POLITECNICO DI MILANO 1863
School of Architecture Urban Planning Construction Engineering
Master course in Architectural Design and History

Candidate: Alessandro Zivelonghi 967134
Supervisor: Prof. Francesco Fassi
Co-supervisor: Kai Zhang

Academic year: 2022-2023

POINT CLOUD CLASSIFICATION

The Case of Milan Cathedral

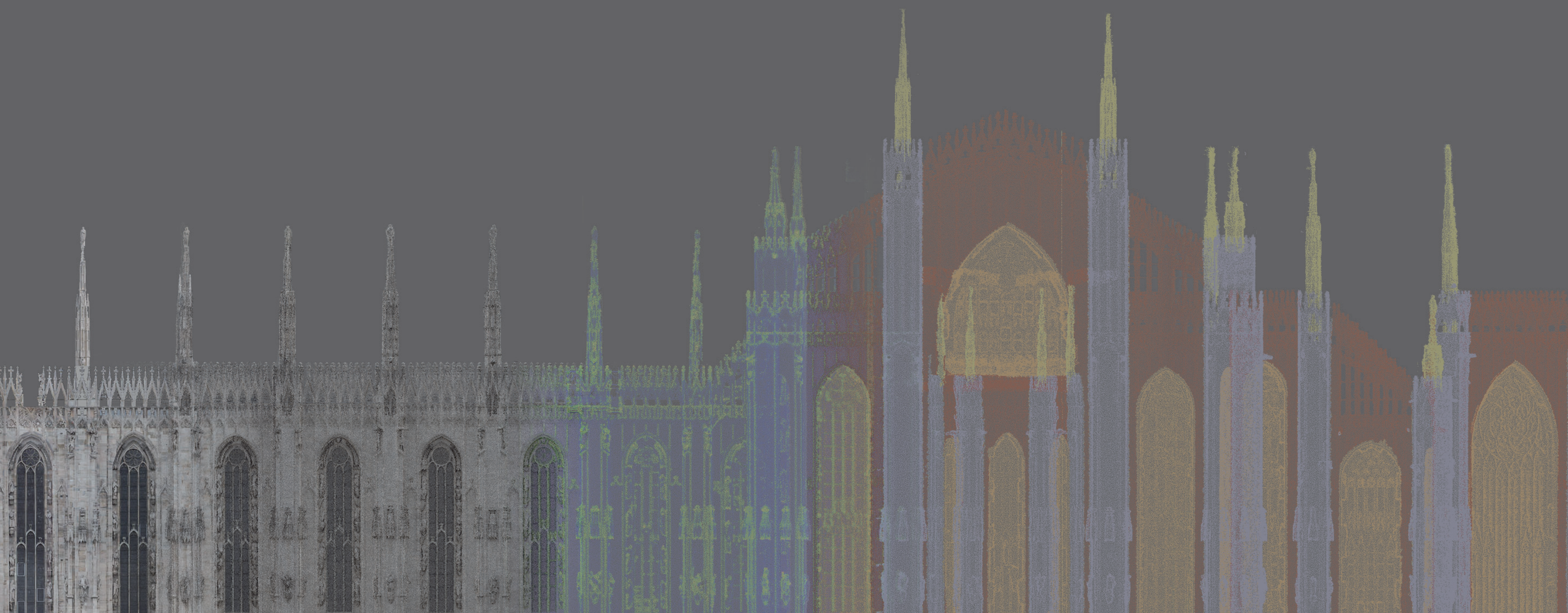

TABLE OF CONTENT

ABSTRACT

ENGLISH

This work is part of a ten-year collaboration between the Veneranda Fabbrica del Duomo di Milano and the Politecnico di Milano, 3D SurveyGroup, DABC. The basis of this work dates back to several years ago, specifically when the 3D SurveyGroup of the Politecnico started to use the TLS and the photogrammetry, besides other techniques, as the two primary sources from which point clouds are obtained. Therefore, this work represents an additional contribution to the ongoing project. (et al. Achille, C., Fassi, F., Mandelli, A., Perfetti, L., Rechichi, F., Teruggi, 2020)

The importance of surveying processes in the field of heritage conservation and restoration is well established and widely used worldwide. It is evident that nowadays, surveying techniques are able to solve almost any problematic situation in terms of data acquisition and data elaboration. The major gap still concerns the use of data at a professional level in daily practice. In these terms, an example would be the fact that point clouds are often used as a transitional tool between the surveys and the manual realization of 3D models, rather than exploiting its other potential and unique features.

Consequently, this investigation may be considered as part of the contemporary national and international research effort, so that the raw products of 3D acquisitions (point clouds) can be easily used professionally. In particular, the thesis utilizes the artificial intelligence (AI) system and examines the techniques for the classification of 3D point clouds to be used as real 3D models structured for different purposes of the professional practice in the Cultural heritage's world. (et. al. Grilli, E. and Remondino, F.: 'Classification of 3D Digital 'Heritage', Remote Sens., 11, 847, 2019. https://doi.org/10.3390/rs11070847

In addition to the more obvious preservation benefits, a lesser-known potential of 3D digital data is that it permits a comprehensive interpretation of architecture, beginning with an understanding of the elements' semantics and proceeding through an examination of their structural and compositional roles, to a consideration of the building's forms and its relationship to space, both real and imagined.

The Milan Cathedral constitutes a unique opportunity for this study mainly because it is both a complex and extensive case study that has proven to be one of the most difficult in the field of surveying and because it necessitates the classification of 3D point clouds of all the architectural elements, which are all of remarkable peculiarity and complexity.

ITALIAN

Questo lavoro fa parte di una collaborazione decennale tra la Veneranda Fabbrica del Duomo di Milano e il Politecnico di Milano, 3D SurveyGroup, DABC. Le basi di questo lavoro, infatti, risalgono a molti anni fa, quando il laboratorio 3Dsurvey group del Politecnico ha iniziato a utilizzare il TLS e la fotogrammetria (oltre a varie altre tecniche): le due fonti primarie da cui si ottengono le nuvole di punti su cui si concentra questo lavoro. Pertanto, questo lavoro rappresenterà un piccolo mattone in un progetto in continuo sviluppo. (et al. Achille, C., Fassi, F., Mandelli, A., Perfetti, L., Rechichi, F., Teruggi, 2020)

L'importanza dei processi di rilievo nel campo della conservazione e del restauro del patrimonio è ben consolidata e ampiamente utilizzata in tutto il mondo. È anche chiaro che oggi le tecniche di rilievo possono risolvere quasi tutte le situazioni in termini di acquisizione ed elaborazione dei dati. La grande lacuna rimane oggi l'utilizzo dei dati a livello professionale nella pratica quotidiana, è esemplare in questi termini fatto che le nuvole di punti sono spesso utilizzate come strumento di passaggio tra il rilievo e la realizzazione manuale di modelli 3D.

La ricerca presentata in questo lavoro si inserisce negli sforzi di ricerca nazionali e internazionali affinché i prodotti grezzi delle acquisizioni 3D (nuvole di punti) possano essere facilmente utilizzati a livello professionale. In particolare, la tesi indaga, attraverso l'uso di sistemi di intelligenza artificiale (AI), le tecniche per classificare le nuvole di punti 3D da utilizzare come veri e propri modelli 3D strutturati per diversi scopi durante la pratica professionale nel mondo dei beni culturali. (et. al. Grilli, E. e Remondino, F.: "Classificazione del patrimonio digitale 3D", Remote Sens., 11, 847, 2019. https://doi.org/10.3390/rs11070847

Oltre ai più ovvi benefici per la conservazione, una potenzialità meno nota dei dati digitali 3D è quella di consentire un'interpretazione completa dell'architettura, a partire dalla comprensione della semantica degli elementi, procedendo attraverso il loro ruolo strutturale e compositivo, fino alla considerazione delle forme dell'edificio e del suo rapporto con lo spazio, sia reale che ideale.

Il Duomo di Milano rappresenta un'opportunità unica per questo studio, sia perché è un caso di studio complesso ed esteso che si è rivelato uno dei più difficili nel campo del rilievo, sia perché richiede la classificazione delle nuvole di punti 3D di tutti gli elementi architettonici, tutti di notevole peculiarità e complessità.

PART 0 Introduction

THE CATHEDRAL OF MILAN: HISTORY OF CONTINUOUS PROGRESS

Evolution and progress can be considered a continuous succession of events and elements that constitute "history" over time. This statement finds its highest expression in artistic disciplines, which have always played a dual role throughout human development: that of inspiring social development and of being a mirror of this progress. However, it is worthwhile to identify what "ancient" means: when does something cease to live its time and become an "ancient" entity host to the present? Narrowing the focus of this research to the art of architecture alone, one would think that "ancient" is what embodies knowledge and information that cannot be found in contemporaneity. It is, therefore, natural to ask ourselves when a work becomes "antique" and according to which logic the adjective may be attributed to something.

The Cathedral was for sure a significant symbol for Milan and its citizens, more than for its religious value, for its being a sign of rehabilitation and certification of the centrality of the city, which was losing its importance due to internal political disputes among the Visconti family. During the 1386, the construction of the Duomo triggered a series of social and economic mechanisms without precedent: new models of site management and materials have given the enormous architectural challenge to be accomplished. As in the past, the Dome is again today a forerunner of new models of study and analysis. So, even centuries later, the Milan's Cathedral provides us the opportunity to progress and evolve new approaches and techniques.
The fact that the Duomo is the subject of this work is undoubtedly not coincidental; whether it is 1386 or 2023, the same church, over the years, confirms itself as the symbol of centrality and progress.

"Torniamo all'antico e sarà un progresso."
"Let's go back to the past and it will be a progress."
Giuseppe Verdi.

Giuseppe Verdi's words resonate in these terms: 'Let us return to the ancient: it will be progress.' According to this philosophy, with the ability to investigate, recognize and respect what is "ancient", we have the possibility of projecting it directly into our present and learning the lesson of the past. These are the most appropriate and meaningful words to describe this work and the research behind it: starting with the symbolic significance that the Duomo has had throughout history, and then concluding with the analysis of the individual indivisible architectural elements that compose it. In other words, an investigation that starts from far away and then becomes an intimate knowledge of the Cathedral, using modern technology to detect and get to know the ancient forms and elements. Gothic ornaments and decorations were considered essential to the realization of a Cathedral worthy of being called such; the reality is that it carries way more meanings in the elements of which it is composed. Every single ashlar that is part of an architectural element encloses within itself a history and a meaning that goes beyond the passage of time. Each of these pieces encompass a sense of completeness in their cohesion and in their belonging to something greater, and in their entirety travel through time to communicate the ancient to the present and the future.
This is probably the most intimate and important value that I have learnt to identify and appreciate in historical architecture, and it is thanks to modern surveying technology and the work done in this paper that I have had the opportunity to go beyond the canonical perception of human senses of space and forms in architecture. The use of point clouds, as well as the possibility of creating new forms of representation and relief of historical heritage, has provided me a new system of analysis and perception, which has revealed sides apparently indistinguishable to the human eye alone. The possibility of understanding how individual elements are composed and how they intersect with each other is one of the assets that impressed me most about point clouds.

STATE OF THE ART

Point cloud segmentation of cultural heritage is an important research area in 3D computer vision. It aims to automatically extract meaningful information from point clouds captured by 3D scanning technologies, such as laser scanners, photogrammetry, or structured light scanning. The goal is to segment the point cloud into meaningful regions, such as walls, windows, doors, and ornaments, to ensure the creation of accurate and detailed 3D models of cultural heritage sites.
The increasing demand and size of heritage survey data have led to developing new systems to manage this data. Over the last few years, studies have shown that artificial intelligence should be used to cope with this condition and quickly perform tasks that would take humans much longer to do.

