath the contemporary canopy, where limited photographic information in the training images constrains radiometric reconstruction.



Figure 8. Close-range 3DGS renders of a secondary architectural area of the Fontana d'Ercole ensemble: Lichtfeld Studio matched configuration LFS_M ((**left**), 54 min) and Jawset PostShot matched configuration PS_M ((**right**), 38 min). The scene encompasses four distinct material and temporal layers: the modern brick pavement in the foreground; the restored brick-and-stone stepped element at the center; the original historic rubble masonry wall in the background; and the contemporary corten steel canopy at the upper register. In the red boxes the most notable differences between the two 3DGS reconstructions.

5. Discussion

5.1. The Accessibility Argument

The visual sequence presented in Section 4.3 confirms that the proposed pipeline is accessible in a stronger sense than software cost alone. The acquisition was performed by a photogrammetry-trained operator with a sub-250 g consumer UAV operating within the EU Open Category A1 regulatory framework; camera alignment and mesh generation were completed in free photogrammetric software through a graphical interface; and 3DGS training was executed in open-source software without any command-line interaction at any stage. This supports an operational-accessibility claim about the graphical toolchain, not a claim that 3DGS requires less acquisition knowledge or processing than photogrammetry. Accessibility here denotes the removal of programming and command-line barriers; it does not remove the need for competent image-network design.

The results also suggest that accessibility does not require sacrificing photogrammetric reliability. RealityScan produced a coherent camera alignment and generated a textured

mesh whose visual quality, for the purposes of rapid archival documentation, was not inferior to the Metashape output shown in Figures 2 and 3, at a fraction of the processing time. More fundamentally, Figure 7 shows that the quality of the subsequent Gaussian reconstruction is strongly conditioned by the upstream pose estimation. PostShot's internal alignment generated 97.3k sparse points in 13 min, a nearly twentyfold reduction relative to the 1.9M points produced by RealityScan in 3 min: the externally initialized model produces a spatially coherent, sharply defined scene, while the internally initialized model shows a visually apparent degradation in scene coherence, statue definition, and architectural legibility. As discussed in Section 4.3, this sparse-point disparity is an observable correlate rather than an established sole cause of the quality gap, and the specific pose-estimation factors responsible remain to be isolated in future work. The most robust accessible workflow is therefore not determined by the fewest possible software steps but by the shortest fully graphical sequence that preserves a dense and accurate pose initialization.

The comparison between Lichtfeld Studio and PostShot introduces a more specific cost-versus-benefit question. At matched parameters, PostShot completed the MCMC reconstruction 16 min faster than Lichtfeld Studio, and the central statue comparisons in Figures 5 and 6 confirm that both tools reproduce the Hercules figure, pedestal reliefs, and water reflections with broadly comparable visual fidelity. However, Figure 8 reverses this picture for the secondary architectural area. Lichtfeld Studio's LFS_M preserves the legibility of the modern brick pavement, the restored stonework, the historic rubble masonry, and the corten canopy with greater spatial consistency and without localized artifacts, while PostShot's PS_M introduces Gaussian floaters on the central wall section and produces a less resolved foreground pavement in which individual brick units are not clearly distinguishable. The paid tool does not provide a visual-quality advantage in this scene; the free tool performs more consistently across the range of surface types present.

The relevant distinction is not free versus paid software in the abstract, nor open-source code versus proprietary distribution in isolation. The operative criterion is whether a toolchain can be installed, understood, and deployed rapidly by the practitioners most likely to be present before, during, or immediately after a crisis, and whether it produces consistent results across the material and geometric complexity of real heritage sites rather than on idealized benchmarks. Lichtfeld Studio currently satisfies both requirements as a free or donation-based graphical 3DGS environment; PostShot satisfies the accessibility requirement as a commercial environment with additional workflow conveniences, but the results presented here suggest that the quality advantage of the paid option is not consistent across scene types and cannot be assumed in advance. Taken together, the timings and outputs argue against treating 3DGS as a competitor to SfM-MVS. The same 282 images supported both products, so field-acquisition time was identical. RealityScan produced the complete SfM-MVS reconstruction in approximately 14 min, while 3DGS optimization required an additional 30–54 min depending on configuration; neither representation therefore held a general processing-time advantage in this experiment. Their evidentiary functions are different: the mesh provides an explicit surface for metric inspection, mapping, and GIS/HBIM integration, whereas 3DGS supports real-time, view-dependent visual interpretation of material appearance. For emergency documentation, the recommended workflow is to preserve the source photographs and camera solution and derive both outputs whenever resources permit.

