AN INNOVATIVE APPROACH FOR THE SEMANTIC SEGMENTATION OF SURVEYED BUILDING FACADES LEVERAGING ON ARCHITECTURAL DRAWINGS

M.B Trivi[1], G. Mazzacca[2,3], M. Griffo[1], S. Malek[2], R. Battisti[2], F. Remondino[2], C. Bianchini[1], E. Chiavoni[1]

[1] Dipartimento di Storia, Disegno e Restauro dell'Architettura, Universita' la Sapienza, Roma, Italy
Email: (mariabelen.trivi, marika.griffo, carlo.bianchini, emanuela.chiavoni)@uniroma1.it

[2] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Email: (gmazzacca, smalek, rbattisti, remondino)@fbk.eu

[3] Dept. Mathematics, Computer Science and Physics, University of Udine, Italy

Commission II

KEY WORDS: Architectural drawing, architectural built heritage, building materials recognition, construction techniques recognition, machine learning, deep learning

ABSTRACT:
Traditionally, drawing products created from 3D surveying activities have been the universal medium of communication used by architects. This has resulted in a vast repository of graphic documentation that serves as a testament of the architectural heritage. The embedded information found in elevations, plans and sections holds considerable value, and it can be seamlessly integrated into the intricate graphics produced during large-scale data acquisition processes. The core objective of this research is to investigate how the information coming from the large amount of existing architectural technical drawings can support 3D heritage classification processes and avoid time-consuming annotation of materials and construction techniques of historical building facades. Starting from available sets of drawings, AI-based methodologies are applied for the annotation of orthoimages and point clouds in order to obtain a predictive model that can recognize classes of materials and construction techniques in a large amount of data. The predicted classes also allow the automatic creation of vector drawing representing the facades of new buildings, providing a novel tool to facilitate the processes of analysis and conservation of architectural heritage.

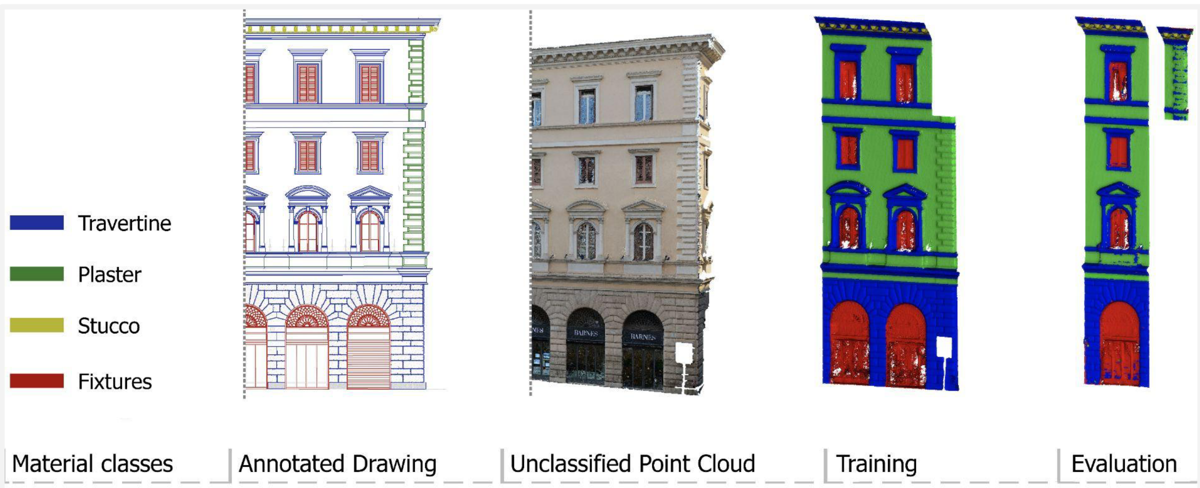

Figure 1: The proposed 3D classification process based on technical architectural drawings to avoid data annotation and support automatic sematic segmentation of historical building facades.