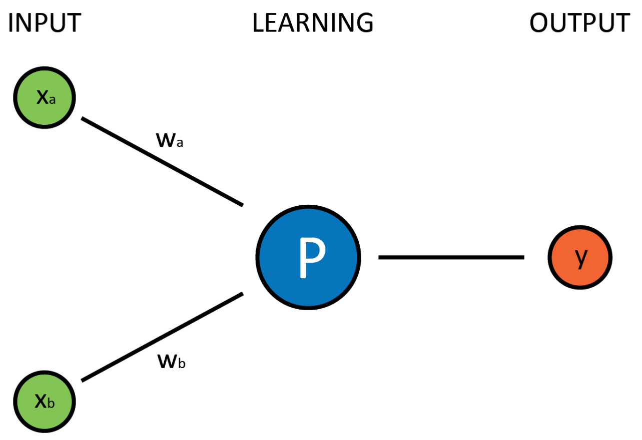

fig.1

fig.1 Abstract scheme of how perceptron works.

Artificial intelligence (AI) may be defined as a discipline that studies whether and how it is possible to develop intelligent computer systems capable of simulating the capacity and behavior of human thought. There are several detailed definitions of it but the one that is here useful to outline is the difference between Narrow AI and General AI. The former is a type of artificial intelligence that focuses on solving specific tasks, and only specializes in the field of research in which it is employed. The second AI, on the other hand, is capable of managing any issue or situation. Considering the complexity of the data and answers that the General AI has to process, it is seen as the closest approximation to the human brain, making its development extremely complex and slow.
Narrow AI has been the subject of numerous studies and implementations in the field of automatic classification, mainly due to the possibility of focusing on a single area of research. In recent years, the generic model that has been most successful is the family of Neural Networks (NN), whose main advantage is its modularity and the consequent possibility of being decomposed and recomposed according to the tasks to be performed. The most basic NN model (fig. below) is the perceptron, which consists of two input nodes, an intermediate processing node (kernel) and an output node that constitutes the final result of the processing. The data received from the input nodes is influenced by the weight-value (called Wa and Wb) that the network assigns to the feature of that node.

Most studies on automatic classification have involved the so-called Machine Learning, an AI system that learns through information provided by the operator. Its primary use concerned the 2D image because of the easy extraction of information, which are necessary for both the learning and the subsequent classification through Convolutional Neutral Networks (CNN). CNNs are neural networks consisting of layers of nodes through which information is processed and passed on and are a subset of machine learning.

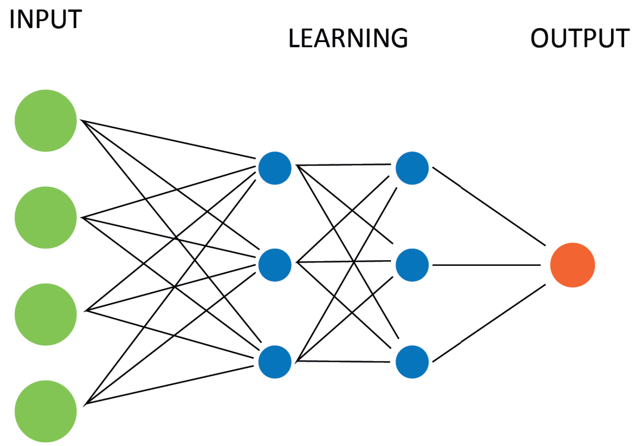

fig.2

fig. 2 Abstract scheme of how CNN works

The main development of this system in the field of 2D is a consequence of the fact that an image can be read as a grid of pixels in which it has only three characteristics (R, G, B) and in which the grid composition always has the same characteristics of distance and position between pixels. Such systems are called 'structured' and, as such, it is easier to establish a fixed analysis system for these features. 3D point clouds, unlike 2D images, do not have a regular structure, as each point in the cloud is surveyed by the laser independently of all the others; therefore, there is no automatic relationship regarding the positions of the points. This necessitates the integration of feature assignment systems that can provide additional useful information for feature detection. (Cao, Y. and Scaioni, M.: A Pre-Training Method For 3d Building Point Cloud Semantic Segmentation, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-2-2022, 219–226, 2020.)
As a consequence of the recent overwhelming success of convolutional neural networks (CNNs) for image analysis Y. Wang et. al. (2019) suggests the value of implementing the abilities of CNNs to the world of point clouds. Wang proposes a new neural network module called 'EdgeConv' suitable for high-level CNN-based tasks for point clouds, which include classification and segmentation.
Compared to previous modules that operate in extrinsic space or treat each point independently, EdgeConv investigates by incorporating the local neighborhood information of each point. Furthermore, EdgeConv is differentiable and can be inserted into existing architectures.

A study by Thomas et. al. (2019) (28) showed that it is possible to use kernel points as convolution filters and to operate on the points without transformations. Convolution weights are learned from kernel points and their neighbors in Euclidean space. On the other hand, GCNs

(Global Convolutional Networks) can naturally extract geometric information from their surroundings by exploiting Graph Neural Networks (GNNs) as a representative method, which is based on the use of CNNs to extract data from the graphs of each point.

In the same paper by Y. Wang Li et al. (2019) (30) mentioned before, he proposes a method that combines GCNs and attention mechanisms to segment point clouds in CAD models. The method is based on a graph representation of the point cloud, where each point is a node in the graph, and the edges between the nodes represent the spatial relationships between the points.
Recent studies conducted by Teruggi et al. (29) have shown that basing the first approach on the automatic recognition and categorization of architectural elements using machine learning can be a winning choice. This makes it possible to obtain cloud models based on the semantic meaning of the elements, concentrating any manual modelling operations on already categorized data and simplifying the creation of models.
According to Teruggi et al. (29), Machine Learning (ML) and Deep Learning (DL) methods, as opposed to manually performed operations, are objective, replicable and repeatable to other data not belonging to the same case study. Standard supervised ML techniques involve algorithms taking as input some manually annotated parts of the point cloud, along with the so-called "geometric features", and/or radiometric attributes selected by the operator to facilitate learning and distinction of the classes to be segmented. On the other hand, DL strategies include the automatic generation of features, which learn through large amounts of annotated input data.
The direct classification of point clouds is a field that has not yet been fully developed. Consequently, this kind of work aims to implement the database of knowledge and experience related to this classification method. That is to say that the objective of this thesis is to contribute to the implementation of the purpose mentioned and to provide new insights into the cataloguing and maintenance of cultural heritage.
In summary, recent research in point cloud segmentation of cultural heritage has demonstrated the potential of deep learning techniques to achieve accurate and efficient segmentation of complex and noisy scenes. These methods have shown promising results in achieving state-of-the-art performance on benchmark datasets. They are expected to be increasingly important in creating accurate and detailed 3D models of cultural heritage sites.

Chapter 1: The Digital Survey

POITN CLOUDS

In recent years, modern technologies have spread the use of point clouds to survey historical heritage for analysis, documentation, and conservation purposes. The point cloud consists of a set of points surveyed in the area under investigation and inserted into a three-dimensional reference system.
Point clouds are the first data generated by surveying operations that exploit modern techniques to acquire and digitize buildings forms. Thanks to their ability to detect a significant amount of points in a few minutes, the point clouds provide information on the spatial characteristics of the surveyed elements. The main surveying techniques used in this field are Terrestrial Laser Scanning (TLS) and photogrammetry. The first one uses a LiDAR system to detect the distance of objects points by calculating the time it takes a laser pulse to hit them and return back to the source. Based on the return time of the laser signal, it can determine the distance of the point from the origin and, based on the measurement of the zenith and azimuth angles according to which the instrument is oriented, it can define the position of each point within a Cartesian reference system by assigning x, y, and z coordinates. In addition to these position-related characteristics, the TLS is able to provide information about the intensity and the reflectance of different materials by comparing the strength of the incoming signal to the launched one. The main advantage of using this instrument is the ability to survey millions of points per second in a 360° field of acquisition around the instrument position.
Photogrammetry, on the other hand, is a surveying technique that allows the metric modelling based on the stereometric information derived from several frames that capture the same surface portions. From a practical point of view, the principal advantage of this type of survey regards the use of a digital camera, a much cheaper and more flexible tool than TLS. While the TLS needs to remain stationary in a fixed position for several minutes, taking a single frame with the camera requires a few tenths of a second (under optimal environmental conditions). This allows greater flexibility in the photogrammetric survey as, for example, allowing the acquisition from moving cranes or UAVs.
Latterly, the speed, the completeness, and the ease of use of these surveying methods have extended their use on a large scale, both for the simplest and most primitive geometries and for architectures containing more complex decorative elements.

In general, 3D point clouds are data obtained by surveying each point as a single element in space and then placing it in a three-dimensional reference system with all the others. In these terms, point clouds are often used as a transition tool between surveying and the manual realization of 3D models, as they provide very precise metric information for individual points.

With the development of surveying techniques and tools, the amount and dimension of surveying sites have increased exponentially, bringing to attention the opportunity to enrich this data with a hierarchical or semantic structure of points.

Chapter 2: Classification

CLASSIFICATION CRITERIA

In point cloud processing, classification means grouping portions of the cloud in which points have common properties into different categories. According to E. Grilli et al., (9) (Classification of 3D Digital Heritage, E. Grilli, F. Remondino, 2019), a correct approach to process historical heritage data is based on three main concepts: segmentation, structuring the hierarchical relationships of the elements and implementation of semantics.
The principle driving this research is that the machine can do the most straightforward things than man, even in a significantly faster and more objective way. All it needs is to learn to do the operations it is asked to do.
This paper will test and discuss the classification method based on Machine Learning which, as stated before, solves problems by using algorithms and statistical models to extract knowledge from data. It is a subset of Artificial Intelligence that exploits algorithms for the learning and subsequent resolution of specific tasks on a model to be classified. In our specific case, it will be provided with information from which to learn what features the points belonging to the different categories display. In this way it will be able to examine the data points to be classified and, based on the characteristics it finds in each, assign it to the category it considers most suitable.

In general, it is possible to distinguish three types of Machine Learning: supervised Learning, unsupervised Learning, and reinforcement Learning. (13) (V. Gupta, V. K. Mishra, P. Singhal, and A. Kumar, "An Overview of Supervised Machine Learning Algorithm," 2022 11th International Conference on System Modelling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 87-92, doi: 10.1109/SMART55829.2022.10047618.)
Supervised learning works on labeled and supervised datasets. It requires to provide manually the information necessary for the algorithm to recognize the elements with their semantic meaning to classify them. So, this working methodology involves extrapolating significant portions from the initial data to manually classify the elements concerned. Based on this study model, the algorithm will train to segment the initial data.
Usually, for classification purposes, machine learning usually exploits the Decision Trees (DT) technique, an instrument in which each piece of data to be classified is examined to draw conclusions on a series of observations. It is a tree model of decisions and possible consequences on the iterative basis of the information available for class identification. The depth of the tree is directly proportional to the amount of information the code has available. This does not always correspond to a better accuracy of the final result for two main reasons: the first is that, given the simplicity of the tree, the variation of a single parameter could significantly alter the final result; the second is that, in more complex situations where the number of labels and features increases, the performance will reverse its trend, taking longer and providing less accurate results.
The Decision Tree algorithm is easy to understand and interpret, but this also makes it a limited method mainly because a single tree is insufficient to produce effective results in complex situations. This is where the Random Forest (RF) algorithm comes in, the one most commonly used by data scientists. RF is an extension of the bootstrap aggregation of Decision Trees and can be used for classification and regression problems. It is a collective learning algorithm, which builds a multitude of Decision Trees instead of following a single branch considered most important as in the Decision Tree. For each joint in each individual tree, the decision is not made on the basis of the feature with the best information match, but based on the best of a randomly chosen subset of n. features. The RF model performs well in the point-based classification approach, even for unbalanced classification datasets.

