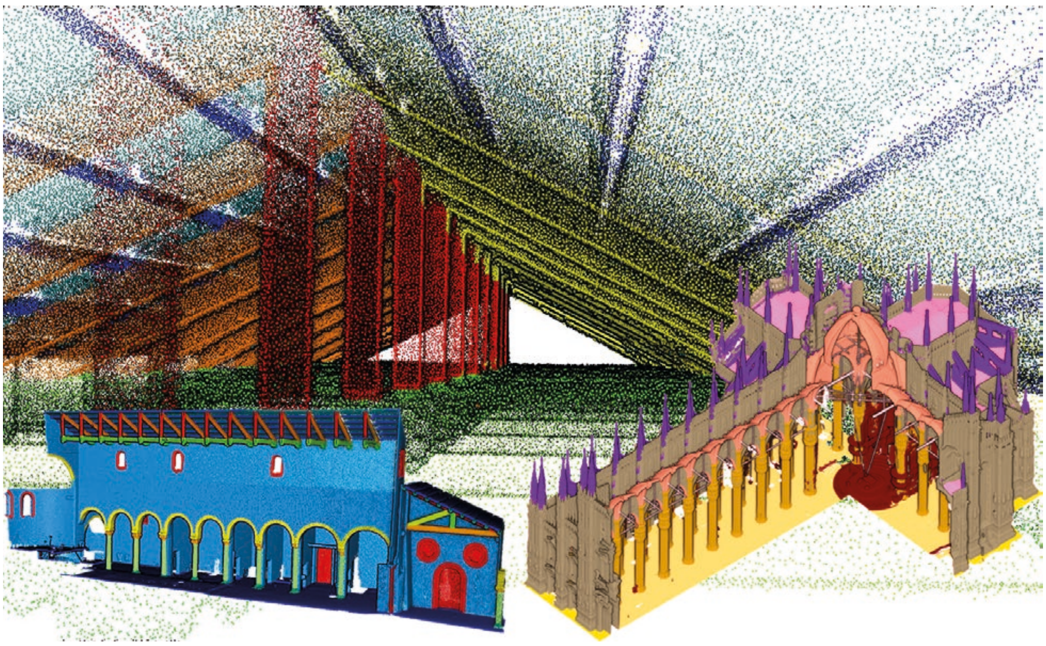
Machine Learning for Cultural Heritage Classification

Michele Russo
Eleonora Grilli
Fabio Remondino
Simone Teruggi
Francesco Fassi

Abstract

Cultural Heritage (CH) assets may be defined as integrated spatial systems composed of interconnected shapes. The classification and organization of geometries within a hierarchical system are functional to their correct interpretation, which is often performed using 3D point clouds. The recurring shapes recognition becomes a crucial activity, nowadays accelerated by Machine Learning (ML) procedures able to associate semantic meaning to geometric data. An interdisciplinary research team [1] has developed a ML supervised approach, tested on the Milan Cathedral and Pomposa Abbey datasets, which presents an innovative multi–level and multi–resolution classification (MLMR) process. The methodology improves the learning activity and optimizes the 3D classification by a hierarchical concept.

Keywords
machine learning, cultural heritage, multi–resolution, hierarchical 3D classification, level of detail.



doi.org/10.3280/oa-686.33

Introduction

Cultural Heritage (CH) assets are complex artifacts whose knowledge passes through analyzing an integrated system of forms interconnected by dependence or proximity relationships. The recognition and classification of 3D data become essential to (re)assign a hierarchical and functional meaning to acquired point clouds. The manual classification activity, which is very time–consuming, can be nowadays replaced by an automatic one based on Artificial Intelligence (AI) approaches, such as Machine Learning (ML) or Deep Learning (DL) methods. These AI approaches have many bottlenecks in the CH field, mainly due to the complexity and variability of the shapes, the reliability of the interpreted data, the scalability of the process and, often, the absence of annotated data. In this paper, a supervised ML method applied to CH is introduced and evaluated. It is based on a Multi–Level Multi–Resolution (MLMR) approach, which considers the various geometric details present in the point cloud. Two complex 3D datasets related to Milan Cathedral and Pomposa Abbey are processed to test the developed methodology and demonstrate its flexibility and efficiency with different scenarios.