5.2. Ethical Dimensions: Recognizability in the Gaussian Field

The radiometric fidelity that makes 3DGS valuable for documentation simultaneously constitutes its most significant ethical risk. A Gaussian field produces a continuous, photo-realistic visual experience in which measured image evidence and algorithmic interpolation

are not automatically distinguishable to the viewer. This condition directly challenges Brandi's principle of *riconoscibilità*: if a digital surrogate obscures the boundary between documented fabric, restored fabric, and computational inference, it risks generating precisely the "false historical evidence" that conservation theory seeks to prevent.

The case study provides a meaningful test of this risk because the Fontana d'Ercole is explicitly a palimpsest. Within a single scene, as documented in Figure 8, four distinct material and temporal layers coexist: modern brick pavement, a restored brick-and-stone stepped element, original historic rubble masonry, and a contemporary corten steel canopy. The Lichtfeld Studio reconstruction preserves the chromatic and textural differentiation between these layers with sufficient fidelity that each can be identified visually in the digital surrogate without reference to the physical site. The modern pavement reads differently from the historic stone, the restored stepped element is distinguishable from the original masonry behind it, and the contemporary canopy is legible as a temporally distinct addition. This is a notable observation from this single case study, based on visual inspection rather than a formal recognizability or usability test with independent viewers: the Gaussian surrogate does not automatically homogenize historical layers into a seamless photorealistic surface, and where photographic coverage and camera geometry are adequate, it can transmit the material cues on which recognizability depends. The same capacity is evident in the close-range statue views of Figures 5 and 6, where the chromatic differentiation between the white marble of the restored Hercules figure, the terracotta brick of the pedestal base, and the greenish water surface is preserved in both Lichtfeld Studio and PostShot outputs.

The same figures also define the conditions under which this conclusion must be qualified. Local Gaussian floaters in PostShot's PS_M model introduce luminance-inconsistent patches on the historic masonry wall that have no correspondence in the photographic record, constituting precisely the kind of artifact that a non-specialist user might interpret as a genuine material feature. Deep shadow zones beneath the contemporary canopy lose surface detail in both reconstructions, not because the Gaussian optimization has failed but because the training images themselves contain insufficient photographic information in those regions. In each of these cases, the 3DGS model's photorealistic visual persuasiveness can exceed its evidentiary transparency: the image looks complete and convincing in areas where its evidentiary foundation is weak or absent.

A practical consequence follows from this ethical reading. Archival 3DGS models should be accompanied by acquisition metadata, training parameters, camera-pose provenance, and, where possible, spatial confidence or uncertainty maps that indicate which regions of the scene are directly supported by dense photographic observation and which are primarily interpolated by the Gaussian optimization. The Gaussian floaters visible in Figure 8 would be identifiable in such a map as regions of low-confidence reconstruction, a disclosure that would allow the conservation professional to treat them appropriately rather than mistake them for material evidence. Current accessible implementations do not yet provide such overlays as standard outputs. Their development represents a necessary step toward aligning 3DGS documentation practice with the conservation principle that archival records must disclose not only what they show but also the evidentiary basis on which they show it.

5.3. Limitations and the Agenda for Future Quantitative Research

The present study remains a qualitative proof of concept. It suggests operational feasibility, visual plausibility, and conservation relevance, but it does not establish the metric accuracy, radiometric stability, or long-term archival reliability required for 3DGS to

function as a disciplinary standard. The results should be read as evidence that accessible 3DGS is worth validating rigorously, not as a substitute for that validation.