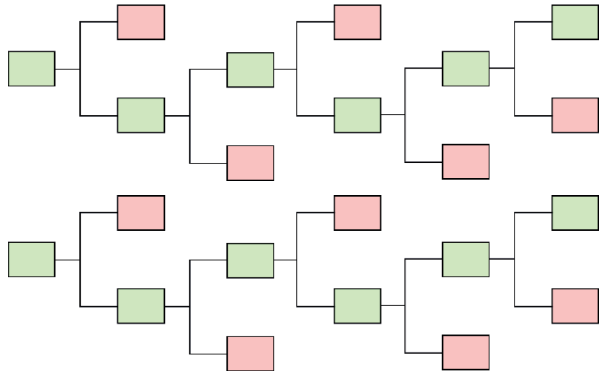
fig.3

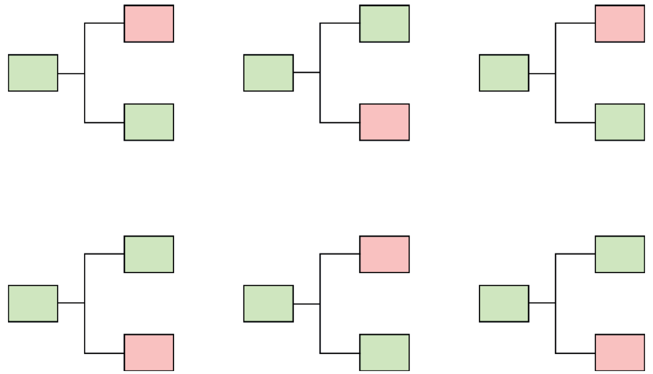
fig.4

fig.3 Schematic functioning of two the Decision Tree (in green positive findings to continue, in red negative findings that interrupt the branch)

fig.4 Schematic functioning of the Random Forest (in green are positive hits to continue from, in red are negative hits that interrupt the branch)

As anticipated before, the classification method adopted was the Random Forest (RF). This is a supervised classification algorithm developed by Leo Breiman (2001) that exploits a set of classification trees, obtains a prediction from each tree, and lastly selects the best solution by voting. Two parameters are required to activate the forest of trees: the number of Decision Trees to be generated (Ntree) and the number of variables to be selected and evaluated for the best subdivision during tree growth (Mtry).

This algorithm was selected for this paper mainly because of some of its characteristics, which were explored by E. Grilli et. al. (12) (Grilli, E., Farella, E. M., Torresani, A., and Remondino, F.: "Geometric Features Analysis for The Classification of Cultural Heritage Point Clouds", Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W15, 541-548, 2019):

- RF is considered a relatively accurate and robust method due to the number of Decision Trees involved in the process.

- RF offers a useful feature selection indicator. In particular, it shows the relative importance or contribu-tion of each feature in the prediction: it automatically calculates the relevance score of each feature in the training phase, then scales the relevance so that the sum of all scores is equal to 1.

MULTI-LEVEL AND MULTI-RESOLUTION APPROACH (MLMR)

In accordance with Teruggi et al. (29) (Teruggi, S.; Grilli, E.; Russo, M.; Fassi, F.; Remondino, F.: "A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification."), a multi-level and multi-resolution approach was applied, both considering the size of the data to be processed and the presence of numerous elements at different scales. This approach involves dividing the work into three classification levels, modifying the resolution of the cloud from time to time to adapt it to the scale of elements to be recognized at each stage.
One of the primary keys to achieving a proper classification is to design the development of the work in order to balance the time as much as possible, and to find the correct proportion between manual and automatic operations. This is because when dealing with clouds of millions of points, it is clear that it will never be possible to achieve 100 per cent accuracy, and certainly, trying to get closer and closer to this value would be a waste of time. So, this entire work is not based on a single level of classification but on a stratification of increasingly detailed operations, providing the possibility of correcting any inaccuracies in subsequent steps, with more suitable point intensities.
From the following images it is possible to appreciate the use of the MLMR approach in the case of a section of the Cathedral's intentions. At the first level (left image) there is a 50mm resolution cloud in which the main elements were segmented. At the second level (right) the classes of pillars and vaults derived from the previous level were in turn segmented into other sub-categories of elements.

fig.5 Example of different layers of classification

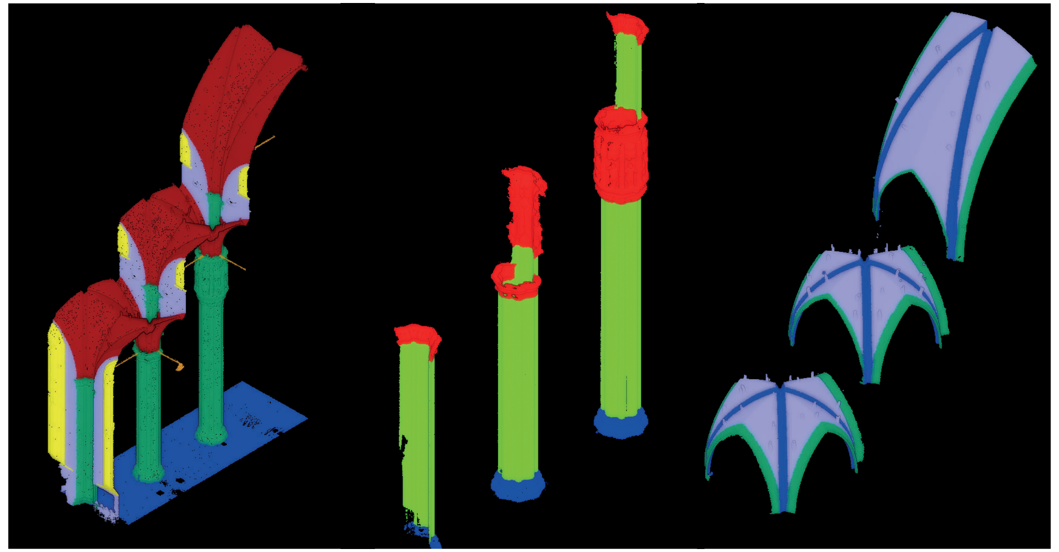

fig.5

FEATURES AND PERCEPTURAL FIELD

Classification by Machine Learning exploits the Random Forest, which learns from study models what characteristics the points belonging to the different classes have, and then analyses the data to be segmented and decides, based on the geometric characteristics it analyses, which category to assign each point in the cloud to.
It has previously been stated that classification is based on the recognition of common geometric links among points belonging to the same architectural element to be identified. The fact that clouds are composed of points with only spatial (x, y, x) and physical (intensity, reflectance, etc.) characteristics means that they do not carry, by default, information on the geometric relations among points; they instead constitute entirely separate entities.
For each point, we calculate a weight that measures the presence (meaning the manifestation) of the feature. Our feature classification is based on estimates of surface variation using local neighborhood covariance analysis. It is possible to apply a classification approach that works directly on point clouds, analyzing the effectiveness of geometric covariance features calculated on spherical neighborhoods at various radius dimensions to support the classification.
In order to assign geometric features to each point, it is necessary to take advantage of a Cloud Compare tool that allows us to compute, for each point, the incidence of the geometric features as n. perceptual fields that correspond to the survey radius to be considered.

The tested properties refer to the covariance matrix calculated within a local area of a point in 3D space. The values of the properties indicate the main linear (1D), planar (2D) or volumetric (3D) structure of the point cloud in its vicinity. (12) (Grilli, E., Farella, E. M., Torresani, A., and Remondino, F.: "Geometric Features Analysis for The Classification of Cultural Heritage Point Clouds", Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W15, 541-548, 2019. They also provide additional features and help to discriminate planes, edges, angles, lines, and volumes. These features describe the local spatial distribution of 3D points. (6) (Nesrine Chehata, Li Guo, Clément Mallet. Airborne Lidar Feature Selection for Urban Classification Using Random Forests. Laserscanning, Sep 2009, Paris, France).

The Perceptual Field

The diagrams below demonstrate that, depending on the perceptual field, in the form of a radius, the behavior of the points changes in relation to those contained within this set. (et. al. Rusu, R. B., Blodow, N. and Beetz, M.: "Fast Point Feature Histograms (FPFH) for 3D registration," 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, pp. 3212-3217, 2009)
As mentioned earlier, this is a three-dimensional environment in which the points assume a scattered order in the reference system, so the perceptual field actually a spherical field calculated with a ray with center at the point being investigated. (fig.6)

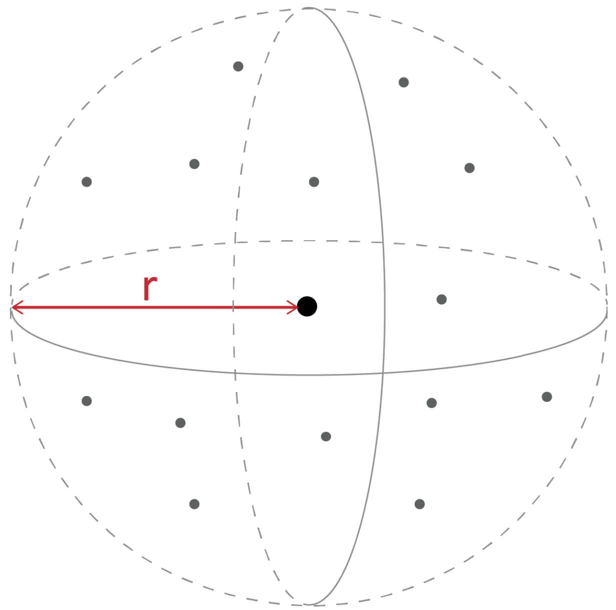

fig.6

aking Verticality as an example, with a radius of 0.50 cm the point under study shows a constant increasing verticality in relation to the other points under study within the radius, so the characteristic verticality of the point under analysis will show a rather high value. Analyzing the verticality of the same point with a larger radius of analysis, one appreciates that the point in question has a softer behavior than before; since the added points soften its verticality, it will show a lower value than before.

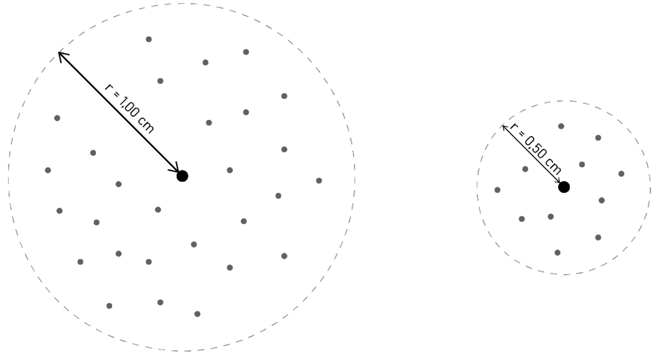

fig.7

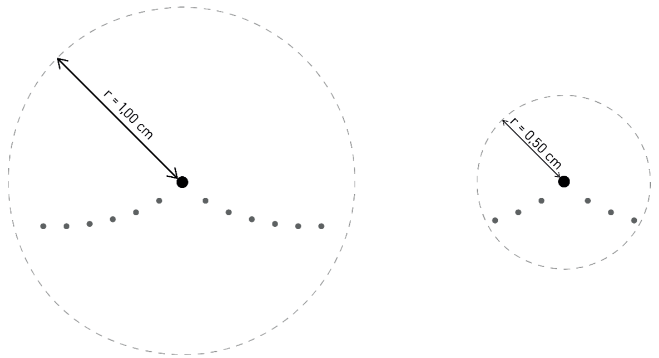

fig.8

fig.6 3D rapresentation of perceptural field of a point

fig.7 **Plan** scheme of perceptive field

fig.8 **Section** scheme of perceptive field fig.9 Graph showing the directly proportional relationship between the size of the radii and the number of points included

fig.10 Graph showing the directly proportional relationship between the increasing of the accuracy and the time necessary to the process

As illustrated in the diagram 1, an increase in the spherical perceptual field, within which to investigate the geometric relations of a point with those around it, corresponds to a logical exponential increase in the quantity of points that will be included within this range. Consequently, this leads to an increase in the accuracy of the final result given the enormous amount of geometric information among the individual points. It is clear, in this sense, that there is a radius limit below which it is appropriate to study the relationships between the points (green dot), while beyond which one would obtain results with too many points. As we will consider later, an increase in the information contained in the model, more often than not, corresponds to a difficulty for the algorithm in identifying which are the best characteristics to exploit in order to determine the class in which to include each point. This is because the Random Forest algorithm will use the information from every radius of investigation chosen, even those that are not really useful at each individual scale of the process. For this reason, it is advisable to analyze and prepare a data with appropriate information to the scale of the elements to be classified.