State of the Art

Several investigations performed to classify (or semantically segment) 3D point clouds in the architectural heritage field using automatic ML and DL methods. Grilli et al. [2018, pp. 1-8] presented a supervised ML approach to transfer classification data from 2D textures to 3D models, whereas Grilli et al. [2020] used a Random Forest (RF) classifier with geometric features to derive architectural classes from point clouds. In the DL domain, Pierdicca et al. [2020] trained the ArCH dataset (http://archdataset.polito.it/) with a Dynamic Graph Convolutional Neural Network (DGCNN) using meaningful features (colour, normals, and HSV), providing promising results. A comparison of ML and DL techniques for the classification of architectural point clouds [Matrone et al. 2020,] shows that similar accuracy results can be achieved. However, ML requires much less time and does not need large 3D datasets in the training phase. For this reason, we hereafter present a supervised ML approach adapted to the different geometric levels of detail and architectural classes.

The Case Studies and Classification Purposes

Two datasets, with different dimensional and morphological characteristics but presenting similar architectural elements, were selected for validating the methodology. The first case study is the Milan Cathedral (fig. 1) which was digitally recorded in the last decade with

Fig. 1. External and internal photos of the Milan Cathedral and Pomposa Abbey, with details of the monumental capitals of the Cathedral and the wooden roofs of the Abbey (authors' images).

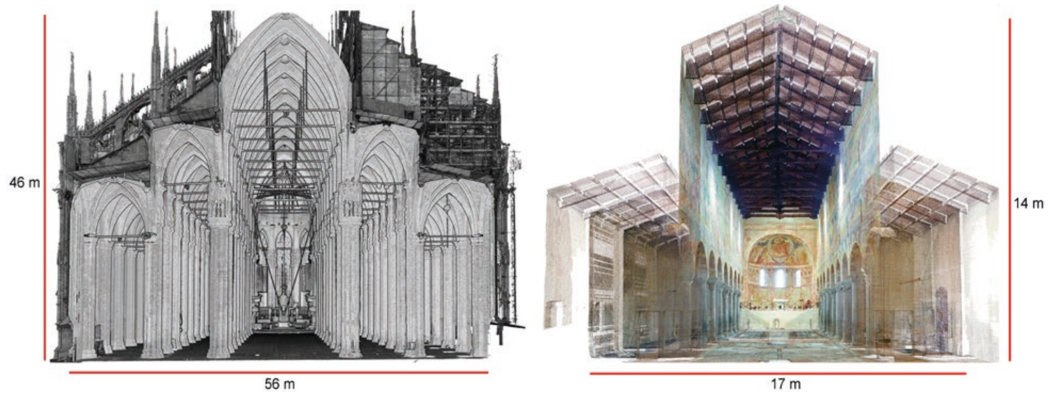


Fig. 2. A view on the point clouds of the two datasets: The Milan Cathedral (left) and the Pomposa Abbey (right).

several integrated acquisition campaigns to generate parametric models [Fassi et al. 2011, pp. 462-487], and define a complete 3D point cloud (fig. 2) at a uniform average resolution of 5 mm [Achille et al. 2020, pp. 331-341]. The classified point cloud may facilitate the 3D data exploration, allowing the integration between archival sources and surveyed data on a web–based BIM–type platform, which can be consulted in situ or remotely. This data organization can also allow multi–scale planning and implementation of conservation and management projects and the quick extraction of 2D representations already classified. The second case study is the Pomposa Abbey (fig. 1) surveyed in 2014 to generate a complete 3D dataset (fig. 2) at a uniform average resolution of 2 cm [Russo et al. 2014, pp. 305-312]. In this scenario, the 3D classification activity can foster access to the system's knowledge, supporting its graphic restitution and the monitoring activities at different scales. Besides, it can facilitate the "quantification" of the building, collecting helpful information for planning a conservation intervention and evaluating the transformations over time.

The Methodological Workflow

The high level of complexity of the case studies highlights two different bottlenecks: on the one hand, the processing of massive datasets cannot be simplified unless losing the level of detail useful in the element recognition. On the other hand, the high number of semantic classes raises the management complexity and reduces their identification accuracy [Teruggi et al. 2020]. An iterative methodology [Grilli et al. 2020] has been developed to overcome these bottlenecks, classifying 3D data in multiple steps according to their information levels (fig. 3). The proposed hierarchical structure is referred to the data density, the morphological and compositional complexity, and the classification purpose. At each level of detail (LOD), the workflow foresees two working steps:
1) The selection of 'covariance features' [Blomey et al. 2014] extracted within specific spherical radii, for the automatic recognition of local geometric characteristics of 3D datum.
2) A small manual annotation to train a Random Forest algorithm [Breiman 2001, pp. 5-32], associating each portion identified by the features to architectural meanings.
The training dataset's selection evaluates the presence of the elements to be classifiied.

