The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W12-2026
11th Intl. Workshop 3D-ARCH "3D Virtual Reconstruction and Visualization of Complex Architectures", 10–12 February 2026, Ancona, Italy

Archival analog drawings for semantic segmentation of Roman Architectural Heritage using Deep Learning

María Belén Trivi [1]

[1] Dipartimento di Storia, Disegno e Restauro dell'Architettura, Sapienza Università di Roma, Italy - mariabelen.trivi@uniroma1.it

Keywords: Analog architectural drawings, Semantic segmentation, Deep Learning, Heritage documentation, Roman architecture.

Abstract

The present study aims to investigate whether the graphic code embedded in analogue architectural drawings—characterised by standardised textures and conventions—provides sufficient semantic information to support a simple, robust, and reproducible deep-learning-based segmentation approach, even under conditions of limited annotated data. The research focuses on a corpus of historical drawings preserved in the *Archivio dei Disegni e Fototeca of the Dipartimento di Storia, Disegno e Restauro dell'Architettura* at Sapienza Università di Roma, in which materials and construction techniques of Roman architectural heritage are represented through encoded graphic patterns and conventions. Starting from a limited set of 19 manually annotated drawings, a reproducible pipeline based on a U-Net architecture with a ResNet-34 backbone is developed, combining tiling strategies, data augmentation, and high-resolution inference. The results show high Overall Accuracy and Weighted IoU values, confirming the model's ability to interpret the implicit graphic language of the drawings, even in conditions of strong class imbalance and limited data availability. Inference on unseen drawings demonstrates an acceptable degree of generalisation, opening new possibilities for the automatic semantic digitisation of historical graphic archives. The study highlights the potential of analogue architectural drawings as a structured source of knowledge for artificial intelligence applications in the documentation, analysis, and conservation of the built heritage.

1. Introduction

1.1 Architectural drawings as encoded knowledge

The Roman palimpsest presents a highly stratified urban, architectural, and archaeological fabric, the result of a succession of historical periods, construction techniques, and diverse materials. This complex superposition of layers of history poses a significant challenge in the analysis, management, and preservation of this cultural heritage.
In this context, architectural drawings represent a crucial interpretative tool, as they not only record the geometry of heritage structures but also encode information about materials, construction systems, states of conservation, and historical phases through graphic textures and conventional symbols.
The discretization of this information contained in the drawings can become a valuable resource for training Artificial Intelligence (AI) classification algorithms, in which each class to be identified is represented by a distinct graphic code.
In recent years, a growing number of studies have undertaken the challenge of applying automated segmentation and classification methods based on AI to digital models of cultural heritage (Hess et al. 2017; Grilli et al., 2019; Croce et al., 2021; Samhouri et. al. 2022; Zhao et al., 2024; Wu et al., 2025; Dong et al., 2024). These studies have shown that the main difficulties lie in the specificity of each case, the heterogeneity of classification criteria, and the lack of annotated training data, which often hinder the replicability and generalization of AI methods across different contexts.
Currently, several scientific contributions focus on the automated generation of new vector drawings; however, they rarely address the exploitation of the semantic and material information embedded in existing drawings (Betsas and Georgopoulos, 2022; Betsas et al., 2024; Liu et al., 2017).
In this regard, the work of Trivi et al. (2024) has demonstrated that the use of existing digital vector drawings can significantly optimize the segmentation and classification of orthophotos and point clouds, thanks to their structured and interpretable nature.
The challenge addressed by the present study, however, is to employ as the primary training data a set of analog drawings, in which the annotative and representational components are intrinsically fused within the graphic artifact itself, unlike digital drawings where semantic information is explicitly organized into separate and structured layers. In particular, drawings preserved in the *Archivio dei Disegni e Fototeca of the Dipartimento di Storia, Disegno e Restauro dell'Architettura* (DSDRA) at *Sapienza Università di Roma* (Chiavoni, 2014; Docci, 2017) constitute a large, coherent, and highly structured corpus, encoding multiple layers of information through standardized graphic textures and symbolic conventions (Figure 1).

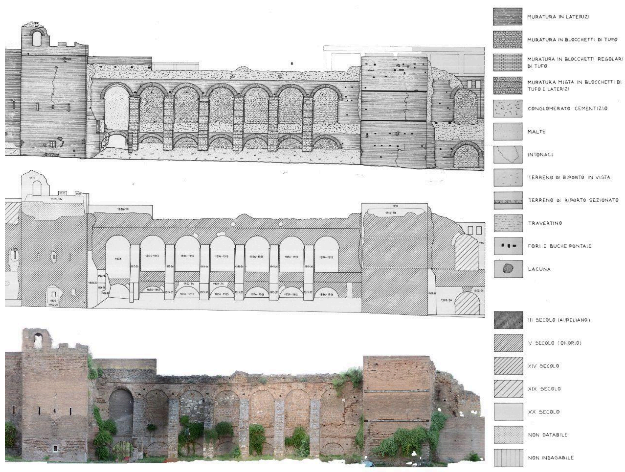


Figure 1. Archival drawing of the Aurelian Walls (1:100) showing graphic codes for construction materials (top), historical phases (centre), and legends (right). Course *Restauro dei Monumenti*, 1991–1992, Prof. G. Carbonara; drawings by

https://doi.org/10.5194/isprs-archives-XLVIII-2-W12-2026-487-2026

Santa and Caltabiano (*Archivio* DSDRA - Sapienza; Fondo CA, no. 95, plates 4–8). Bottom: orthophoto by the author.