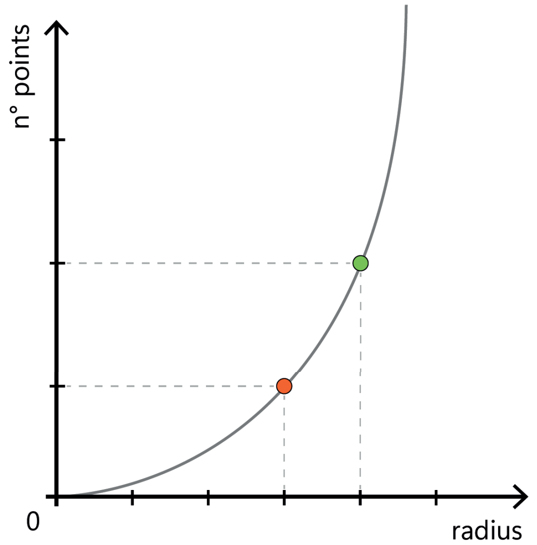

fig.9

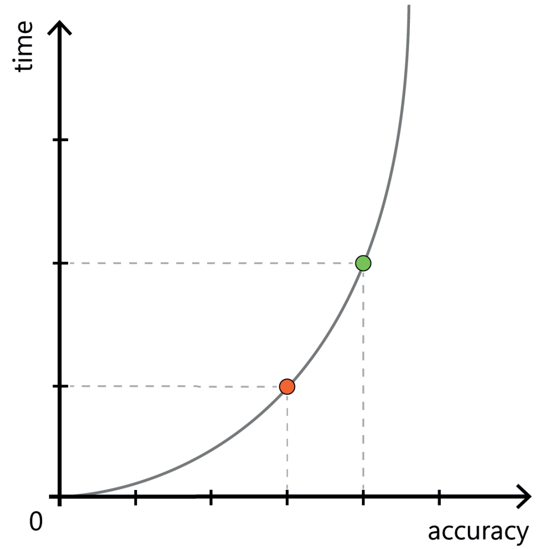

fig.10

fig.11 Spectrum of the presence of a feature in a point

fig.12 Anisotrophy witht radius 0.1m

fig.13 Anisotrophy witht radius 0.5m

The characteristics selected to achieve our aim are: Anisotropy, Planarity, Linearity, Surface Variation, Sphericity and Verticality. All the following examples are based on a 50mm resolution point cloud.
As the graph below shows, the presence of a feature manifests itself in direct proportion to the variation of the color in the spectrum.

fig.11

Anisotropy: property whereby the value of a physical quantity, all other conditions being equal, depends on the direction being analyzed. From the comparison of the following images, it is possible to understand the behavior of this feature as the radius varies. In the first case (anisotropy 0.1m) almost all the elements that make up the interior of the windows and the details of the capitals are visible. In the second (anisotropy 0.5m) the window details almost completely disappear, and the capitals are distinguishable as a single element with no internal details.

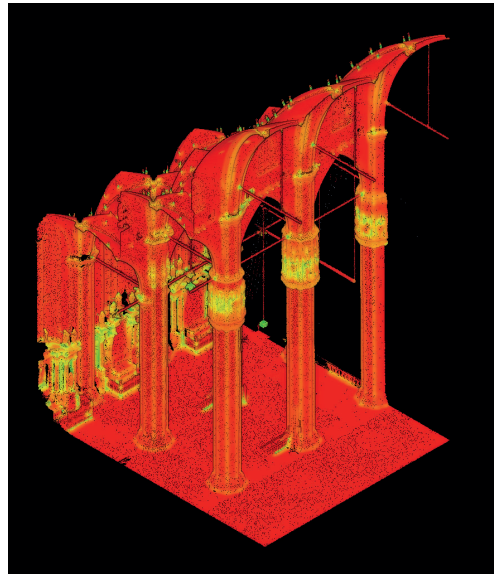

fig.12

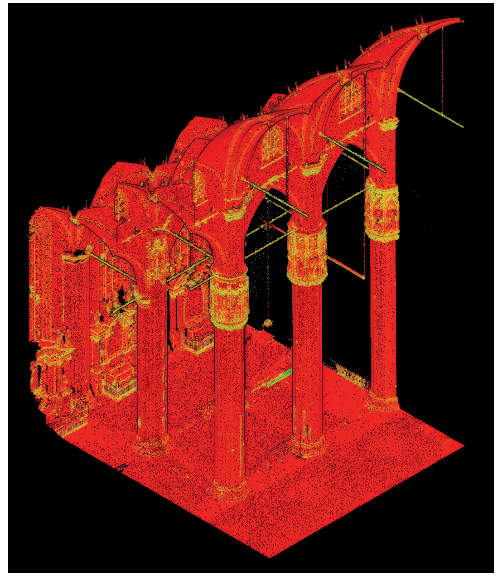

fig.13

fig.14 Planarity witht radius 0.1m

fig.15 Planarity witht radius 0.2m

Planarity: property by which a point 'P' lies on a given reference plane, generated on the average basis of the other points taken into consideration. By comparing the following figures, it is possible to understand the behavior of this feature even at the smallest radius. At a radius of 0.1 m (figure dx) almost all points show a non-linear behavior due to the type of materials and the wear of the materials with which the Dome is made. At a radius of 10 cm, in fact, every point will show this non-planarity due to the imperfections of the materials and not to the composition of the architectural elements. By increasing the radius to 0.2 m (second figure), it is possible to distinguish the shapes of the main architectural elements, since the radius makes it possible to see the relationships between several points that lie on the same plane.
With this radius, for example, it is possible to distinguish points along the shaft of pillars or along vaults as planar due to the greater number of points examined.

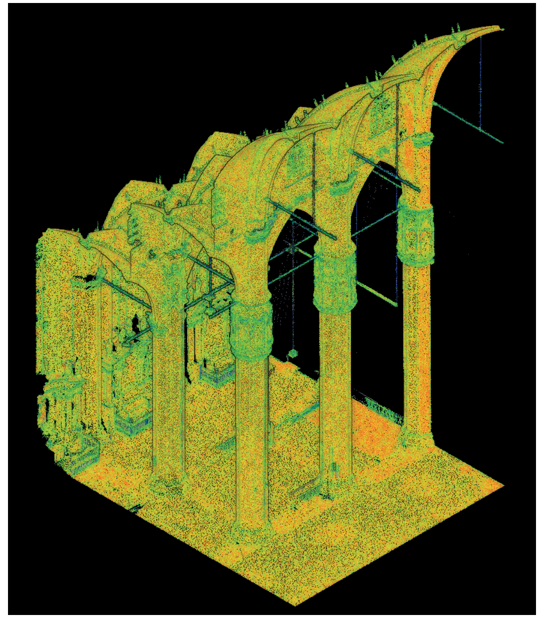

fig.14

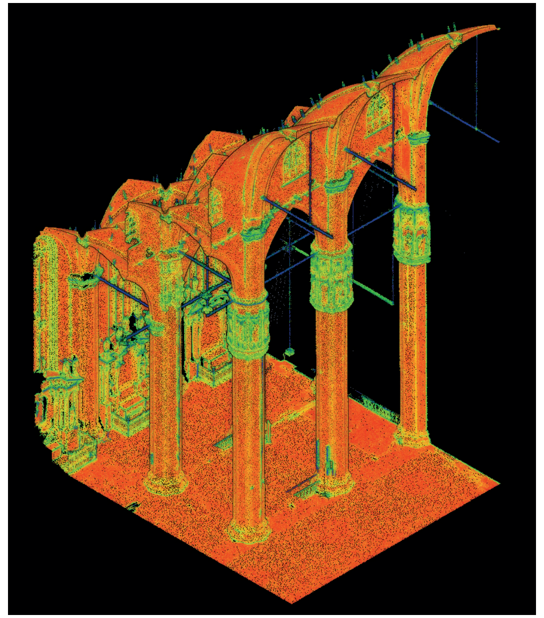

fig.15

Linearity: the property of a point 'P' to occupy an aligned position in relation to a reference generated on the base of the average of the other points considered. The representation of this feature is similar to the previous one but, since it is referred to a linear reference and not a plane, it provides more useful information in the case of long objects that develop in a linear manner.

fig.16 Linearity witht radius 0.1m

fig.17 Linearity witht radius 0.3m

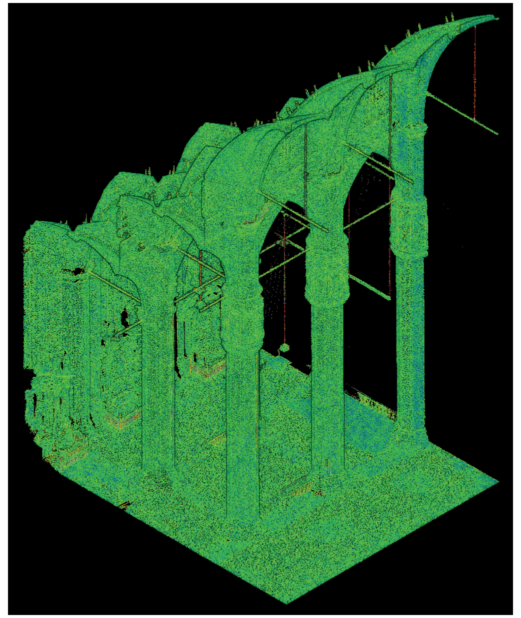

fig.16

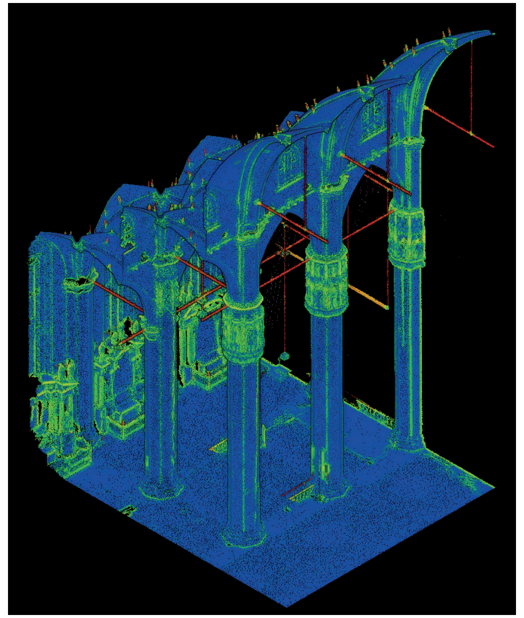

fig.17

fig.18 Surface Variation with radius 0.1m

fig.19 Surface Variation with radius 0.5m

Surface variation: the property of a point 'P' to vary its position in relation to others in the perceptual field by taking only its position as a reference. This feature is one of the most useful in case studies such as this, as its minimal radius allows the visualization of many decorations and non-architectural objects. By extending its radius to 0.5m (second figure), or in general to increasing values, it offers the possibility of neglecting some elements smaller than its radius from the visualization.