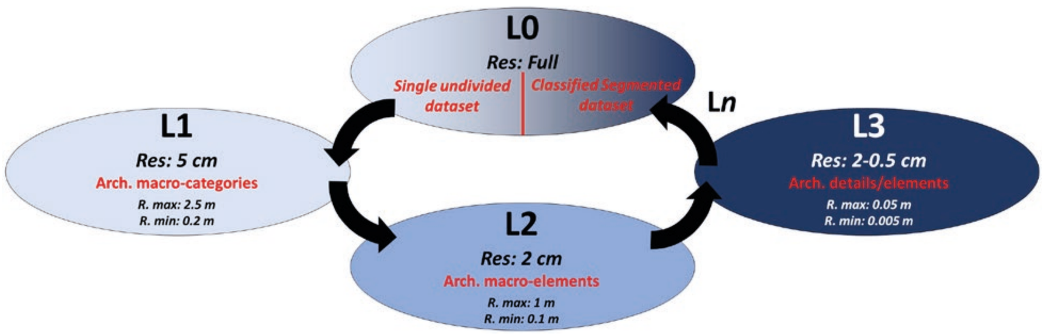


Fig. 3. Schema of the MLMR iterative process.

Experimentation and Results

The classification process refers to the following three–level of details (fig. 4):
– In the first level (L1), a point cloud subsampled at 5 cm, with min/max radius of the features between 20 cm and 2.5 m, was processed, subdividing the churches into architectural macro–categories;
– In the second level (L2), after transferring the L1 classification to the 2 cm resolution point cloud, features extracted with radii between 10 cm and 1 m were used to split the architectural elements into macro–elements;
– In the third level (L3), receiving the L2 subdivision, features with radii of 0.5 and 5 cm were used on the 3D point cloud with a 5 cm density for the Cathedral and 2 cm for the Abbey. This allowed identifying the single architectural monolithic and technologically coherent components.
Both the processing time and the metrics commonly used in ML to define reliability of the results ("Precision," "Recall," and "F1 score" [Goutte et al. 2005, pp. 345-359]), were analyzed to evaluate the classification performance (tab. 1).

Table 1. Timing and metric summary for the two datasets according to the three classification levels. (*) 18 Core Processor; (**) 12 Core Processor.

	Milan Cathedral*			Pomposa Abbey**		
	L1 (5 cm)	L2 (2 cm)	L3 (0.5 cm)	L1 (5 cm)	L2 (2 cm)	L3 (2 cm)
Features computation (min.)	1500			30		
Annotation (min.)	500			60		
Training (sec.)	363	17	142	5	1	4
Classification (sec.)	43	12	174	2.7	1	29
Precision (%)	94.7	99	92	95.3	98	95.8
Recall (%)	95	98	88.5	95.1	97.7	95.7
F1 Score (%)	93.8	99.3	91.8	95.1	97.8	94.6

The achieved results highlight the importance of using point clouds with a level of detail (geometric resolution) and density suitable to support subsampling or backward interpolation processes consistent with identifying architectural elements. Moreover, if the features radii affect only the time in shapes research and the complexity of the architectural connections affects just the classification process, the geometrical density and the processor capacities affect the whole timing workflow (tab. 1). The reported quality metrics show the possibility of obtaining excellent results quickly, identifying even very complex geometric structures.

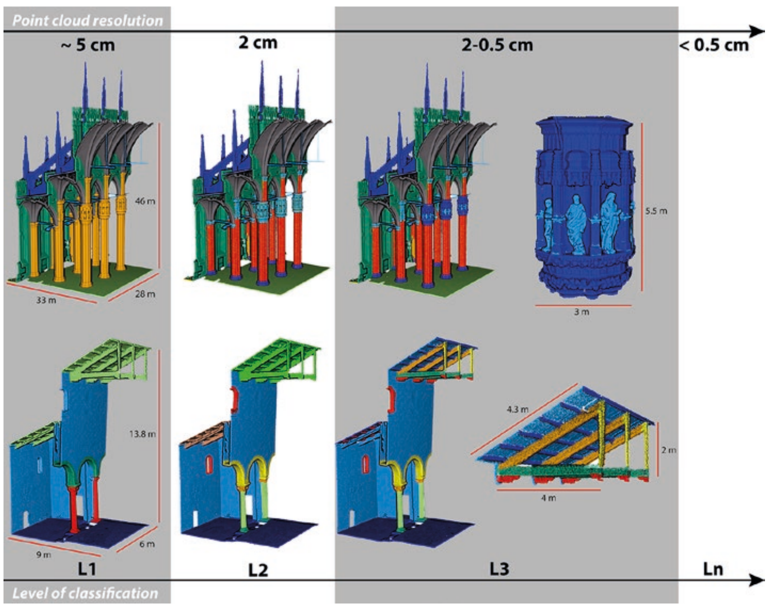


Fig. 4. The classification process of the Cathedral (top) training portion and the Basilica (bottom) according to the resolutions and levels of detail sought by the features.