Given the frequent scarcity of labeled data for AI training in heritage contexts, these documents represent a valuable and largely untapped resource for developing machine learning methods aimed at the interpretation and semantic segmentation of architectural heritage.

1.2 Aim of the study

The main objective of this study is to investigate whether the graphic code embedded in analog architectural drawings, characterized by standardized textures and conventions, provides sufficient semantic information to support a simple, robust, and reproducible deep-learning-based segmentation approach, even in conditions of limited annotated data.

Within this framework, the study proposes an experimental method capable of automatically identifying the graphic codes related to materials and construction techniques represented in analog drawings of Rome’s architectural and archaeological heritage. The method generates pixel-wise semantic masks from archival drawings, forming a structured semantic representation that can support subsequent processes of vectorisation and multi-scale analysis.

While the long-term objective of the research includes the projection of this semantic information onto orthophotos and three-dimensional data, the present contribution focuses on the 2D segmentation stage as a necessary preliminary step.

2. Methodology

2.1 Dataset selection and methodological framework

The adopted methodology is based on the selection of a coherent corpus of analogue architectural drawings preserved in the *Archivio* DSDRA at Sapienza Università di Roma, characterized by consistent graphic conventions and material codification.

The drawings are processed through an experimental semantic segmentation pipeline based on deep learning techniques, aimed at the automatic identification and classification of material-related graphic patterns at the pixel level. The output of this process consists of multi-class semantic masks representing construction materials and techniques encoded in the original drawings.

To handle high-resolution drawings, the pipeline employs convolutional neural networks (CNNs) and follows three main stages: preprocessing via tiling, model training, and post-processing for reconstruction.

The proposed workflow is designed to be replicable and adaptable to other archival drawing collections that adopt standardized graphic symbols and conventions, enabling its application to different architectural typologies and supporting heritage documentation and analysis workflows.

2.2 Graphic coding of materials and construction techniques in archival drawings

The *Archivio* DSDRA preserves a broad and diverse collection of architectural drawings, characterized by the symbolism and graphic conventions established by various Roman schools of representation.

In particular, the drawings produced within the course *Restauro Architettonico*, under the direction of Professor Giovanni Carbonara, share a common graphic expression, formally defined in the *Guida agli elaborati grafici* (Carbonara, 1990), conceived to standardize the visual communication of restoration projects.

Drawings created according to these conventions display a visual coding system capable of conveying information about the materials and construction techniques characteristic of Roman architecture through distinct textures and patterns. These representations, accompanied by legends, are clear and universally interpretable, allowing precise understanding by any reader (Figure 2).

This shared and structured graphic language enabled the creation of a class schedule, in which the categorization criteria precisely reproduce the materials and construction techniques defined in the original legends of the drawings.