Sphericity: property of a point 'P' to move in a spherical line in relation to the others contained in the perceptual field of investigation. This feature offers the possibility of distinguishing, at various scales, the spherical course of sets of points. With a radius of 0.1 m (figure 1), this feature provides a detailed visualization of all decorative and small elements, such as statues in capitals or window frames, while it flattens out all groups of points that make up large elements such as pillar shafts, walls or vaults. By expanding the radius to 1.0 m (figure 2), the Sphericity feature starts to show the shape of the grooves that characterize the section of the shafts given their diameter of approximately 3 m.

fig.20 Sphericity witht radius 0.1m

fig.21 Sphericity witht radius 1.0m

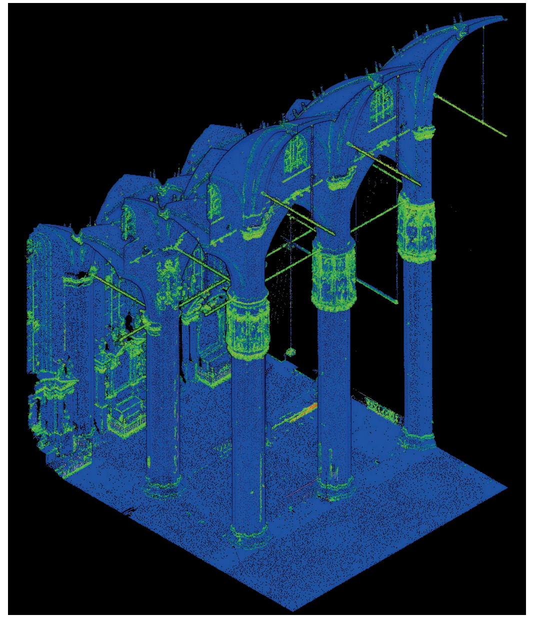

fig.18

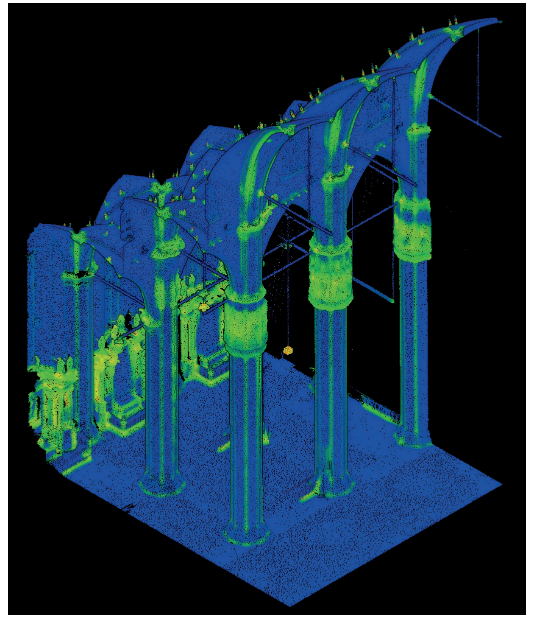

fig.19

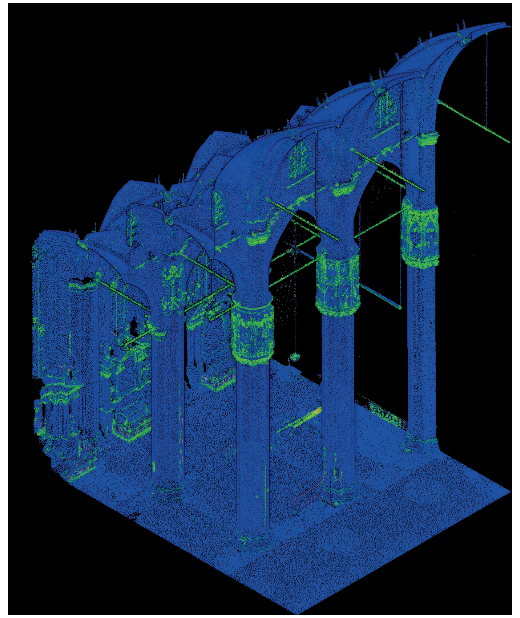

fig.20

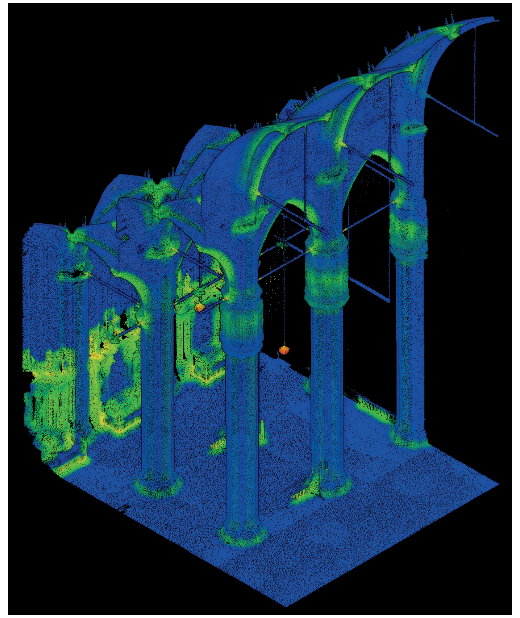

fig.21

fig.22 Verticality with radius 0.1m

fig.23 Veritcality with radius 1.0m

Verticality: property of a point 'P' to move along the vertical axis z in relation to the others contained in the perceptual field of investigation. As can easily be understood, this characteristic is manifested by the presence of points along the z-axis within the perceptual field. At a radius of 0.1 m (figure 1), the statues of the giant capitals are visualized clearly enough to perceive their shape, while at a radius of 1.0 m (figure 2), the capitals are barely distinguishable from the shaft of the pillars. Using such a high perceptual field, larger than the size of objects such as the capital statues, the feature will show the points of the capital as all vertical to each other, as it will be inclined to neglect all variations along the z-axis smaller than the set radius.

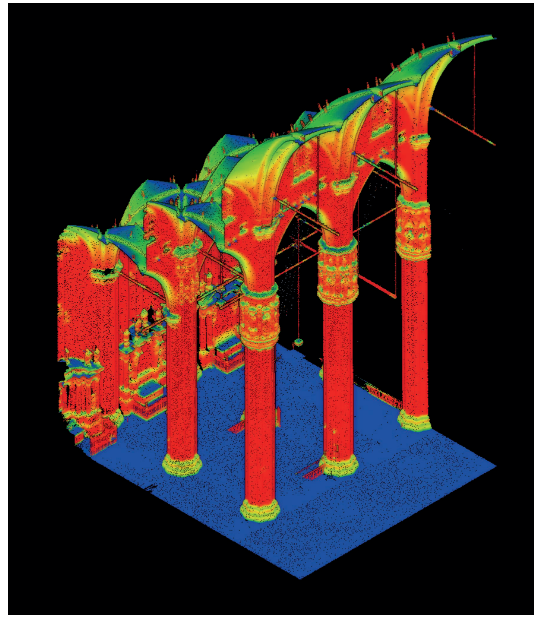

fig.22

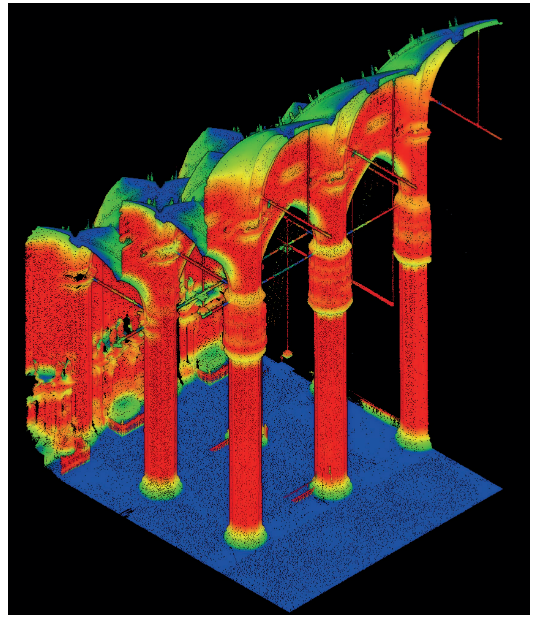

fig.23

These six selected features were calculated on spherical neighborhoods at various radius dimensions, to explore the different relationships as a function of the different geometric properties of the elements under study. (10) (Grilli, E., Farella, E. M., Torresani, A., and Remondino, F.: Geometric Features Analysis for The Classification of Cultural Heritage Point Clouds, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII- 2/W15, 541-548, 2019).

STUDY MODELS: TRAINING SET AND EVALUATION SET

In Supervised Machine Learning, the operator provides the algorithm with a set of labelled data from which it learns the features to be searched for in the model that has to be classified. The input data, classified manually, is called the "training set" and it is useful to provide information regarding the geometric characteristics of the points belonging to the different classes. On the other hand, the "evaluation set" refers to the data used to have an initial feedback on the ability of the code to recognize the elements. Lastly, the data on which the algorithm will then process the result is called the "dataset". In concrete terms, the code will learn from the training set which type of points belongs to the different classes, and on the basis of their geometrical features, it will analyze the features of each point of the dataset and will place them in the most appropriate category provided by the input data.

In light of this, the importance of choosing which part of the initial data has to be extrapolated in order to classify the elements manually becomes clear. Indeed, it must be a representative portion in which all elements to be classified are well defined and repeatable throughout the entire model. Considering the example of pillars in the interior classification, we can appreciate three different types of pillars repeated for the entire nave. In this case, it will be appropriate to define a training set in which all the different types of pillars appear; otherwise, if even one is missing, the code may run into problems in classifying that pillar, since redundant information regarding the specific composition of that type of pillar would not be provided.

fig.24 Example of extrapolation and manual classi4icationof study model from the dataset

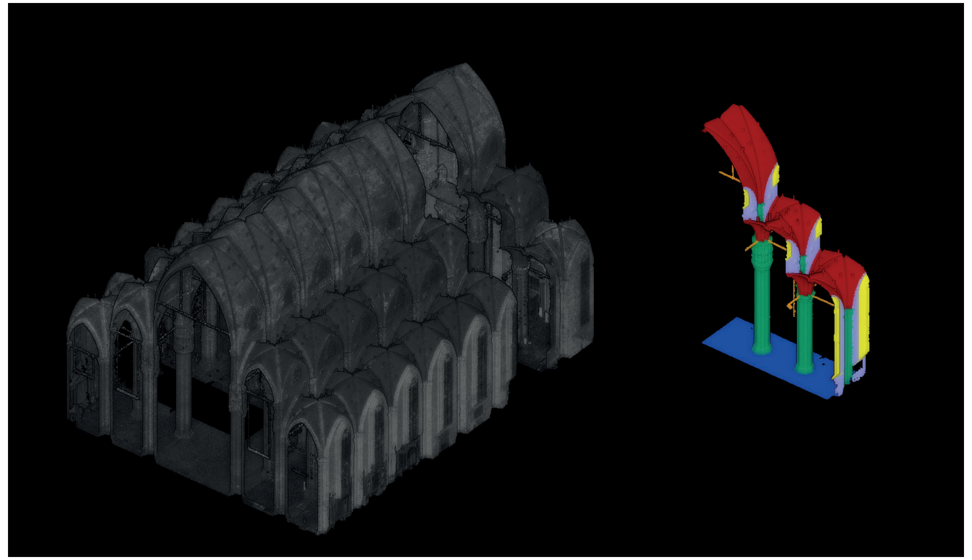

fig.24

CONCEPT OF GENERALZIATION

The main objective arising from this is to minimize the time required for manual operations (carried out by the operator) and entrust the machine with as much work as possible so as to reduce processing times and obtain accurate, more objective and repeatable results over time. In this perspective, the possibility of generalizing operations plays a role of primary interest: the correct approach and the quality of realization of the study models allow them to be generalized to other portions of the cloud or even to other architectures. In our case, this means being able to work in such a way that the study models generated for the classification of the nave can be repeated for other portions of the interior. As this is one of the first cases in which this type of work is carried out on a Gothic architecture of this size, it should be clarified that the main objective will not be to generalize the classification of the data set using a single model for the entire church. Rather, it is to investigate what is the best balance in the MLMR approach between manual and automatic operations in order to obtain the most accurate results possible with the least amount of operator time. The complexity of the Cathedral's architectural elements, as well as the unique ornaments and decorations on almost every surface of the church, highlight the necessity of an initial subdivision of architectural types between interior and exterior. In fact, inside the Cathedral, despite a strong presence of decorations such as statues and frames, it is possible to clearly recognize the main architectural elements that compose it, since most of its pure forms are on display. As far as the exterior is concerned, particularly the masonry, the clear difference from the interior is evident: here, in fact, every element that makes up the façades is completely covered in relief decorations.

PART II Chapter 3: The Case Study

DESCRIPTION OF THE SOURCE DATA

The starting point of the classification work was a point cloud with 5 mm resolution of the entire cathedral, both interior and exterior. For the purpose of this research, the work was divided into two main parts: interior classification and exterior classification. For the classification of the interior, the data was divided into three parts: the nave, the transept and the apse. Architecturally and structurally, all the interiors of the Cathedral can be identified as repeatable elements, since, for example, the vaults have the same composition in the nave, transept and apse. What varies is the way in which the macro-elements are positioned in relation to each other. From the following sections one can see the repetition of the same architectural elements such as pillars, walls, windows and vaults in both areas of the interior. Similarly, one can see that the longitudinal and transverse compositions of the interior do not correspond. In the nave we have three orders of windows, the main ones at the bottom on the perimeter wall, the central ones on the pillars dividing the nave from the side and middle aisles, and finally the third order, higher up, on the division between the middle and middle aisles.