Conclusions and Future Research

In this paper, a new iterative strategy for supervised automatic ML classification of 3D point clouds of complex Cultural Heritage is presented. Few annotated 3D data were necessary and very detailed semantic segmentation results could be achieved. The cognitive contribution in the supervision phase is crucial in the correct definition of classes and the choice of training and validation sets. These steps are also critical to adapt the general process to the specific case study and different purposes.
In the future, the relationship between classification levels, cloud resolution, and feature search radii will be more investigated, defining a general multi–scenario approach. Besides, the introduction of photogrammetry into the process as a tool to acquire an additional level of detail may be of particular interest. Scan–to–BIM and reality–based modelling from classified data may be specific topics to analyses, supporting the point cloud seg–mentation purposes. A final goal concerns the creation of a classification framework that is more user–friendly for non–experts in the field, broadening its application to different disciplinary areas.

Notes

[1] The presented research is the result of the joint work of five authors. M.R. took care of the Introduction and Conclusions, E.G. prepared the State of the art, the methodological workflow and run the case studies, S.T. supported the methodological workflow and experiments, F.F. and F.R. supervised the work and reviewed the paper. All authors shared the analysis of experiments and results.

References

Achille Cristiana, Fassi Francesco, Mandelli Alessandro et al. (2020). From a Traditional to a Digital Site: 2008-2019. The History of Milan Cathedral Surveys. In *Digital Transformation of the Design, Construction and Management Processes of the Built Environment.* Cham: Springer, pp. 331-341.

Blomley Rosmarie, Weinmann Martin, Leitloff Jens et al. (2014). Shape distribution features for point cloud analysis–a geometric histogram approach on multiple scales. In *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2 (3), pp. 9-16.

Breiman Leo (2001). Random forests. In *Machine learning*, 45 (1), pp. 5-32.

Fassi Francesco, Achille Cristiana, Fregonese Luigi (2011). Surveying and modelling the main spire of Milan Cathedral using multiple data sources. In *The Photogrammetric Record*, 26 (136), pp. 462-487.

Goutte Cyril, Gaussier Eric (2005). March. A probabilistic interpretation of precision recall and F–score, with implication for evaluation. In *European conference on information retrieval.* Berlin: Springer, pp. 345-359.

Grilli Eleonora, Dininno Domenica, Marsicano Lucia et al. (2018). Supervised segmentation of 3D cultural heritage. In *3rd Digital Heritage International Congress (DigitalHERITAGE)*. USA: IEEE, pp. 1-8.

Grilli Eleonora, Remondino Fabio (2020). Machine Learning Generalisation across Different 3D Architectural Heritage. In *ISPRS International Journal of Geo–Information*, 9 (6), 379, pp. 1-19.

Matrone Francesca, Grilli Eleonora, Martini Massimo et al. (2020). Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation. In *ISPRS International Journal of Geo–Information*, 9 (9), 535, pp. 1-22.

Pierdicca Roberto, Paolanti Marina, Matrone Fracesca et al. (2020). Point Cloud Semantic Segmentation Using a Deep Learning Framework for Cultural Heritage. In *Remote Sensing*, 12 (6), 1005, pp. 1-23.

Russo Michele, Manferdini Anna Maria (2014). Integration of image and range–based techniques for surveying complex architectures. In *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* 2 (5), pp. 305-312.

Teruggi Simone, Grilli Eleonora, Russo Michele (2020). A Hierarchical Machine Learning Approach for Multi–Level and Multi–Resolution 3D Point Cloud Classification. In *Remote Sensing*, 12 (16), 2598, pp. 1-27.

Authors
Michele Russo, Dept of History, Representation and Restoration of Architecture, Sapienza University of Rome, m.russo@uniroma1.it
Eleonora Grilli, 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), grilli@fbk.eu
Fabio Remondino, 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), remondino@fbk.eu
Simone Teruggi, 3DSurveyGroup – ABCLab – Politecnico di Milano, simone.teruggi@polimi.it
Francesco Fassi, 3DSurveyGroup – ABCLab – Politecnico di Milano, francesco.fassi@polimi.it

Isbn 9788835125280