The first limitation concerns geometric accuracy. The 3DGS reconstructions are bounded by the accuracy of the camera poses supplied by RealityScan, and the present paper does not evaluate those poses against an independent TLS, total-station, or GNSS reference. Such a comparison is not straightforward for a Gaussian field in the way it is for a polygonal mesh: a 3DGS model does not correspond to a specific physical surface but constitutes an irregular representation of a radiance field, in which the appearance of a given surface from a given viewpoint is typically formed by the blended contribution of several overlapping Gaussians rather than a single, unambiguous 3D coordinate. While many works have been using CloudCompare's M3C2 plugin [44] as a tool for measuring distance between 3DGS-derived point clouds and the TLS reference, direct measurement of lengths or point coordinates within a 3DGS scene is not a simple extension of mesh-based metrology. Future work must quantify camera-pose accuracy, model scale stability, and local surface deviation on control and check points, particularly on low-texture masonry, smooth stone, water-adjacent surfaces, and thin sculptural elements, the conditions in which both emergency heritage documentation and photogrammetric initialization are most demanding. The 20x difference in sparse point density between RealityScan and PostShot's internal alignment further motivates a systematic evaluation of how pose-initialization density affects final reconstruction accuracy across different surface types.

The second limitation concerns image-quality evaluation. The comparisons presented here are visual and interpretive. They should be formalized through held-out view evaluation using standardized metrics, such as PSNR, SSIM, and LPIPS, with the original 3DGS reference implementation serving as a methodological performance ceiling. Such testing should be stratified by surface category: central sculptural detail, broad architectural geometry, material stratification, optically complex surfaces such as water, and high-contrast lighting zones including deep shadow and sky-adjacent regions. The figures suggest that tool choice and pose source affect these categories differently: PostShot shows marginally sharper statue rendering at matched parameters, while Lichtfeld Studio shows greater consistency on flat low-relief surfaces, and a single global quality score would be insufficient to characterize the documentary consequences of tool selection.

The third limitation concerns illumination and acquisition conditions. The dataset was acquired under direct solar illumination with pronounced cast shadows and high overall brightness. This combination produces strong luminance gradients across the sculptural and architectural surfaces and limits surface detail recovery in deep shadow zones beneath the contemporary canopy. Future acquisition protocols should investigate the effect of diffuse or overcast illumination on 3DGS reconstruction quality for heritage surfaces, as well as strategies for HDR image acquisition that extend the dynamic range available to the Gaussian optimization in high-contrast outdoor scenes. Additionally, moving elements in the scene introduce photometric inconsistencies in the training data that the current Gaussian optimization cannot resolve. Acquisition protocols for rapid documentation should address this through flight timing, image filtering, or robust loss functions.

The fourth limitation is archival. A 3DGS model is not yet a stable conservation record in the same sense as a georeferenced point cloud, orthophoto, or textured mesh. Viewer dependencies, evolving file formats, absent metadata standards, and the lack of embedded uncertainty visualization all constrain its immediate archival robustness. Future research should address preservation formats, metadata and paradata schemas, uncertainty disclosure, and interoperability with established heritage documentation environments including GIS, HBIM, and institutional digital repositories.

Finally, the accessible pipeline evaluated here is limited to aerial exterior acquisition. Smartphone LiDAR, handheld video, 360° imagery, and indoor image sequences can extend the same accessibility logic to interiors, narrow urban spaces, and sites where UAV deployment is impossible. A comparative evaluation of smartphone- and UAV-based accessible 3DGS against TLS ground truth, on the same site and with a common evaluation methodology, represents the most direct extension of this work.

6. Conclusions

This paper argues that accessible 3D Gaussian Splatting can make a meaningful contribution to the rapid documentation of cultural heritage by addressing a failure mode that remains insufficiently theorized in conservation practice: informational amnesia. When a site is destroyed without a record capable of sustaining future interpretation, the loss is not only material but documentary. Preventing that second loss requires timely acquisition of a sufficiently complete image record. Once captured, the same dataset can be processed into complementary derivatives: an SfM-MVS mesh for metric and geometric uses and a 3DGS scene for view-dependent visualization. The principal contribution evaluated here is the operational accessibility of the GUI-based pipeline, not an inherent superiority or speed advantage of 3DGS.

The case study of the Fontana d'Ercole at the Reggia di Venaria Reale indicates that a pipeline meeting these criteria is technically plausible. Using consumer UAV imagery, RealityScan for rapid graphical pose estimation, and Lichtfeld Studio for GUI-based Gaussian optimization, the study produced a photorealistic and navigable surrogate of a complex UNESCO World Heritage asset without commercial software or command-line interaction. The resulting models captured sculptural detail, water reflections, architectural depth, and material stratification across the four distinct temporal layers of the restored ensemble at a level that supports their use as rapid visual documentation products. The same dataset also generated a detailed textured mesh, and the two products should be retained together because they answer different documentary questions.