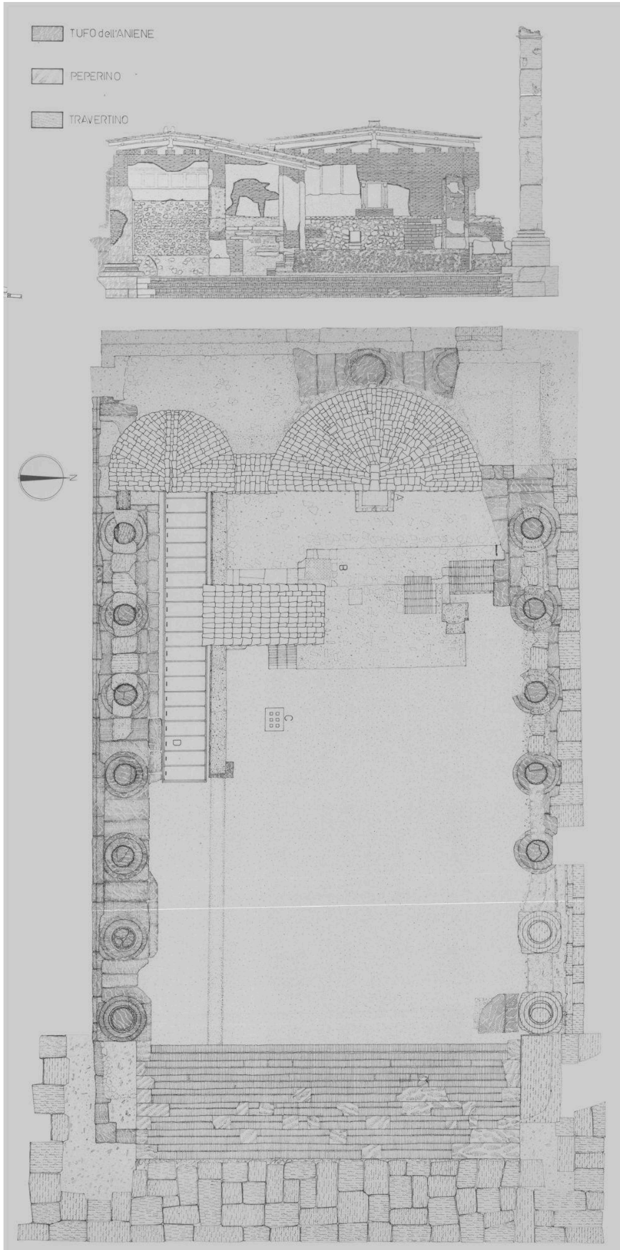
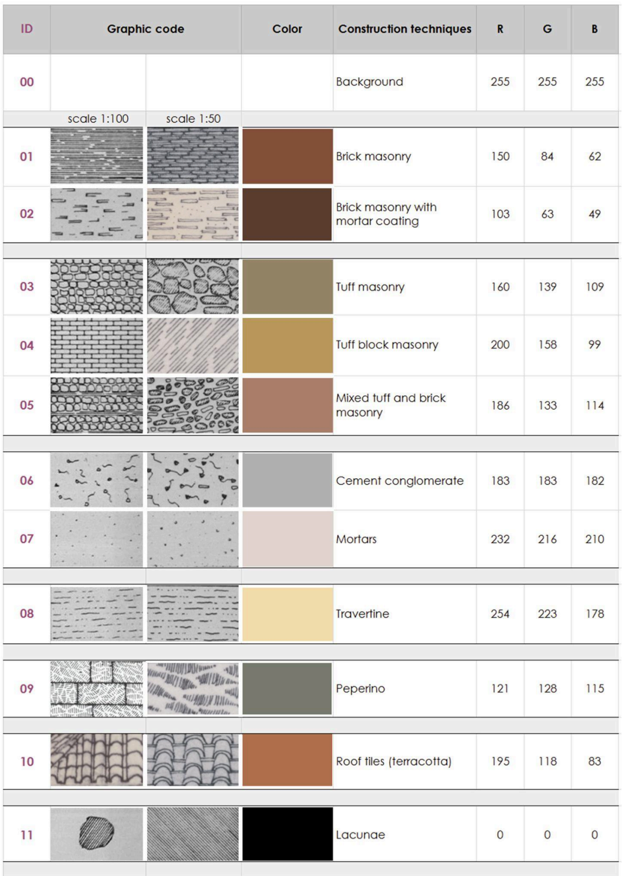


Figure 2. Archival architectural drawings of Temple A, Largo di Torre Argentina (Rome): elevation and roof plan. Produced within the course Restauro dei Monumenti, Academic Year 1987–1988, under the supervision of Prof. G. Carbonara. Drawings by students Abbondati and Marchetti.
(*Archivio DSDRA, Sapienza Università di Roma;* Fondo CA n. 76, plates nos. 5 and 6).

2.3 Definition of the semantic class schedule

Based on the legends included in the analogue drawings, a semantic schedule of classes was defined, representing the construction materials and techniques most frequently found in Roman architecture.
The selected classes correspond to materials naturally available in the Roman territory and therefore widely employed across different historical periods. The schedule includes: *Brick masonry; Brick masonry with mortar coating; Tuff masonry; Tuff block masonry; Mixed tuff and brick masonry; Cement conglomerate; Mortars; Travertine; Peperino; Roof tiles; Lacunae* (areas of material loss or absence).
This schedule defines the set of classes used in the automatic classification experiment. Each class represented by a graphic code was assigned a specific RGB color, used to generate labeled masks that translate the graphic language of the drawings into structured digital information (Figure 3).



ID	Graphic code (scale 1:100)	Graphic code (scale 1:50)	Color	Construction techniques	R	G	B
00				Background	255	255	255
01				Brick masonry	150	84	62
02				Brick masonry with mortar coating	103	63	49
03				Tuff masonry	160	139	109
04				Tuff block masonry	200	158	99
05				Mixed tuff and brick masonry	186	133	114
06				Cement conglomerate	183	183	182
07				Mortars	232	216	210
08				Travertine	254	223	178
09				Peperino	121	128	115
10				Roof tiles (terracotta)	195	118	83
11				Lacunae	0	0	0

Figure 3. Schedule of materials and construction techniques illustrating the standardized graphic codes adopted for each class at scales 1:100 and 1:50, together with the corresponding RGB color values used for the generation of semantic masks.

2.4 Evaluation metrics

Model performance was quantitatively assessed by comparing the predicted segmentation masks P with the manually annotated ground truth masks G.
The following metrics were computed:

- Overall Accuracy (OA), indicating the proportion of correctly classified pixels;
- Mean Intersection over Union (mIoU), computed as the average IoU across all classes;
- Weighted Intersection over Union (wIoU), where each class IoU is weighted by its pixel frequency, addressing class imbalance.