1. INTRODUCTION

The information stored in architectural drawings can be a powerful source to be integrated with Artificial Intelligence (AI) methods to semantically enrich 2D/3D digital heritage. In particular, the challenge of 3D heritage classification has been open for some years with few automated processes available (Poux et al., 2017; Grilli & Remondino, 2019, Grilli & Remondino, 2020, Matrone et al., 2020; Pierdicca et al., 2020; Yang et al., 2023). Digital heritage often presents peculiar characteristics that hamper replicability and generalization of AI methods across different scenarios. Especially in the architectural field, the presence of unbalanced classes, the size and uniqueness of the monuments as well as the level of detail that architectural analysis of buildings usually requires pose hard challenges for AI-based classification approaches. Another severe drawback consists in the lack of sufficient training data which, linked to the previous challenges, contributes to making these kinds of approaches even more challenging. In this scenario, the use of previously produced architectural drawings to generate a sizeable amount of training data can be very useful to boost the implementation of AI methods in the field.
Knowledge and conservation of architectural heritage presents a growing number of modalities and techniques. The integration of semantic information into 2D and 3D data through automated classification methods has been a crucial advancement in digital production, aiming to enhance the analysis, monitoring, and conservation methodologies of cultural assets. In recent years, advances in Machine Learning (ML) and Deep Learning (DL) algorithms had a decisive impact on performing these tasks. Specifically, the classification of building facades has been explored in various studies (Korc and Forstner, 2009; Tyleček

and Šára, 2013; Hess et al., 2017). In Teboul et al. (2013), the combination of Shape Grammars and Reinforcement Learning achieved positive results in the segmentation of 2D facade images based on compositional architectural elements. Valero et al. (2018) developed innovative algorithms for the automatic segmentation of individual masonry units through the use of geometric and colour data acquired by laser scanning devices. Grilli et al. (2019) compare the precision metrics obtained in the segmentation of point clouds from various archaeological and architectural heritage cases using Machine and Deep Learning. Generally, the process needs manual labelling of classes in the learning phase as unsupervised methods are still not available for heritage 3D data. In our case, the idea focuses on exploring the possibility of replacing the manual annotation operation with automatic labelling using the thematic information present in the existing vector drawings of the selected case studies. At the same time, a data fusion is performed involving the radiometric features of the orthoimage and the geometric features of DSM and SHP files, aiming to achieve the automatic creation of new vector polygons.

1.1 Objectives of the study

The aim of the presented work is two-fold:

- to investigate how the information coming from a large amount of existing architectural technical drawings can support supervised 3D classification processes and avoid time-consuming annotation of materials and construction techniques of historical building facades: this can be accomplished using the information that comes from the drawings and projecting these kinds of annotations onto point clouds and orthoimages. In this way, it is possible to generate a significant amount of annotated data and train Machine Learning (ML) / Deep Learning (DL) models, speeding up the semantic segmentation of architectural-related point clouds and orthoimages.
- to test whether the use of architectural technical drawings as input in the training phase allows the automatic creation of new vector files representing the façades of other buildings for which no drawing exists.

The automated recognition of materials and construction techniques and the automatic obtaining of vector representations can be an extremely valuable tool for preserving, restoring and monitoring the abundant architectural heritage found throughout urban landscapes.

2. METHODOLOGY

2.1 Materials and construction techniques

The investigation used CAD technical drawings for annotation and both unclassified point clouds and orthoimages from photogrammetric surveys for training and validation. A significant quantity of two-dimensional CAD technical drawings has been collected throughout the years. The variety of materials and construction techniques is represented using recurrent textures and patterns. This information can help AI methods in their task of automatically recognising classes within building facades. To facilitate this, a proper schedule was conceived, identifying the diverse categories of materials and construction techniques within this architectural typology. The schedule summarises the general categories of materials and construction techniques commonly found in most palaces. Specifically, during the learning phase, we selected buildings with more complexity and a greater variety of categories to help the algorithm acquiring as much information as possible simplifying the subsequent prediction task. In this way, the technical drawings of the most representative facades were schematised by dividing the layers based on the classes represented in the schedule. This resulted in the labelled information used to train the ML/DL models to perform their tasks.

Some architectural studies (Pallottino, 1990; 1992), specify in detail the various construction techniques and materials that constitute the elements present in the facades of this architectural typology. In this way, by taking these works as a reference and conducting on-site observations of selected and analogue buildings in the urban context of Rome, a synthesised schedule of the materials and construction techniques most frequently found on their façades was carried out (Figure 2). As a result, *travertine, plaster, stucco and fixtures* were identified as classes of materials. The *fixtures* class, in particular, included the internal elements of the fixtures, such as glass, shutters or grilles, which involve a greater complexity due to the presence of reflections in the glass or the variation of the position of the shutters. These aspects can be a significant obstacle in the development of an automated process, so it was decided to consider this class autonomously, avoiding possible errors that could hinder the experimental phase.