Several specific findings extend beyond the proof-of-concept demonstration. First, the quality of accessible 3DGS depends more on the upstream pose initialization than on the choice of Gaussian optimization tool. RealityScan's pose estimation produced nearly twenty times more sparse points than PostShot's internal alignment in a third of the time, and this difference in initialization density propagated directly into, though not established as the sole cause of, visually apparent scene coherence, gaps, and sharpness in the final model. For rapid documentation, a two-step graphical pipeline, free SfM tool for pose estimation followed by open-source 3DGS tool for Gaussian optimization, outperformed the simpler single-application alternative. Second, accessible 3DGS does not require a paid tool to meet the qualitative threshold for speditive visual documentation. Lichtfeld Studio, available as a free or donation-based Windows installer, produced more spatially consistent results than Jawset PostShot in the materially complex secondary architectural area of Figure 8, where PostShot introduced Gaussian floaters and an insufficiently resolved foreground pavement. The processing time advantage of PostShot at matched parameters is real, but the visual-quality advantage is not consistent across scene types.

The ethical conclusion is equally important. The Fontana d'Ercole's material stratification provides a meaningful test of the recognizability principle. The Lichtfeld Studio reconstruction preserves the chromatic and textural differentiation between the layers in the digital surrogate, providing evidence that, at least in this case, 3DGS does not homogenize the historical fabric into a seamless photorealistic surface. Yet the same results document the conditions under which this guarantee fails: Gaussian floaters introduce radiometrically inconsistent patches and deep shadow zones lose evidentiary support. These limitations do

not, however, exhaust the operational value of the models. Beyond their archival function, rapid 3DGS reconstructions offer a navigable surrogate through which conservators and structural engineers can make preliminary assessments of material degradation, identify zones of active loss or instability, and prioritize physical intervention before direct access is possible or safe. Emergency 3DGS models should therefore be understood as powerful but interpretive objects that serve a dual purpose, namely immediate operational support and longer-term archival record, and in both cases they should be accompanied by acquisition metadata, processing parameters, and, in future implementations, spatial uncertainty maps that make the limits of the photographic record visible rather than concealed by the model's photorealistic appearance. Accordingly, the 3DGS scene should be interpreted alongside, not instead of, the metrically controlled photogrammetric model.

The quantitative work remains to be done: metric validation against TLS or controlled survey data, formal image-quality benchmarking across accessible and reference implementations, systematic evaluation of pose-initialization strategies under varied acquisition conditions, and the development of archival standards for Gaussian radiance fields in heritage documentation contexts. The contribution of this paper is to show why that work is necessary and how 3DGS can form an additional interpretive layer within a broader documentation package. If the aim is to prevent informational amnesia, the preferred emergency deliverable is not a single "best" representation but an archival bundle containing source images, metadata and paradata, a metrically controlled photogrammetric product, and a 3DGS visualization, with the limitations and uncertainty of each made explicit. The Fontana d'Ercole case supports the operational feasibility of this complementary approach while remaining a qualitative proof of concept.

Author Contributions: Conceptualization, Methodology, Processing, Analyses, Writing—original draft preparation, A.M.; Writing—review and editing, Supervision, F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This publication is part of the project PNRR-NGEU which has received funding from the MUR–DM 630/2024.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. ICOMOS. *Heritage at Risk: World Report 2016–2019 on Monuments and Sites in Danger*; Hendrik Bäßler Verlag: Berlin, Germany, 2020.
2. Calantropio, A.; Chiabrando, F.; Sammartano, G.; Spanò, A.; Teppati Losè, L. Uav Strategies Validation and Remote Sensing Data for Damage Assessment in Post-Disaster Scenarios. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-3/W4*, 121–128. [CrossRef]
3. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *Photogram. Rec.* **2014**, *29*, 144–166. [CrossRef]
4. Kerbl, B.; Kopanas, G.; Leimkühler, T.; Drettakis, G. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Trans. Graph.* **2023**, *42*, 139. [CrossRef]
5. Clini, P.; Nespeca, R.; Angeloni, R.; Coppetta, L. 3D representation of Architectural Heritage: A comparative analysis of NeRF, Gaussian Splatting, and SfM-MVS reconstructions using low-cost sensors. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *XLVIII-2/W8-2024*, 93–99. [CrossRef]
6. Vandenabeele, L.; Häcki, M.; Pfister, M. Crowd-Sourced Surveying for Building Archaeology: The Potential of Structure from Motion (SfM) and Neural Radiance Fields (NeRF). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-2-2023*, 1599–1605. [CrossRef]
7. Chiabrando, F.; Di Lolli, A.; Patrucco, G.; Spanò, A.; Sammartano, G.; Teppati Losè, L. Multitemporal 3D Modelling for Cultural Heritage Emergency during Seismic Events: Damage Assessment of S. Agostino Church in Amatrice (RI). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-5/W1*, 69–76. [CrossRef]