3. Experiment

3.1 Case studies and source drawings

The experimental dataset is composed of a selection of analogue architectural drawings referring to three representative architectural artefacts located within the urban fabric of Rome.
The first case study concerns Temple A of the Area Sacra di Largo Argentina, also known as the Temple of Juturna (or of the Lares Permarini), originally built in the 3rd century BC and extensively reconstructed during the 1st century BC.
The second case study focuses on a specific portion of Tower XXIII, constructed around the 5th century AD as a continuation of the Aurelian Walls (3rd century AD).
The third case study is the Church of San Urbano alla Caffarella, built in the 6th century AD over the remains of a 2nd-century AD Roman temple.
Despite differences in chronology, function, and architectural typology, the three case studies represent typical examples of Roman architecture characterised by a complex stratification of construction phases, materials, and techniques accumulated over time. Such stratified conditions are often difficult to interpret directly on-site, making architectural drawings a fundamental interpretative tool.

3.2 Archival drawings and graphic conventions

The selected artefacts were documented through architectural drawings produced by students of the course *Restauro Architettonico* coordinated by Giovanni Carbonara at Sapienza Università di Roma between the academic years 1985-1986, 1987-1988, and 1991-1992).
Each architectural complex is represented through multiple drawings - including plans, sections, elevations, and construction details - at scales ranging from 1:100 to 1:50 and 1:20.
Although the drawings were produced by different students and in different academic years, they share a consistent graphic code, in which construction materials and historical phases are represented through standardized graphic patterns. These patterns do not aim at mimetic realism but rather function as symbolic discretisations, enabling the reader to distinguish materials that are visually difficult to differentiate in reality.
This standardisation transforms the drawings into a structured visual language, making them particularly suitable for computational interpretation and machine learning–based analysis. A total of 19 drawings were selected and manually annotated, generating labelled masks for each semantic class (Figure 4). The dataset includes seven drawings of the Largo di Torre Argentina complex (four elevations, two sections, and one plan); eight drawings of the Aurelian Walls (two elevations, four sections, and two plans) and four drawings of the Caffarella site (four elevations).

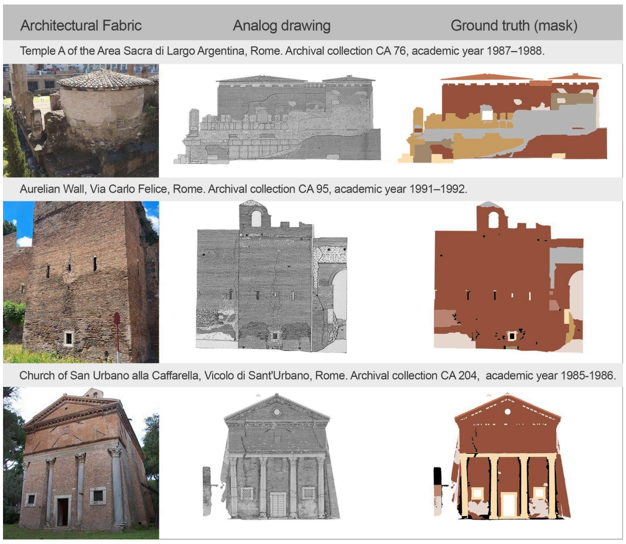


Figure 4. Selected case studies.

3.3 Ground truth generation and image preparation

The analogue drawings were digitised using a high-resolution scanner available at the Department of History, Representation and Restoration of Architecture of Sapienza, producing raster images with heterogeneous resolutions and dimensions.
To ensure consistency across the dataset, all drawings were resized to a common resolution of 7000 × 2500 pixels, corresponding to the largest scanned plate. The original proportions of each drawing were preserved, leaving white background areas where necessary.
Each semantic class was assigned a unique RGB value, and a total of 19 semantic masks were manually generated, one for each selected drawing, using Adobe Photoshop in indexed colour mode. This approach ensured pixel-perfect correspondence between the defined semantic classes and the mask values (Figure 4).

3.4 Tiling and data augmentation

Given the large size of the input images, a sliding-window tiling strategy was adopted. Each original image *Ior* and its corresponding ground truth mask *Morig* were divided into fixed-size patches of 512 × 512 pixels.
To avoid information loss at image boundaries, dynamic padding with a constant white background was applied to the right and bottom margins, ensuring that image dimensions were divisible by 512.
Data augmentation techniques were applied to mitigate overfitting, given the limited number of original samples. Using the Albumentations library, random rotations (90°), horizontal and vertical flips, and normalisation were applied to the extracted patches.
This strategy expanded the initial dataset of 19 high-resolution drawings into more than 1,300 training samples, significantly increasing the robustness of the training process.

3.5 Model architecture

The U-Net architecture was selected, given its effectiveness in pixel-wise semantic segmentation tasks.
The encoder was implemented using a ResNet-34 backbone pre-trained on ImageNet, enabling transfer learning and facilitating the extraction of meaningful visual features related to material textures and graphic patterns.
The decoder progressively reconstructs the spatial resolution through skip connections, producing a pixel-wise semantic prediction.
The final layer applies a Softmax activation function to assign each pixel to one of the 12 semantic classes.