Regarding construction techniques, the identified classes include *rusticated travertine, carved travertine, particular elements travertine, smooth plaster, rusticated plaster, carved stucco*, and once again, *fixture* as a general category.

The fundamental aspect of constructing this schedule is related to the identification of the geometric and morphological characteristics through which these construction techniques are represented in the various two-dimensional drawings that exist of these buildings. In this regard, geometric features derived from the representation of this architectural language are added to the radiometric features involved in the machine learning process.

01_Material	02_Construction Technique
A01. Travertine	B01a Rusticated
	B01b Carved
	B01c Particular Elements
A02. Plaster	B02a Smooth
	B02b Rusticated
A03. Stucco	B03a Carved
A04. Fixtures	B04a Fixtures

Figure 2: Schedule of materials and construction techniques.

2.2 3D semantic segmentation pipeline

The first proposed pipeline (Figure 3) leverages the drawing information to generate labelled 3D annotated data of the facades, eventually aiming at training a neural network to perform the semantic enrichment of not yet annotated facades' point clouds.

At first, we generate masks comprising the classes that represent the materials and the construction techniques by reprojecting the CAD vector drawings onto orthoimages. Subsequently, we align

the generated masks to the palace facades' photogrammetric point clouds to reproject the classes of choice onto the point clouds via interpolation. Following the reprojection, the annotated point clouds are divided into training and evaluation sets. These sets are then fed into the deep learning algorithm together with sensor- and geometric-based features, which help the network to learn how to identify the selected classes. Eventually, we use the generated models to predict the classes on new unseen data.
The prediction phase relies on Point Transformer (PT) (Zhao et al., 2021), a method that accomplishes fundamental scene understanding tasks by leveraging the self-attention operator. The PT architecture can improve its capabilities for tasks involving large scenes with millions of points by implementing the self-attention mechanism locally. In addition, the network can employ class-balancing techniques to account for the existence of underrepresented classes.

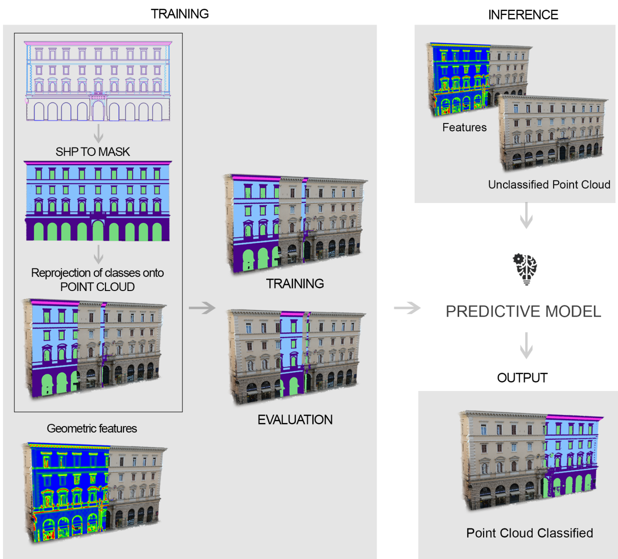

Figure 3: The proposed 3D semantic segmentation pipeline.

.