8. Wilson, L.; Rawlinson, A.; Frost, A.; Hepher, J. 3D Digital Documentation for Disaster Management in Historic Buildings: Applications Following Fire Damage at the Mackintosh Building, The Glasgow School of Art. *J. Cult. Herit.* **2018**, *31*, 24–32. [CrossRef]
9. Apollonio, F.I.; Fantini, F.; Garagnani, S.; Gaiani, M. A Photogrammetry-Based Workflow for the Accurate 3D Construction and Visualization of Museum Assets. *Remote Sens.* **2021**, *13*, 486. [CrossRef]
10. Martin-Beaumont, N.; Nony, N.; Deshayes, B.; Pierrot-Deseilligny, M.; De Luca, L. Photographer-Friendly Work-Flows for Image-Based Modelling of Heritage Artefacts. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *XL-5/W2*, 421–424. [CrossRef]
11. Denard, H. (Ed.) *The London Charter for the Computer-Based Visualisation of Cultural Heritage, Version 2.1*; King's College London: London, UK, 2009. Available online: https://londoncharter.org/ (accessed on 30 July 2026).
12. Barzaghi, S.; Bordignon, A.; Gualandi, B.; Heibi, I.; Massari, A.; Moretti, A.; Peroni, S.; Renda, G. A Proposal for a FAIR Management of 3D Data in Cultural Heritage: The Aldrovandi Digital Twin Case. *Data Intell.* **2024**, *6*, 1190–1221. [CrossRef]
13. MrNeRF. *LichtFeld Studio*, version 0.5.2. Available online: https://github.com/MrNeRF/LichtFeld-Studio (accessed on 5 June 2026).
14. Postshot. *Jawset Visual Computing*, version 1.1.0. Available online: https://www.jawset.com/ (accessed on 5 June 2026).
15. UNESCO World Heritage Centre. *Residences of the Royal House of Savoy*; UNESCO: Paris, France, 1997. Available online: http://whc.unesco.org/en/list/823 (accessed on 5 June 2026).
16. Brandi, C. *Theory of Restoration*; Nardini Editore: Florence, Italy, 1963.
17. ICOMOS. *The Venice Charter for the Conservation and Restoration of Monuments and Sites*; ICOMOS: Charenton-Lebon, France, 1964.
18. ICOMOS. *The Nara Document on Authenticity*; ICOMOS: Charenton-Lebon, France, 1994.
19. Jokilehto, J. *A History of Architectural Conservation*; Butterworth-Heinemann: Oxford, UK, 1999.
20. Council of Europe. *Framework Convention on the Value of Cultural Heritage for Society (Faro Convention)*; Treaty Series No. 199; Council of Europe: Faro, Portugal, 2005.
21. EASA. *Commission Implementing Regulation (EU) 2019/947 of 24 May 2019 on the Rules and Procedures for the Operation of Unmanned Aircraft*; Official Journal of the European Union: Luxembourg, 2019.
22. Murtiyoso, A.; Grussenmeyer, P.; Suwardhi, D. Technical Considerations in Low-Cost Heritage Documentation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W17*, 225–232. [CrossRef]
23. Zhang, N.; Lan, X. Everyday-Carry Equipment Mapping: A Portable and Low-Cost Method for 3D Digital Documentation of Architectural Heritage by Integrated iPhone and Microdrone. *Buildings* **2025**, *15*, 89. [CrossRef]
24. *Epic Games Reality Scan*, version 2.0. Available online: https://www.realityscan.com/ (accessed on 5 June 2026).
25. Schönberger, J.L.; Frahm, J.-M. Structure-from-Motion Revisited. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA*; IEEE: New York, NY, USA, 2016; pp. 4104–4113.
26. McAvoy, S.; Tanduo, B.; Spreafico, A.; Chiabrando, F.; Rissolo, D.; Ristevski, J.; Kuester, F. An Archival Framework for Sharing of Cultural Heritage 3D Survey Data: OpenHeritage3D.org. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *XLVIII-2-2024*, 241–248. [CrossRef]
27. *3D Heritage Online Presenter*; Visual Computing Laboratory-ISTI-CNR: Pisa, Italy. Available online: https://3dhop.net/ (accessed on 5 June 2026).
28. Lanzara, E.; Catello, D.; Casciello, M. Scan to Hbim for Complex Reflective Metal Artefacts. 3d Digitisation and Restoration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLVIII-2/W1-2022*, 121–128. [CrossRef]
29. Debevec, P.; Yu, Y.; Borshukov, G. Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping. In *Proceedings of the Eurographics Rendering Workshop, Vienna, Austria, 29 June–1 July 1998*; Vienna University of Technology: Vienna, Austria, 1998; pp. 105–116.
30. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.T.; Ramamoorthi, R.; Ng, R. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *arXiv* **2020**, arXiv:2003.08934.
31. Croce, V.; Billi, D.; Caroti, G.; Piemonte, A.; De Luca, L.; Véron, P. Comparative Assessment of Neural Radiance Fields and Photogrammetry in Digital Heritage: Impact of Varying Image Conditions on 3D Reconstruction. *Remote Sens.* **2024**, *16*, 301. [CrossRef]
32. Li, Z.; Müller, T.; Evans, A.; Taylor, R.H.; Unberath, M.; Liu, M.-Y.; Lin, C.-H. Neuralangelo: High-Fidelity Neural Surface Reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023*; IEEE: New York, NY, USA, 2023; pp. 8456–8465.
33. MipMap Desktop-Professional Desktop 3D Reconstruction Software-MipMap. Available online: https://na.mipmap3d.com/products/mipmap-desktop (accessed on 30 July 2026).
34. Pix4D. Closing the Loop: Gaussian Splatting Comes to PIX4Dmatic. 11 May 2026. Available online: https://www.pix4d.com/blog/gaussian-splatting-pix4dmatic (accessed on 30 July 2026).