3.6 Training configuration

The model was implemented in PyTorch using the Segmentation Models PyTorch (SMP) library and trained on a GPU-accelerated environment (NVIDIA Tesla T4).
Cross-Entropy Loss was adopted as the loss function, suitable for multi-class pixel-wise classification tasks.
The Adam optimiser (Kingma and Ba, 2015) was used with an initial learning rate of 1×10−4.
The network was trained for 20 epochs with a batch size of 4, and model checkpoints were saved periodically to monitor convergence.

3.7 Inference and reconstruction

During inference, full-resolution drawings cannot be directly processed by the network. Therefore, the same tiling strategy was applied to the input image, which was scanned using a sliding window of 512×512 pixels.
Each tile was independently processed by the trained U-Net, and the resulting local predictions were assembled through a stitching algorithm to reconstruct the full semantic segmentation map at the original resolution.
Padding margins added during preprocessing were removed in the final output

3.8 .Quantitative evaluation

The proposed pipeline was quantitatively evaluated on 19 full-resolution architectural drawings, each independently processed through the tiling-based inference and reconstruction workflow.
For each drawing, the predicted semantic mask was compared with the manually annotated ground truth using Overall Accuracy (OA) (the higher the better) , Mean Intersection over Union (mIoU) (the higher the better), and Weighted Intersection over Union (wIoU) (the higher the better).
Across the evaluated drawings, the model consistently achieved high overall accuracy values, ranging between approximately 96% and 99%. This result reflects the model's ability to correctly classify the majority of pixels, particularly for dominant material classes and background areas.
Mean IoU values varied across drawings, typically ranging between 71% and 80%, highlighting a heterogeneous behaviour across semantic classes. This variability is expected given the strong class imbalance and the visual similarity between certain construction materials represented through analogous graphic patterns.
Weighted IoU values remained consistently high (generally above 94%), confirming that the segmentation performance is particularly robust for the most frequently occurring materials in the drawings.
A class-wise analysis of the Intersection over Union reveals a clear distinction between dominant and rare material classes.
Structural materials characterised by well-defined and recurrent graphic patterns—such as brick masonry, regular tuff blocks, travertine, and background—achieved consistently high IoU values, often exceeding 0.85–0.95 across multiple drawings.
Lower IoU values were observed for rare or visually ambiguous classes, including mixed masonry types, cement conglomerates, and specific mortar representations.

These classes typically occupy limited pixel areas and present graphic similarities with neighbouring materials, making both manual annotation and automatic classification more challenging.
This behaviour confirms that the model effectively captures the semantic structure encoded in the architectural drawings, while highlighting the limitations imposed by class imbalance and dataset size.

3.9 Representative case studies

Figures 6-7 present representative examples of the segmentation results for selected drawings, including input images, manually annotated ground truth masks, and model predictions at full resolution.
The visual comparison confirms the quantitative metrics, showing accurate delineation of structural elements, continuity of material regions, and coherent reconstruction of complex stratigraphic conditions.
Most classification errors are localised at material boundaries or within small regions corresponding to rare classes, rather than affecting the overall interpretability of the drawingsks.

Drawing 1 Largo Argentina	
Overall Accuracy (OA)	0.9696 (96.96%)
Mean IoU (mIoU)	0.7306 (73.06%)
Weighted IoU (wIoU)	0.9524 (95.24%)

Table 1. Quantitative evaluation of construction technique segmentation for Drawing 1 (Largo di Torre Argentina).

Drawing 1_Largo Argentina		
Semantic Class	IoU	Pixels
0. Background	0.9881	11455858
1. Brick masonry	0.8911	904568
2. Brick with mortar coating	0.1596	43878
3. Tuff masonry	0.8276	454992
4. Tuff block masonry	0.9792	2530894
5. Mixed tuff and brick	0.1864	126475
6. Cement conglomerate	0.3803	361489
7. Mortars	0.6109	44179
8. Travertine	0.9485	1427512
9. Peperino	0.9898	24324
10. Roof tiles	0.8811	86780
11. Lacunae	0.9243	39051

Table 2. Class-wise segmentation metrics on the test sample, showing the IoU score and total pixel distribution for each semantic label for Drawing 1 (Largo di Torre Argentina).