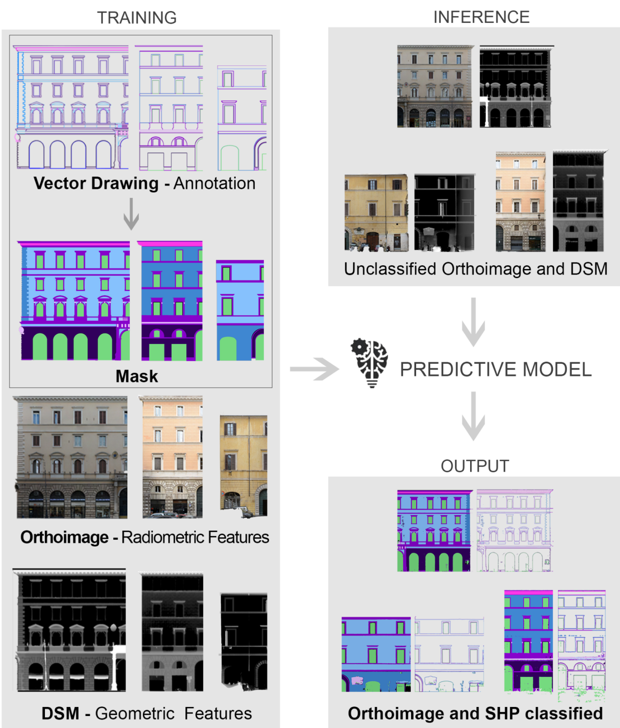

Figure 4: The 2D segmentation and vectorisation pipeline.

2.3 2D semantic segmentation and vectorisation pipeline

The second proposed pipeline (Figure 4) proceeds in the same manner as the first, converting the drawings into masks representative of the selected classes. However, in this case, the end goal is to generate vector-based drawings, therefore the semantic enrichment is carried out directly at 2D level, with no information passed to the 3D point cloud. At the same time, to take advantage of the 3D reconstruction, the geometric information is transferred into the network by using the available facades’ Digital Surface Model (DSM). After preparing the masks with the annotated classes, orthophoto and DSM are injected into the deep learning network (Figure 5). By doing so, we aim to enhance the algorithm's results combining depth information from the DSM with features extracted from the orthophoto.

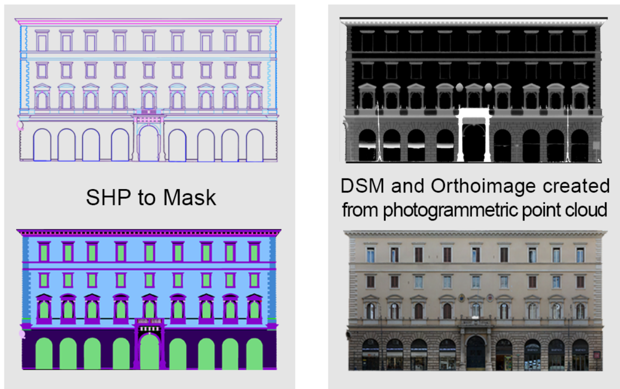

Figure 5: Data preparation for the training phase.

The prediction step relies on DeepLabV2 (Adam et al. 2021), which is an improved version of the origin DeepLab (Chen et al. 2017). It is based on ResNet152 (He et al. 2015) as the backbone network which was trained using the ImageNet database. The main advantage of DeepLab is the use of the dilated convolutions concept, also called *atrous*. This new concept allows the model to capture wider context in images without reducing their resolution. It has also demonstrated robustness against various image perturbations and high-class variance, among other factors.
The predicted masks are then used to generate a contour map of the classes’ area borders and, from that, automatically create the vector-based drawings of the predicted classes, already divided into multiple shapefiles, one for each represented class.

2.4 Evaluation metrics

In order to objectively evaluate the segmentation results of both pipelines, we decided to use the following metrics, calculated by comparing the predictions’ results to the ground truth data:
The Overall Accuracy (the higher the better):

$$\mathrm{OA} = \frac{\sum_{i=1}^{C} TP_i}{N} \quad (1)$$

The mean of Intersection over Union (the higher the better):

$$\mathrm{mIoU} = \frac{1}{C}\sum_{i=1}^{C} \frac{TP_i}{TP_i+FP_i+FN_i} \quad (2)$$

The weighted Intersection over Union (the higher the better):

$$\mathrm{wIoU} = \sum_{i=1}^{C} \frac{w_i \times TP_i}{TP_i+FP_i+FN_i} \quad (3)$$

with
C: the total number of classes.
N: the total number of samples.
Ni: the total number of samples of class “i” (i = 1, …, C).
True positive (TPi): the number of correctly predicted samples of

class "i".
False Positive (FPi): the number of samples of other classes (not class "i") which are predicted as class "i".
False Negative (FNi): the number of samples of class "i" that is wrongly predicted to other classes.
wi: the weight of the class "i" which represents the sample percentage of class "i" over all the samples.