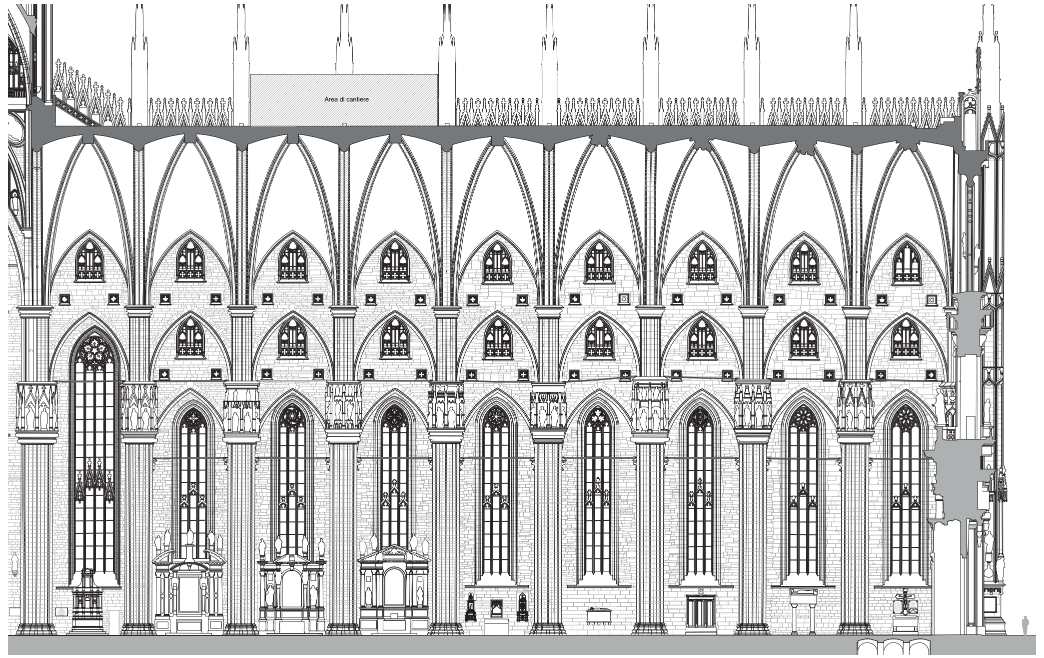

fig.25

fig.25 Secgtion of central nave

As for the cloud part of the apse, it was isolated from the rest of the interior due to its unique composition and not suitable for automatic classification. For this reason, it will not be investigated.
In general, the data on the interior of the church is presented as a fully complete and homogenous cloud, thanks to the survey work carried out to perfection in recent years.
The starting data for the classification of exteriors is a cloud detected by TLS and interpolated with RGB information from photogrammetry. This additional information from the cloud had a twofold advantage: for manual element recognition operations and as additional data for the learning and subsequent automatic segmentation phase of the algorithm. If from the point of view of the information contained in the data there was an implementation with respect to the interior, from the point of view of the completeness of the survey there was no such correspondence. The exterior of the Cathedral, due to the larger number of decorative elements and direct exposure to the weather, is the part most subject to conservation work. A large part of the north wing of the transept is affected by scaffolding that obscures its surface, making any relief work impossible. For this reason, the external wall datum is missing from this portion of the Cathedral.

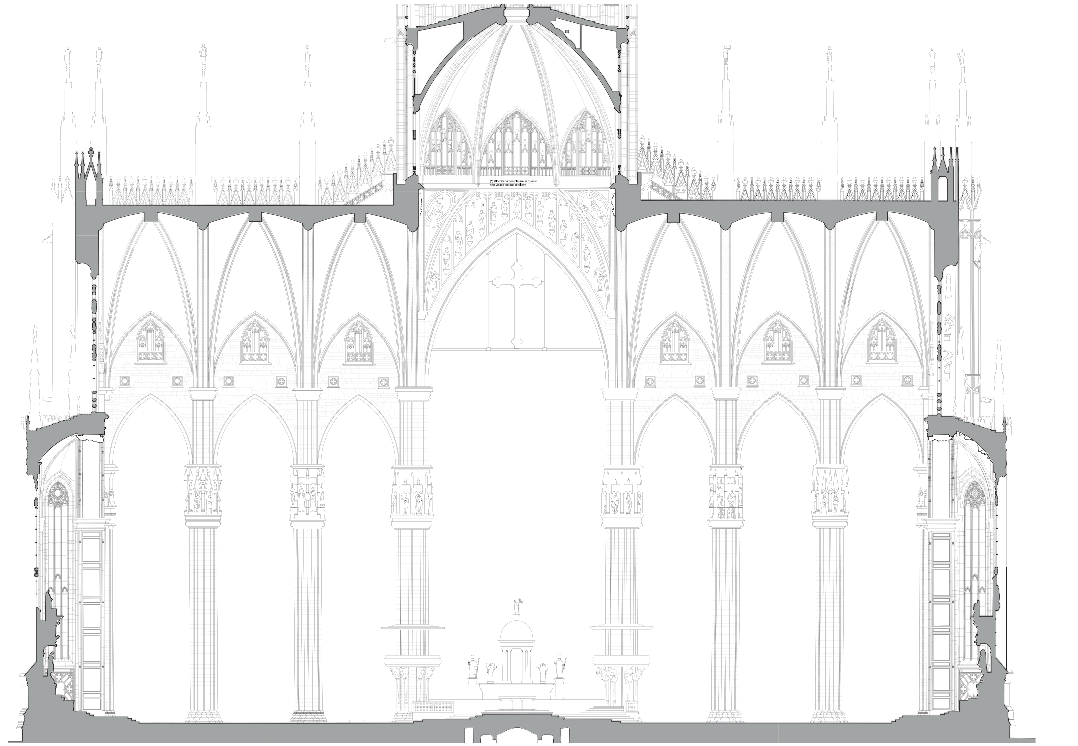

fig.26

fig.26 Section of transept

GENERIC APPROACH FOR CLASSIFICATION

fig.27 Pipeline of generic classification approach

This system entails the repetition of the following pipeline:

1) Depending on the size and characteristics of the macro-categories to be identified, the starting cloud is subsampled to lighten its size. It is essential to carefully evaluate the appropriate resolution to be used in this first phase, as a cloud that is too detailed would unnecessarily increase the processing time, while a cloud that is insufficiently detailed would not allow the operator to correctly identify the elements. Based on this, the first level of classification is performed.

2) The classified cloud obtained from the previous step is interpolated with one of greater intensity, in order to transfer the semantic information by transferring the classes determined in step 1 to a certain circle of points in the higher resolution cloud.

3) The macro categories of interest are in turn subdivided to identify their components (e.g. vaults into ribs, arches and sails).

4) The interpolation of the results of step 2 is repeated again with a new, higher-resolution cloud for the classification of further classes.

GENERIC CLASSIFICATION PIPELINE

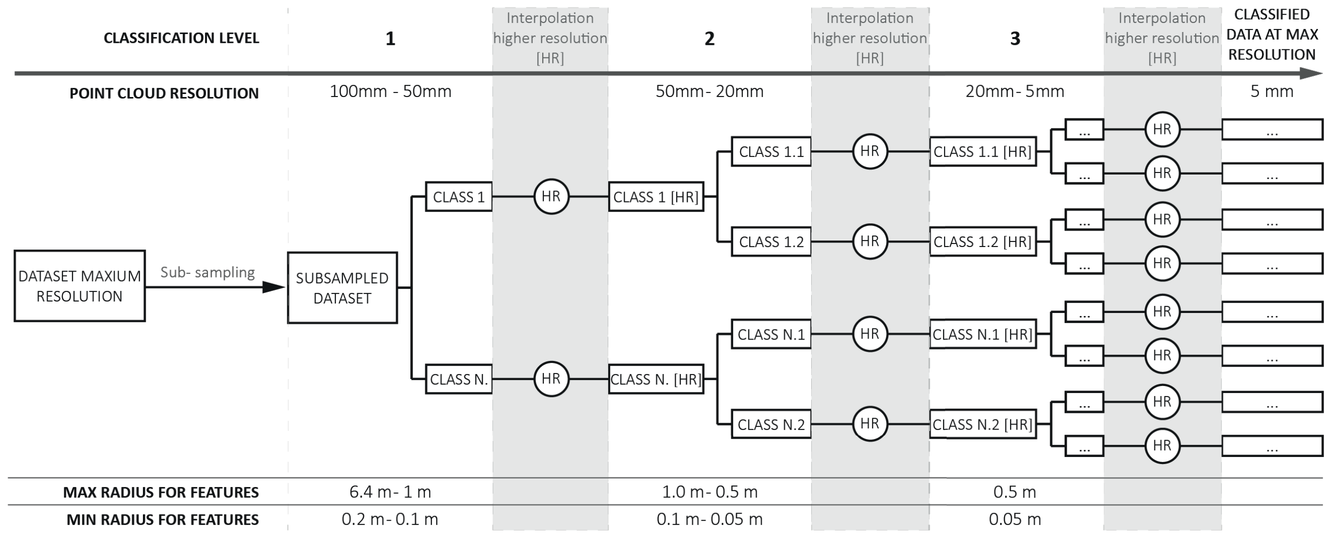

fig.27

CLASSIFICATION OF INTERIORS

fig.28 Scheme of the pipeline for internal classification

Interior Classification Pipeline

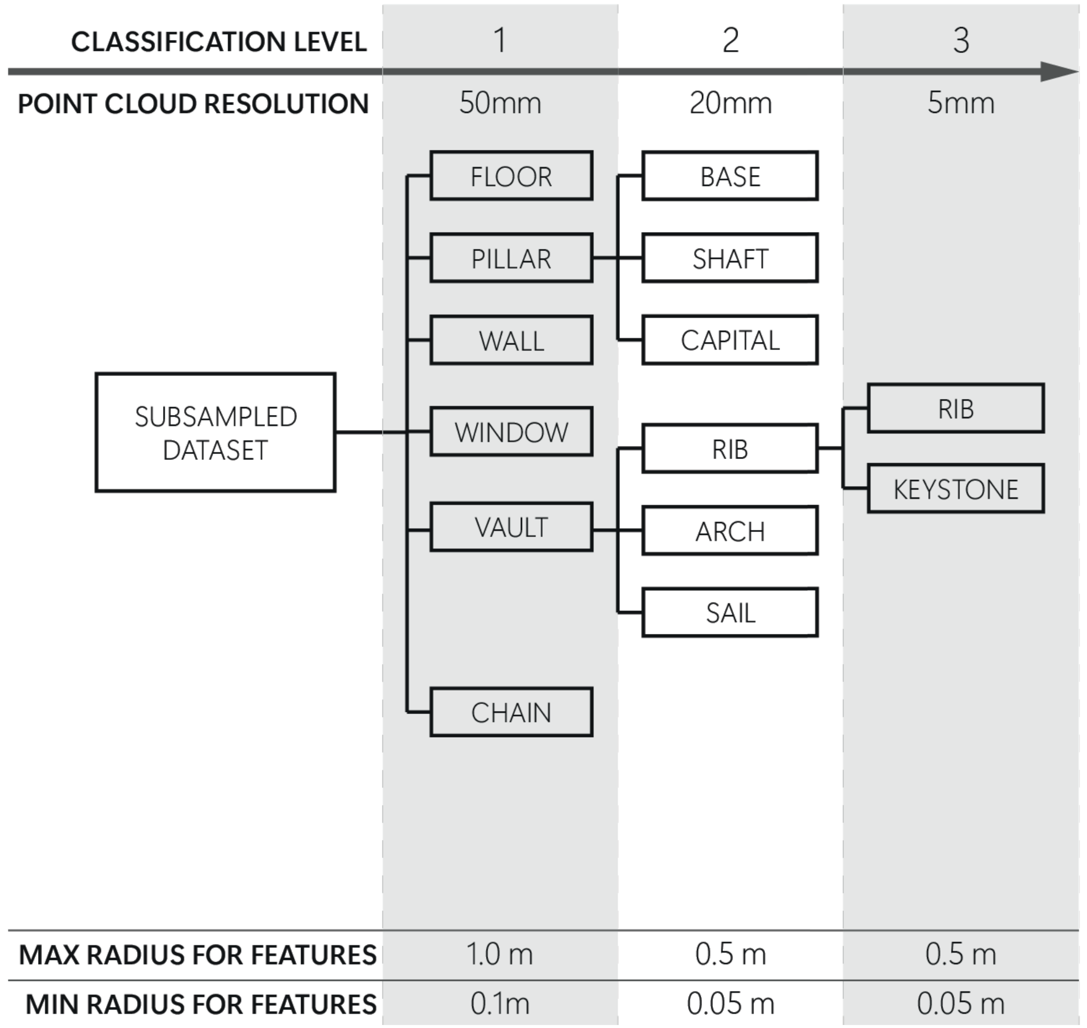

fig.28

CLASSIFICATION LEVEL 1

As already mentioned, the main factors to consider when choosing the resolution of the starting data are the size of the elements to be segmented and the presence or absence of details of different sizes. The classes to be segmented in this first level are: floor, pillars, walls, windows, vaults and chains; they can be considered as large elements if compared to the general size of the decorative elements inside the Cathedral. Following this analysis, in order to perform the automatic classification of the main elements, without the decorations hindering their recognition, a point cloud of 50 mm resolution was formed.