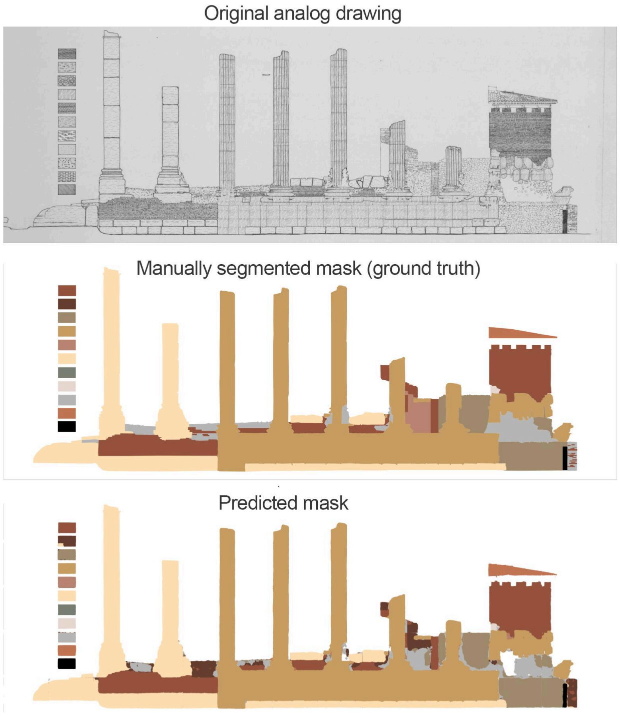


Figure 5. Drawing 1 Top: original drawing; middle: manually segmented mask (ground truth); bottom: predicted mask output.

Drawing 9 Aurelian Wall	
Overall Accuracy (OA)	0.9789 (97.89%)
Mean IoU (mIoU)	0.7146 (71.46%)
Weighted IoU (wIoU)	0.9670 (96.70%)

Table 3. Quantitative evaluation of construction technique segmentation for Drawing 9 (Aurelian Wall.

Drawing 9_Aurelian Wall		
Semantic Class	IoU	Pixels
0. Background	0.9889	13718197
1. Brick masonry	0.9584	2785879
2. Brick with mortar coating	0.2417	145925
3. Tuff masonry	0.9055	272750
4. Tuff block masonry	0.8542	24322
5. Mixed tuff and brick	0.8810	25869
6. Cement conglomerate	0.4869	91909
7. Mortars	0.7811	317174
8. Travertine	0.5383	33488
9. Peperino	0.9517	24719
10. Roof tiles	0.8185	26344
11. Lacunae	0.1692	33424

Table 4. Class-wise segmentation metrics on the test sample, showing the IoU score and total pixel distribution for each semantic label for Drawing 9 (Aurelian Wall).

https://doi.org/10.5194/isprs-archives-XLVIII-2-W12-2026-487-2026

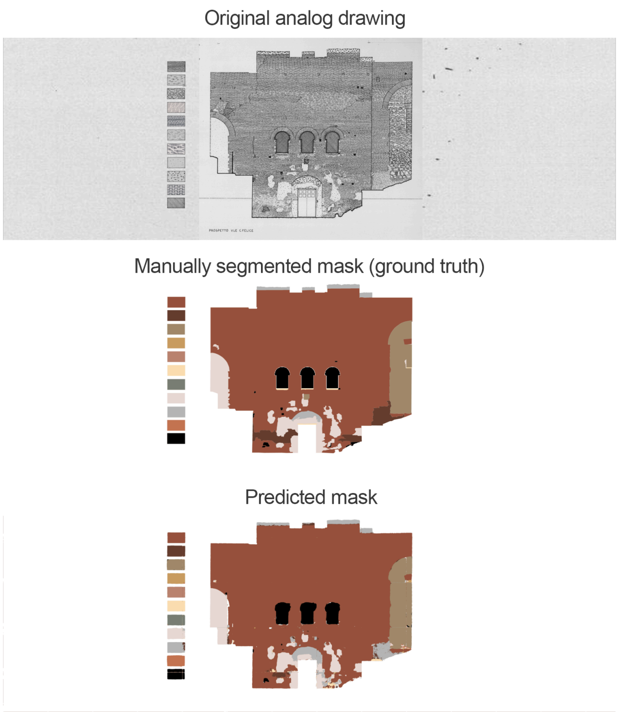


Figure 6. Drawing 9_Aurelian Wall. Top: original drawing; middle: manually segmented mask (ground truth); bottom: predicted mask output.

Drawing 17_Caffarella	
Overall Accuracy (OA)	0.9686 (96.86%)
Mean IoU (mIoU)	0.7105 (71.05%)
Weighted IoU (wIoU)	0.9462 (94.62%)

Table 5. Quantitative evaluation of construction technique segmentation for Drawing 17 (Caffarella)

Drawing 17_Caffarella		
Semantic Class	IoU	Pixels
0. Background	0.9931	10731361
1. Brick masonry	0.9485	3234536
2. Brick with mortar coating	0.6462	27563
3. Tuff masonry	0.3128	78574
4. Tuff block masonry	0.7766	29658
5. Mixed tuff and brick	0.8767	1062727
6. Cement conglomerate	0.2232	48952
7. Mortars	0.5832	594558
8. Travertine	0.9385	179329
9. Peperino	0.8745	26587
10. Roof tiles	0.9076	1351693
11. Lacunae	0.4446	134462

Table 6. Class-wise segmentation metrics on the test sample, showing the IoU score and total pixel distribution for each semantic label for Drawing 17 (Caffarella).