3. EXPERIMENTS

3.1 The case studies

To test the implementation of the proposed pipelines, a large amount of 2D graphic documentation of renaissance and neo-renaissance palaces in Rome is considered. These palaces present an interesting variety of materials and construction techniques, all clearly distinguishable through the available architectural drawings. Consequently, this documentation stands out as a coherent and standardised representation of a specific architectural language, providing a perfect basis for the experimental methods we employ. The data acquisition from these buildings (Figure 6) was carried out using photogrammetric survey campaigns. High-resolution orthoimages and digital surface models (DSM) were generated for each of the surveyed buildings.

Figure 6: Facades orthoimages (left) and photographs of the urban context (right) in which the selected palaces are located.

The Palazzo Severo, also known as Palazzo della Pontificia Accademia Ecclesiastica, is an example of those buildings that in the 19th century had their façade renovated, following the neoclassical canons with typical neo-Renaissance decorations. Its facade spans approximately 42 meters in length and 23 meters in height, while the photogrammetric point cloud obtained during the survey consists of around 27,4M points.
The Palazzo in Piazza Rondanini is a simple example of a Renaissance building built in the 16th century. With a height of 14 meters and a length of around 19 meters, it showcases a rusticated travertine arch at the entrance, sculpted travertine moulding details, and a "false" rusticated plaster at level 0. The point cloud of this palace comprised 42,7M points ca.
The Palazzo Mazzetti was constructed between 1855 and 1860, featuring distinctive rusticated travertine, lintel-framed windows in stone, and a false rusticated plaster on the upper levels. With a height of approximately 23 meters and a length of 24 meters, it yielded a dense point cloud comprising around 40,8M points.

3.2 3D point cloud semantic segmentation results

For all three palace facades', annotated masks were generated from a simplified version of their technical drawings. Both the shapefiles-to-mask reprojection and masks-to-shapefiles generation procedures are performed via Python scripts, utilising the libraries Rasterio, Shapely and Fiona. As explained in Section 3.2, the obtained masks in projected onto the point cloud through interpolation by using the CloudCompare software interpolation tool. A portion of each building is then allocated for the training process, aiming to provide the algorithm with a sufficient quantity of data for efficient learning and effective model development. Approximately 30% of the dataset was allocated to the training set, 20% to the evaluation set, and 50% to the test set. This choice was taken in light of generalization purposes. The algorithm was assisted in converging to a better result by using selected geometric features calculated at different *radii* (Table 1) as well as sensor-based features (RGB radiometric values). The network was trained for 100 epochs with a batch size of 32, using the Adam optimizer method (Kingma and Ba, 2015) and subsampling the input point clouds at 1 cm. The optimizer method used for the training was the Adam method.

Geometric Feature	*Radius (m)*
Normal change rate	0.05
Mean curvature	0.1
Normal change rate	0.1
Mean curvature	0.2
Normal change rate	0.3
Roughness	0.3
Sphericity	0.3
Verticality	0.3

Table 1: Geometric features calculated at different r*adii.*

The pipeline was adopted to predict both materials and construction techniques classes, with the only difference being that we used the materials label as a feature to train the construction techniques model. The reasoning behind this decision is that since the materials and construction techniques are related in a "one to many" manner, following the intervention of a manual operator who corrects the inconsistencies in the materials' prediction, the materials' labelled data can be used to assist the network in the subsequent phase.
As visually noticeable in Figure 7, the facades were for the most part correctly predicted for both materials and construction techniques. Even very under-represented classes like *stucco* in the materials' prediction and the class *particular elements* in the construction techniques prediction have been detected and segmented quite successfully. Looking more closely at the predictions' masks, we can see that the algorithm had a harder time distinguishing among classes with very similar geometric or radiometric characteristics, as noticeable on the Palace of Piazza Rondanini's facade between two different variants of plasters (*rusticated* and *smooth*) or the confusion between *carved travertine* and *particular elements* on Palazzo Severo's facade in Figure 7.