35. Yu, Z.; Chen, A.; Huang, B.; Sattler, T.; Geiger, A. Mip-Splatting: Alias-Free 3D Gaussian Splatting. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: New York, NY, USA, 2024; pp. 19447–19456. [CrossRef]
36. Huang, B.; Yu, Z.; Chen, A.; Geiger, A.; Gao, S. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. In *ACM SIGGRAPH*; Association for Computing Machinery: New York, NY, USA, 2024.
37. Lu, T.; Yu, M.; Xu, L.; Xiangli, Y.; Wang, L.; Lin, D.; Dai, B. Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2024.
38. Carbonara, G. *Avvicinamento Al Restauro. Teoria, Storia, Monumenti*; Liguori: Naples, Italy, 1997.
39. Blaeu, J. *Theatrum Statuum Regiae Celsitudinis Sabaudiae Ducis, Pedemontii principis, Cypri Regis. Pars Prima, Exhibens Pedemontium, et in eo Augusta Taurinorum, & Loca Viciniora, Vol. 1, Apud Heredes Ioannis Blaeu*; Joann. Blaev: Amstelodami, Germany, 1682; p. 138.
40. *Agisoft Metashape Professional*, version 2.3.1. Available online: https://www.agisoft.com/features/professional-edition/ (accessed on 5 June 2026).
41. PlayCanvas SuperSplat Editor. Available online: https://github.com/playcanvas/supersplat (accessed on 5 June 2026).
42. Gabara, G.; Sawicki, P. Multi-Variant Accuracy Evaluation of UAV Imaging Surveys: A Case Study on Investment Area. *Sensors* **2019**, *19*, 5229. [CrossRef] [PubMed]
43. Luhmann, T.; Chizhova, M.; Gorkovchuk, D. Fusion of UAV and Terrestrial Photogrammetry with Laser Scanning for 3D Reconstruction of Historic Churches in Georgia. *Drones* **2020**, *4*, 53. [CrossRef]
44. Lague, D.; Brodu, N.; Leroux, J. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 10–26. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.