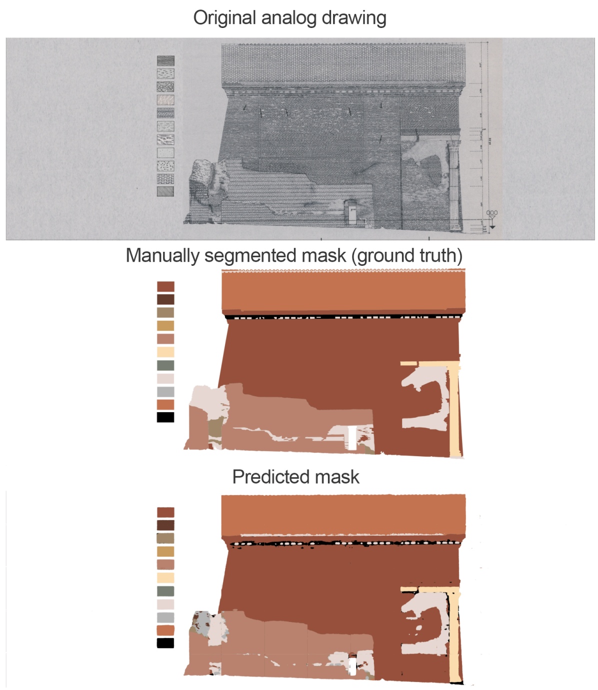


Figure 7. Drawing 17_Caffarella. Top: original drawing; middle: manually segmented mask (ground truth); bottom: predicted mask output.

3.10 Ianference on Unseen Architectural Drawings: Evaluating Model Generalizations

To verify the robustness of the model when applied to unseen data and to assess its potential for the automatic digitization of historical archives, an inference test was conducted on a drawing external to the training and validation datasets (.
The selected case study was a representation of the façade of the Church of San Teodoro al Palatino, drawn at a scale of 1:50 between 1990 and 1991. Unlike the training data, this document did not include a previously generated manual segmentation mask (ground truth), thereby simulating a real-world application scenario in which no human intervention is available.
The model processed the high-resolution image using the previously described tiling and reconstruction pipeline. The visual results demonstrate that the predicted segmentation largely respects the architectural linework boundaries and the background, producing relatively sharp contours for the different material classes. Overall, a good level of semantic coherence is observed; however, some confusion remains between materials with similar graphic representations, such as travertine and mortar.
Despite these limitations, the model is able to correctly identify and segment the main masonry areas without confusing them with the background, even though this specific building was not included in the training set. This indicates a certain degree of generalization capability when applied to unseen architectural drawings, while highlighting the challenges posed by visually similar material patterns.
This result is particularly significant considering the limited size of the training dataset, which consists of only 19 manually annotated drawings (Figure 8).

Figure 8. Top: Inference on an unseen drawing of the façade of the Church of San Teodoro al Palatino (scale 1:50). Bottom: semantic predicted mask.

4. Conclusions

This paper presented a deep learning–based workflow for the semantic segmentation and digitisation of archival architectural drawings, focusing on analogue representations characterised by a highly standardised graphic code. Although the proposed approach relies on a consolidated architecture (U-Net), the contribution of this work lies in its application to a scarcely explored domain: historical architectural drawings produced through consistent representational conventions and preserved in archival collections.

One of the main challenges addressed by this research is the limited availability of manually annotated training data, a condition that is typical when dealing with historical graphic archives. Despite the reduced size of the training dataset (19 manually annotated drawings), the results demonstrate that the model is capable of achieving a high level of semantic coherence, particularly for dominant material classes, and of generalising to unseen drawings produced in different years and by different authors. The inference experiment performed on an external drawing without a manually created ground truth further confirms the robustness of the approach in realistic application scenarios, where manual annotation is often laborious.

An important aspect of this work is that the achieved results were obtained using a non-complex and widely adopted U-Net architecture, without introducing ad hoc or computationally expensive model modifications. This design choice was intentional, as it reinforces the replicability and transferability of the proposed method to other research contexts, including those with limited computational resources or technical expertise. The study also highlights the effectiveness of deep learning techniques in small-data conditions. Through appropriate preprocessing strategies, including tiling and data augmentation, a relatively small number of training samples proved sufficient to train a model capable of learning and reproducing complex graphic rules embedded in architectural drawings. This aspect is particularly relevant in the heritage domain, where the creation of large, fully annotated datasets is rarely feasible.

From a scalability perspective, architectural archives contain an enormous quantity of analogue drawings that lack semantic labelling. The ability of the proposed model to generate visually coherent and semantically meaningful segmentation masks opens new possibilities for the large-scale automatic cataloguing and quantitative analysis of archival material - tasks that would be unmanageable if performed manually.

In terms of accuracy assessment, this study demonstrates that pixel-based metrics alone are insufficient to evaluate the equivalence between an analogue drawing and its digital counterpart. While high Overall Accuracy values are partly influenced by class imbalance (e.g. the predominance of background), metrics such as mean Intersection over Union (mIoU) provide a more reliable measure of segmentation quality. From a disciplinary perspective, equivalence should not be interpreted as a perfect pixel-wise correspondence, but rather as the correct identification and separation of materials, construction techniques, and stratigraphic information necessary for architectural analysis, documentation, and interpretation.