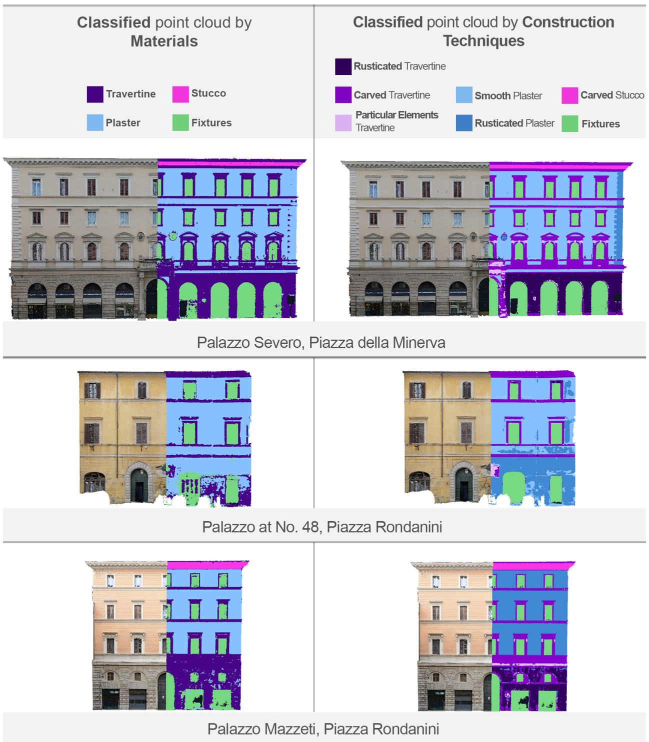

Figure 7: Predicted point clouds for material (left) and construction techniques (right).

Tables 2 and 3 present some metric for the abovementioned tests. The metrics were calculated on each facade's prediction, causing the issue of not having a complete representation of all the classes in every facade. This problem didn't affect the calculation of the OA, not being influenced by it, and wIoU, being the weight of the not represented classes equal to 0. The mIoU was instead estimated by taking the class that was not represented out of the computation.

Dataset	*OA %*	*mIoU %*	*wIoU %*
Palazzo Severo	93.32	88.30	87.52
Palazzo Mazzetti	93.66	88.19	88.18
Palazzo at n. 48	86.12	72.31	77.26

Table 2: Quantitative evaluation for materials predictions.

Dataset	*OA %*	*mIoU %*	*wIoU %*
Palazzo Severo	93.48	79.00	90.12
Palazzo Mazzetti	92.87	78.69	88.71
Palazzo at n. 48	88.12	77.10	80.95

Table 3: Quantitative evaluation for construction techniques predictions.

The metrics show that, even if encouraging results are visible through visual inspection, there still remains a large number of inaccuracies, particularly in the facade prediction for the Palace of Piazza Rondanini. This was to be expected given the type of data we used, facades of buildings, comprising several classes with no recognisable geometric signature. Because of that, it was not possible to take full advantage of the 3D nature of the point clouds.

3.3 2D semantic segmentation and vectorisation results

After creating the annotated masks representing materials and construction techniques of the facades, we divided our dataset into 2 subsets, a training set and a testing set. In order to fairly evaluate the 2D semantic segmentation results and generate models with a higher understanding of the scene, each image was divided into 2 equal parts: the left part of the image (50%) was used for the training set while the other part was used for the testing set. The training dataset was created by cropping randomly over the training set (the left half of each image). Each crop (patch) was of size 512x512 px. The total number of training samples was 9000 samples. Half of them were created using data augmentation (horizontal flipping).
The proposed model was implemented by using the TensorFlow library. The batch size was set to 4 with the maximum number of epochs equal to 100. The optimizer method used in training was the Adam method (Kingma and Ba, 2015) combined with the "categorical cross-entropy" loss function. The initial learning rate was set to 0.0002 and the obtained results are presented in the next tables.
Table 4 and Table 4 reports prediction metrics whereas visual results are shown in Figure 8.

Dataset	*OA %*	*mIoU %*	*wIoU %*
Palazzo Severo	97.14	93.77	94.47
Palazzo Mazzetti	95.68	91.68	91.76
Palazzo at n. 48	97	69.2	94.26

Table 4: Quantitative evaluation for materials segmentation.

Dataset	*OA %*	*mIoU %*	*wIoU %*
Palazzo Severo	97.56	91.62	95.28
Palazzo Mazzetti	96.76	79.63	94.02
Palazzo at n. 48	96.74	91.18	93.74

Table 5: Quantitative evaluation for construction technique segmentation.