In fact, the main elements to be classified are well distinguishable with this point density, as each object belonging to a class has a good number of points to identify its shape, and similarly, the details and ornaments belonging to the main classes should not be composed of a sufficient number of points to be recognized as separate entities.

fig.29 Detail of central nave with characteristic surface variation 0.1m

fig.30 Detail of central nave with characteristic surface variation 1.0m

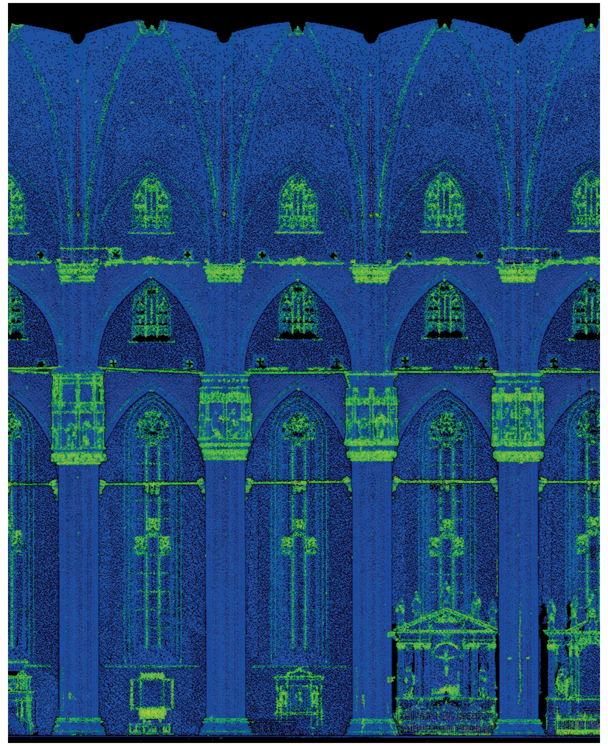

fig.29

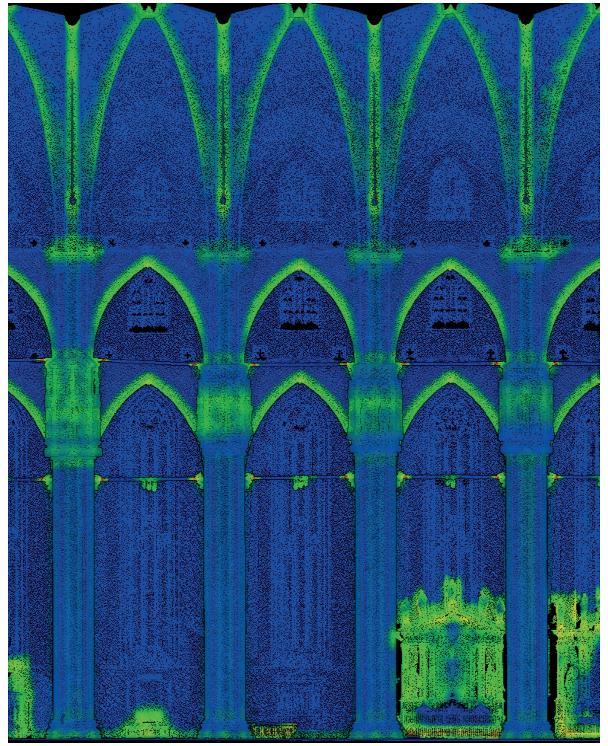

fig.30

These two images depict the same portion of the Cathedral with the characteristic surface variation, the one on the left with a perceptual field of 0.1 m and the one on the right with a perceptual field of 1 m. From this comparison it can be appreciated that the details on the columns are not large enough to be considered as a separate entity from the rest of the element. Examining the image with a perceptual field of 1 m, in fact, it is not possible to precisely identify the details on the pillars, which means that, with this radius, it is possible to distinguish them as a whole from the base shaft and capital. On a practical level this means that if visually one cannot see certain details (e.g. the statues on the capitals) then the code, using this information, will also not be able to distinguish these elements but will be able to identify the pillar as a whole.
At this stage, the main obstacle is the large number of elements that are present in the scene but not connected to the structure and thus to the main architectural elements to be segmented. Objects such as parapets, benches and altars attached to the perimeter walls, are to all intents and purposes obstacles to the correct identification of the forms of the primitive architectural elements. The presence of such elements must therefore be considered in order to find a standardized solution to this type of problem. To simplify subsequent descriptions, these objects will be identified as non-architectural elements.

The first issue tackled was that of non-architectural elements. One solution was to identify and classify all these elements with an additional class to the previous five. According to initial estimations, it was assumed that these elements were sufficiently homogeneous in shape and size to allow the code to identify them as separate entities from the main architectural elements.
The main risk of this approach lies in the position of the non-architectural elements in the reference system in relation to the other objects. The following image depicts the training set and evaluation set realized for this first approach: in green the pillars, in purple the walls, in yellow the vaults and in orange the non-architectural elements. The frontal view of these models clearly shows the fact that all the architectural elements develop for varying heights intersecting each other, while the non-architectural elements all insist in the same band along the z-axis. The presence of a class of elements with such a strong predominance in a single range of coordinates along the vertical axis leads the code to learn that almost all points located in that band belong to the category of non-architectural elements. In the second image, one can see the obvious misclassification of points belonging to the category of walls that have been confused and segmented as non-architectural elements. In this case, it is evident that the classification problem is related to the z-coordinate, since all the points of the perimeter walls below a certain height (the same height as the altars) have been assigned to the category of non-architectural elements.

This analysis brings to attention the fact of how important the incidences of geometric features used for automatic segmentation are and that it is not always possible to act on them to change the final result. In this case, it would have been pointless to remove the z-coordinate from the list of features, as it would have improved the classification of the wall bases to some extent, but would have created numerous problems in the upper part of the cloud, where height is crucial for the identification of pillars and vaults in particular. (Demonstrative images of what happens by removing the z will follow).

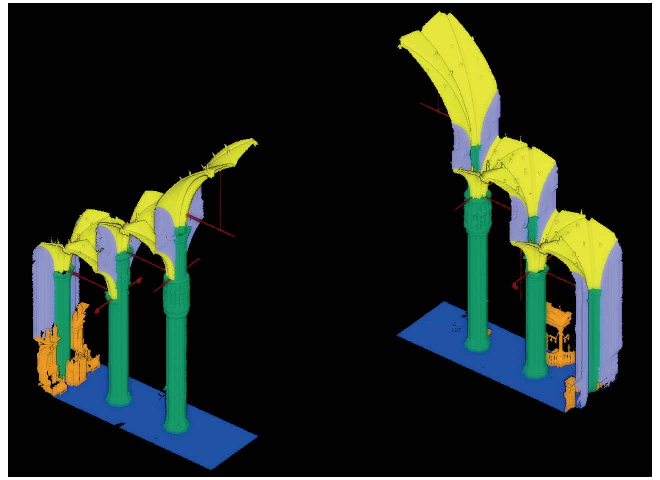

fig.31

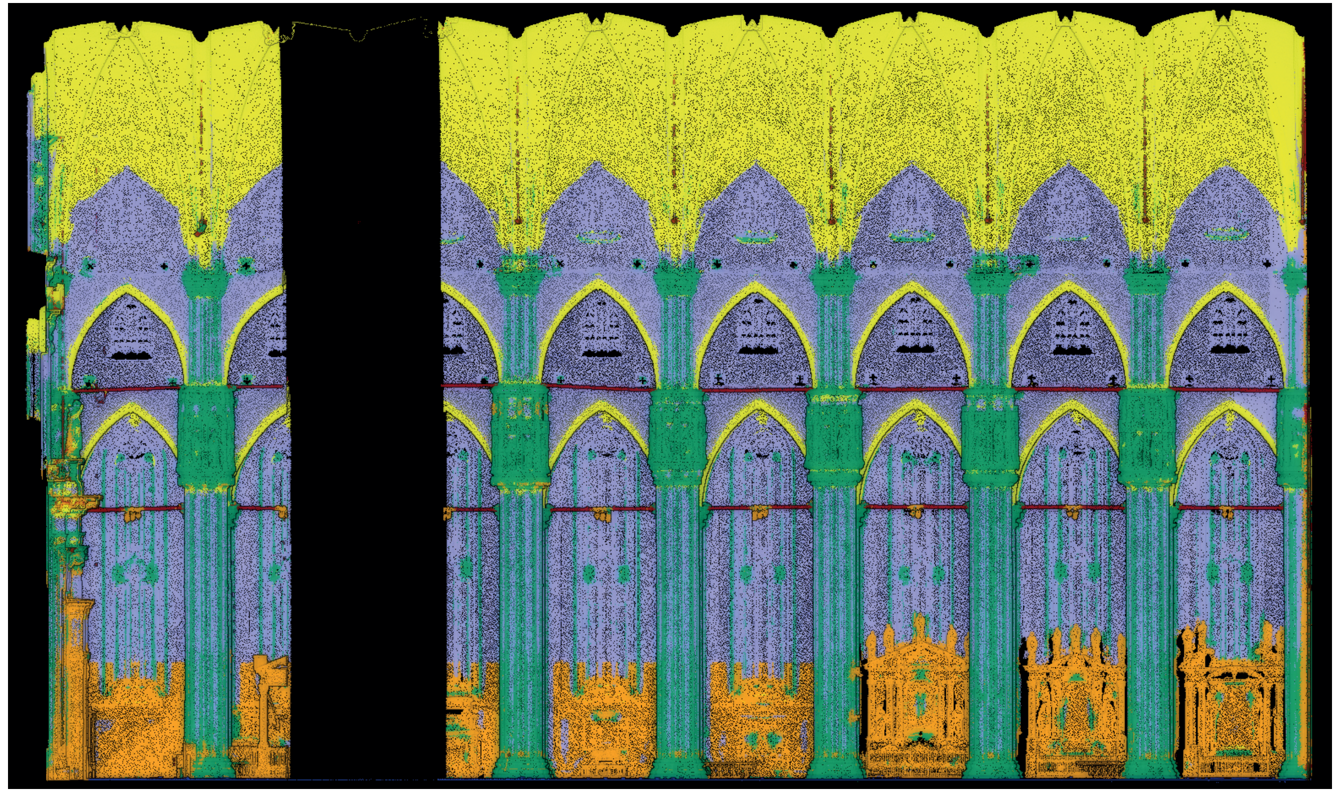

fig.32

fig.31 Training set and evaluation set with non-architectural element class

fig.32 Result of automatic classification with non-architectural element class

Table 1
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	8,651,099
Train set points	612,053
Evalu set points	611,056

Table 1

Table 2
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	35 min
Time processing	550 sec.
Manual reassignment	180 min

Table 2

Table 3
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Floor	0.99	0.99	0.99
Pillar	0.93	0.95	0.94
Wall	0.91	0.86	0.90
Vault	0.97	0.98	0.98
Chain	0.93	0.79	0.86
Non arch. elements	0.86	0.79	0.89
Accuracy avg	0.95		
Macro avg	0.95	0.89	0.91
Weighted avg	0.94	0.95	0.96

Table 3

The second solution consists of removing the architectural elements from the study model, as if they were part of the cloud noise to be segmented, so that they are not included in the automatic data processing phase. This solution simplifies the identification of the main categories by the algorithm, as it lightens the processing of a series of information related to very different elements.
A second point that had to be considered in the design of the work was how to consider windows at this level, whether as a class in their own right or whether to include them in the class of walls for the time being. In fact, due to their position and composition, the windows do not differ greatly in terms of their geometric characteristics in relation to the wall; the depth of the frame in relation to the glass does not differ much from that of the decorative profiles of the internal wall façade. As can be seen above (in the image of the interior set to Surface Variation 1 m), in fact, no distinction is made between the wall and the windows, but the latter appears as a variation-free element, totally represented in blue.
On the subject of windows, the main misclassification error occurs along the cornice and in the keystone of the ogival arch that closes the upper part of the window. In the first case, the splay of the cornice composed of irregular and non-continuous curves led the algorithm to classify those points as a pillar, associating with the irregular line of the nave's pillar section. In the second case, the hypothesis is that the arched shape of the upper part of the cornice, combined with the depth of the splay, led the program to place those points in the vaults category, since taking that portion of the cornice out of context one could actually identify that element as a small pointed vault, given the proportion between the depth and the development of the arch.