Geometric resolution plays a crucial role in preserving the precision of the original representation. In this work, high-resolution scans and a tiling-based inference strategy were adopted to prevent loss of detail and to maintain the original scale relationships of the drawings. This ensures that the semantic information extracted by the model remains consistent with the architectural intent and the survey accuracy of the original documents.

From an application standpoint, the digitisation of these drawings enables faster access to structured, queryable, and reusable semantic information, providing a foundation for further analyses, including large-scale studies of materials and construction techniques in historic urban fabrics.

Furthermore, this study represents the first systematic phase of a broader research process, involving the critical review of archival material and the construction of a reasoned and coherent material abacus derived directly from the graphic legends of the drawings. This preliminary step is fundamental, as it formalises the graphic knowledge embedded in analogue representations and establishes the semantic foundation necessary for future extensions of the workflow.

Future work will focus on expanding the dataset, refining material classes affected by visual ambiguities, exploring semi-supervised or active learning strategies to further reduce the need for manual annotation, and experimenting with the projection of predicted segmentation masks onto point clouds and orthophotos to support integrated 2D - 3D semantic workflows.

While the proposed system does not aim to replace expert interpretation, it provides a robust and scalable tool for the preliminary semantic digitisation of archival drawings, significantly reducing manual effort and enhancing the value of the vast corpus of analogue architectural documentation for conservation, research, and heritage management.

References

Betsas T., Georgopoulos A., Doulamis A., 2024: Incad: 2d vector drawings creation using instance segmentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-2/W4-2024, 2024, 65–72. https://doi.org/10.5194/isprs-archives-XLVIII-2-W4-2024-65-2024

Betsas T., Georgopoulos A., 2022: 3D Edge detection and comparison using four-channel images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII- 2/W2-2022, diciembre 2022, 9–15. https://doi.org/10.5194/isprs-archives-XLVIII-2-W2-2022-9-2022

Carbonara, G., 1990: *Restauro dei monumenti. Guida agli elaborati grafici*. Liguori Editore, Napoli.

Chiavoni, E., 2014: Drawings on paper. Digital historical archives of the former Radaar Department at the University Sapienza School of Architecture in Rome. *SCIRES-IT*, 4(2), 1–12. https://doi.org/10.2423/i22394303v4n2p117

Croce, V., Caroti, G., De Luca, L., Jacquot, K., Piemonte, A., Véron, P., 2021: From the semantic point cloud to heritage-building information modeling: A semiautomatic approach exploiting machine learning. *Remote Sensing*, 13(3), 461. https://doi.org/10.3390/rs13030461

Docci, M., 2017: Storia, disegno e restauro nei materiali d'archivio. Un patrimonio da gestire e condividere. In: Prescia, R. (ed.), *RICerca/REStauro. Sezione 4. Valorizzazione e gestione delle informazioni*. Edizioni Quasar, Roma, 826–837.

Dong, Y., Hou , M., Zhou, M., 2024: MP-DGCNN for the semantic segmentation of Chinese ancient building point clouds. *Research Square*. https://doi.org/10.21203/rs.3.rs-3957836/v1

Grilli, E., Özdemir, E., Remondino, F., 2019: Application of machine and deep learning strategies for the classification of heritage point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W9, 447–454. https://doi.org/10.5194/isprs-archives-XLII-2-W9-447-2019

Hess, M. R., Petrovic, V., Kuester, F., 2017: Interactive classification of construction materials: Feedback driven framework for annotation and analysis of 3D point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 343-347. https://doi.org/10.5194/isprs-archives-XLII-2-W5-343-2017

Liu, C., Wu, J., Kohli, P., Furukawa, Y., 2017: Raster-to-vector: Revisiting floorplan transformation. *Proceedings of the IEEE international conference on computer vision*, 2195-2203. https://doi.org/10.1109/ICCV.2017.241

Samhouri, M., Al-Arabiat, L., Al-Atrash, F., 2022. Prediction and measurement of damage to architectural heritages facades using convolutional neural networks. *Neural Computing and Applications*, *34*(20), 18125-18141. 18141. https://doi.org/10.1007/s00521-022-07461-5(0123456789().,-volV)(0123456789.,-().volV)

Trivi, M.B., Mazzacca, G., Griffo, M., Malek, S., Battisti, R., Remondino, F., Bianchini, C., Chiavoni, E., 2024: An innovative approach for the semantic segmentation of surveyed building facades leveraging on architectural drawings. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2, 445–451. https://doi.org/10.5194/isprs-archives-XLVIII-2-2024-445-2024

Wu, X., Yu, Y., Li, Y., 2025: A damage detection network for ancient murals via multi-scale boundary and region feature fusion. *Heritage Science*, 13(1), 82. https://doi.org/10.1038/s40494-025-01546-9

Zhao, J., Yu, H., Hua, X., Wang, X., Yang, J., Zhao, J., Xu, A., 2024: Semantic segmentation of point clouds of ancient buildings based on weak supervision. *Heritage Science*, 12(1), 232. https://doi.org/10.1186/s40494-024-01232-8