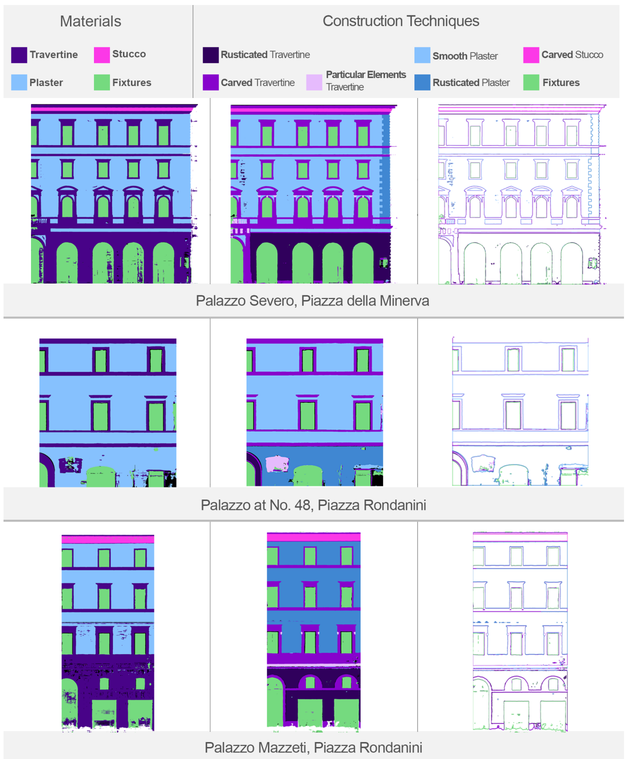

Figure 8: Predicted materials (left) and construction techniques (centre) using the proposed 2D semantic segmentation approach and generated vector drawings (right).

In the final phase, the masks' predictions of the three facades were used to create contour maps of the classes and, from those,

the classes' shapefiles were generated (Figure 8-right).
The metrics and the visual inspection of the inferences reveal a high level of accuracy in the predictions of both materials and construction techniques, exceeding what achieved with the 3D pipeline. The network manages to predict both unbalanced classes and classes with fairly similar geometric and radiometric characteristics with a high degree of success.

4. CONCLUSIONS AND FUTURE WORKS

The investigation aimed to take advantage of the information potential stored in existing technical drawings of buildings' facades. Starting from these architectural drawings, facades' point clouds and their orthoimages, two working pipelines were proposed to produce ready-to-be-used products in the field of architecture and heritage preservation.
The first approach integrates the drawings into a 3D semantic segmentation pipeline aimed at automatically producing labelled point clouds of unseen buildings' facades. The main advantage of this approach consists of automating the generation of training data by transferring the drawings' information to the point clouds, an otherwise time-consuming procedure to perform manually. Despite the lack of relevant geometrical 3D properties in the building facades, the first approach achieved good results in performing the semantic enrichment, automatically generating annotated point clouds in the needed classes. Nevertheless, the use of a DSM as a feature didn't help to improve the accuracy of the predictions.
The second pipeline starts from drawings and facades' orthoimages and create 2D masks exploitable as annotated data. We can then use this data to train a neural network to predict other palaces' facade orthoimages and eventually convert the prediction into technical drawings with the labelled classes. Predictions achieved even higher levels of accuracy, allowing the successful generation of shapefiles for each predicted class. Despite the presence of irregularities and inhomogeneity in the resulting shapefiles, the results seem promising and further development of the procedure could help to improve the outcomes.
The use of technical drawings has proved to be a reliable and fast way to create annotated data from already existing information, significantly reducing the time of data preparation for the training phase. From a wider perspective, this could lead to the creation of a much larger training set by taking advantage of a large corpus of already-made technical drawings.
The vectorization results are promising also on a more conceptual level. We might consider that the technical architectural drawing brings together two main scopes. On one hand, manual drawing has the role of reporting in 2D vectors the metrical and dimensional characters of an object referring purely on captured data. On the other hand, drawing is an interpretation tool of a certain architectural language. This process of interpretation aims at disassembling the building into elements and converting each of them into 2D shapes. These shapes are codified from a geometrical point of view, and they are built referring to geometric constructions. The reported work is located in between these two scopes to evaluate at what extent AI processes can produce a technical architectural drawing able to report not only the metrical and dimensional aspects but also the ones connected to a specialistic interpretation of a building.
First results demonstrate that a certain level of automation in generating a 2D vector drawing with metric accuracy can be reached. Starting from this, the pipeline could be further improved to test the possibility of converting source data into codified geometric shapes. Starting from this, we could be one step closer to a real automatic drawing.

Despite the overall successful performance in taking the most out of the inputted drawings in both pipelines, some issues remain to be addressed:

- Improve the mask creation procedure to be suitably used to create annotated data from detailed, instead of simplified, technical drawings: this could lead to a reduction of the total amount of manual work and overall speed up the procedure.
- Both tested pipelines were trained on a very small amount of training data, reducing significantly the potentialities of the deep learning techniques: upscaling the experiments using a significantly larger amount of data is required to improve predictions' results and generalization.
- Improve the quality of the resulting shapefiles in the 2D pipeline to achieve outcomes closer to the standard used for technical drawings, therefore reducing the need for manual operations to correct possible inaccuracies.

ACKNOWLEDGEMENTS

Laser scanner point clouds and CAD preliminary drawings are the result of student's elaborations at the courses of "Scienza della rappresentazione III" and "Rilievo dell'architettura" conducted by Prof. Alfonso Ippolito in the years 2020/2021 at the Faculty of Architecture, University of Rome La Sapienza, Italy.

REFERENCES

Adam, H., Qiao, S., Yuille, A. L., Collins, M. D., Yuan, L., Yu, Q., Wang, H., Zhu, Y., Weber, M., Cremers, D., Xie, J., Schroff, F., Kim, D., Chen, L.-C. and Leal-Taixe, L., 2021. Deeplab2: A tensorflow library for deep labeling. *arXiv preprint arXiv:2106.09748.*

Agisoft (2023) Agisoft Metashape. https://www.agisoft.com

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A. L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), pp. 834-848

Cloud Compare (Version 2.13.alpha) [GPL software]. 2023. Retrieved from: http://www.cloudcompare.org

Grilli, E., & Remondino, F., 2020. Machine learning generalisation across different 3D architectural heritage. *ISPRS International Journal of Geo-Information*, *9*(6).

Grilli, E., Özdemir, E., & Remondino, F., 2019. Application of machine and deep learning strategies for the classification of heritage point clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-4/W18*, 447–454.

Grilli, E., & Remondino, F., 2019. Classification of 3D digital heritage. *Remote Sensing*, *11*(7).

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980.*

He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. *Proc. CVPR*, 770-778.

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.

Korc, F. & Förstner, W., 2009. eTRIMS Image database for interpreting images of manmade scenes. *Dept. of Photogrammetry, University of Bonn, Tech. Rep*. TR-IGG-P-200901.

Matrone, F., Grilli, E., Martini, M., Paolanti, M., Pierdicca, R., Remondino, F., 2020. Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation. *ISPRS International Journal of Geo-Information*, 9, 535.

Pallottino, E., 1990. Colori di Roma. *Ricerche Di Storia Dell'arte*, 41–42, 129–149.

Pallottino, E., 1992. Il vero e il falso XVI secolo nei rivestimenti dei palazzi romani. Esempi di valutazione e nuove interpretazioni tra Sette e Ottocento. In *Quaderni dell'Istituto di Storia dell'Architettura*. Vol. 2, 799–812.

Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E. S., Frontoni, E. and Lingua, A. M., 2020. Point cloud semantic segmentation using a deep learning framework for cultural heritage. *Remote Sensing*, 12(6), pp. 1005.

Poux, F., Neuville, R., Van Wersch, L., Nys, G.-A., Billen, R., 2017. 3D Point Clouds in Archaeology: Advances in Acquisition, Processing and Knowledge Integration Applied to Quasi-Planar Objects. *Geosciences*, *7*, 96

Teboul, O., Kokkinos, I., Simon, L., Koutsourakis, P., & Paragios, N., 2013. Parsing facades with shape grammars and reinforcement learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(7), pp. 1744–1756.

Tyleček, R. and Šára, R., 2013. Spatial pattern templates for recognition of objects with regular structure. Proc. *German Conference on Pattern Recognition*, 364374. Springer.

Valero, E., Bosché, F., & Forster, A., 2018. Automatic segmentation of 3D point clouds of rubble masonry walls, and its application to building surveying, repair and maintenance. *Automation in Construction*, *96*, 29–39.

Yang, S., Hou, M., Li, S., 2023. Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review. *Remote Sensing*, 15(3):548.

Zhao, H., Jiang, L., Jia, J., Torr, P. H., & Koltun, V., 2021. Point transformer. *Proc. ICCV,* 16259-16268.