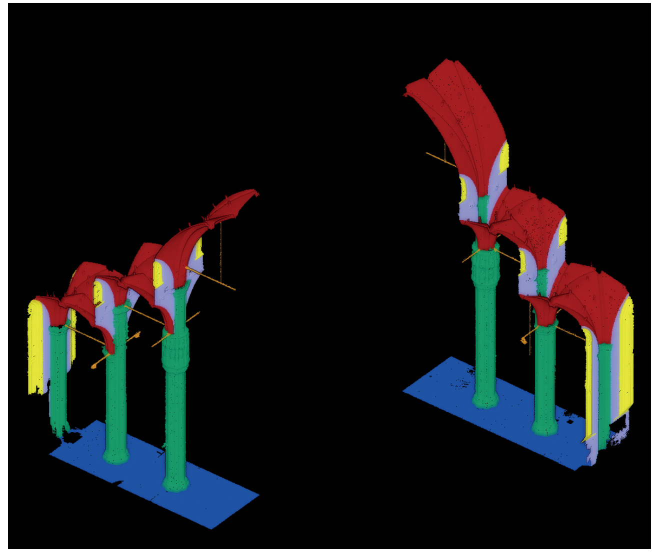

fig.33

fig.33 Training set and evaluation set without non-architectural element class fig.34 Result of automatic classification without non-architectural element class

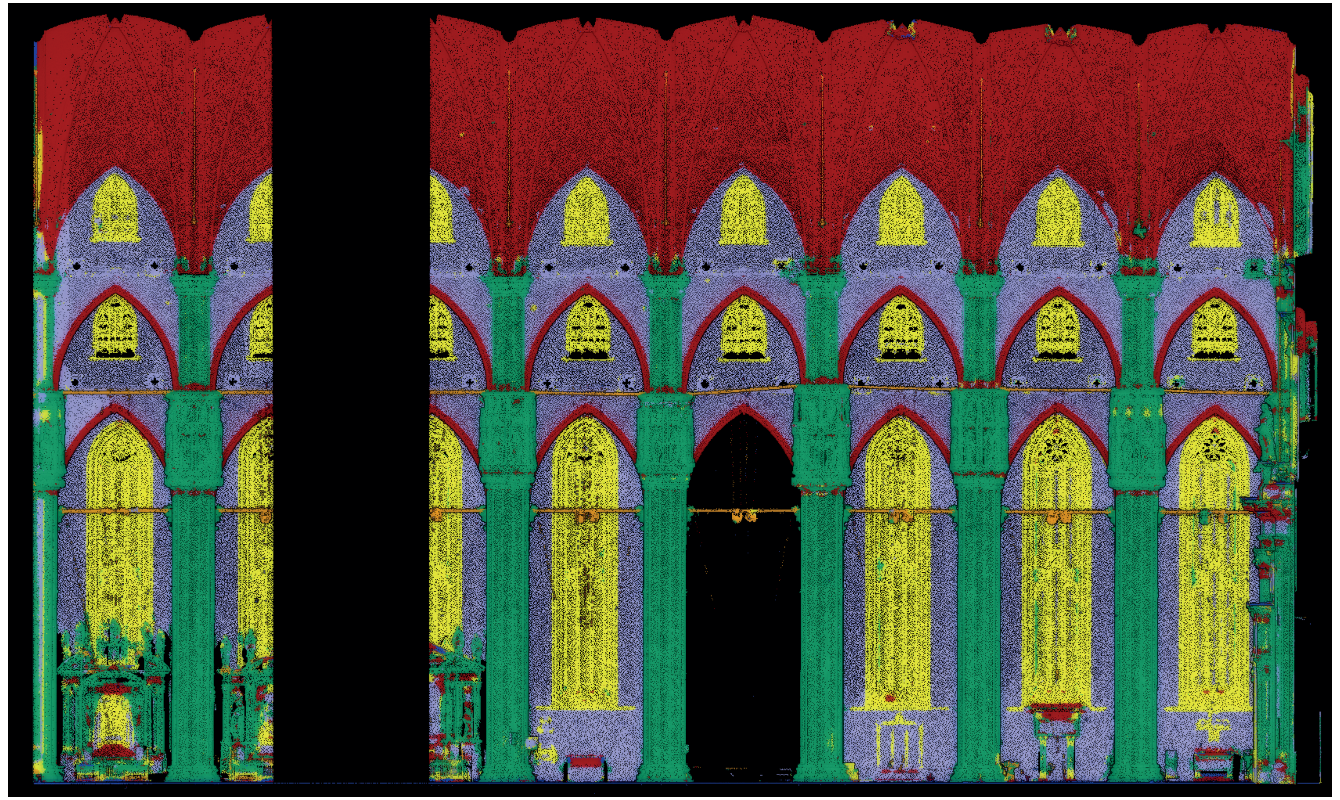

fig.34

In light of these observations, it was decided to proceed by completely eliminating the non-architectural elements within the training model and the evaluation set, so as to remove a lot of information that not only did not help to better identify this category, but also compromised the correct segmentation of the remaining classes. In this way, the non-architectural elements of the dataset ended up being classified as part of the interior. This made it easier to remove them from the rest of the data when cleaning the cloud after segmentation, and the main classes were recognized in a more defined way by the code, relieved of redundant information.

With regard to the classification of windows and walls, the confusion between window frames and pillars was corrected by adjusting the survey radiuses of the characteristics 'Surface Variation' and 'Sphericity'; by removing the smaller ones, it was possible to provide the program the necessary knowledge to better understand that only a large-scale continuous sphericity, such as the actual sphericity of the pillar, should be identified for this category.

By applying these corrections, the classification errors for this category were significantly reduced, providing a more than acceptable result to proceed to the data cleaning phase. With regard to the problems in classifying the keystones of the window arches, it was not possible to find a convincing solution that corrected this detail without affecting the remaining segmentation of the model. For this reason, despite the persistence of this error, the overall result was considered satisfactory in order to be able to proceed to the cleaning phase and the subsequent second level of classification.

fig.35 Axonometric view of the result of automatic classification

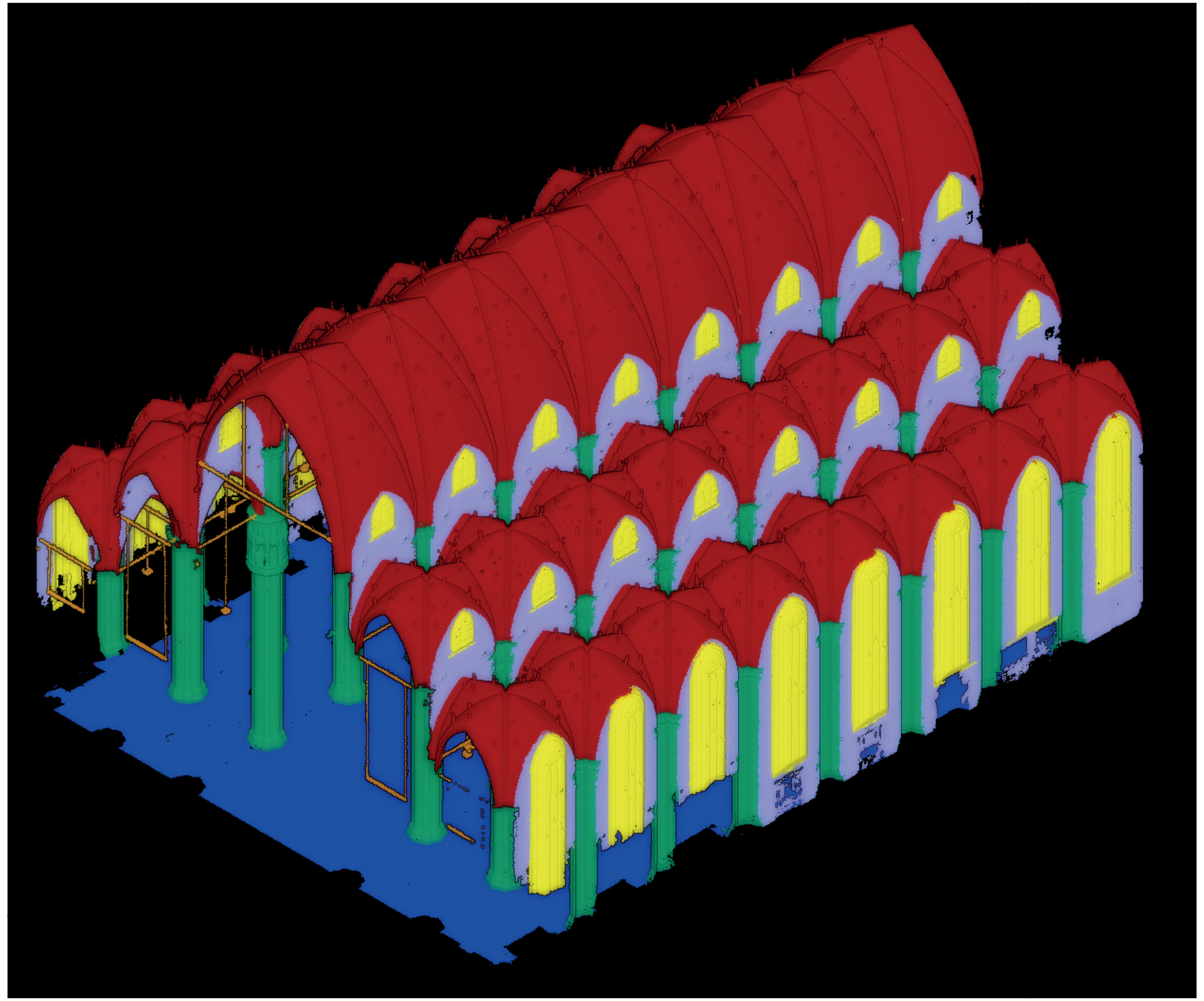

fig.35

Table 4
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	8,651,099
Train set points	515,272
Evalu set points	1,025,246

Table 4

Table 5
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	0* min
Time processing	550 sec.
Manual reassignment	50 min

Table 5

Table 6
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Floor	0.99	0.99	0.99
Pillar	0.96	0.98	0.97
Wall	0.95	0.86	0.90
Window	0.91	0.91	0.91
Vault	0.97	0.98	0.98
Chain	0.93	0.79	0.86
Accuracy avg	0.96		
Macro avg	0.95	0.92	0.93
Weighted avg	0.96	0.96	0.96

Table 6

* Used training and evaluation set realized for classification with the category not architectural element

fig.36 axonometric cross-section of the result of automatic classification

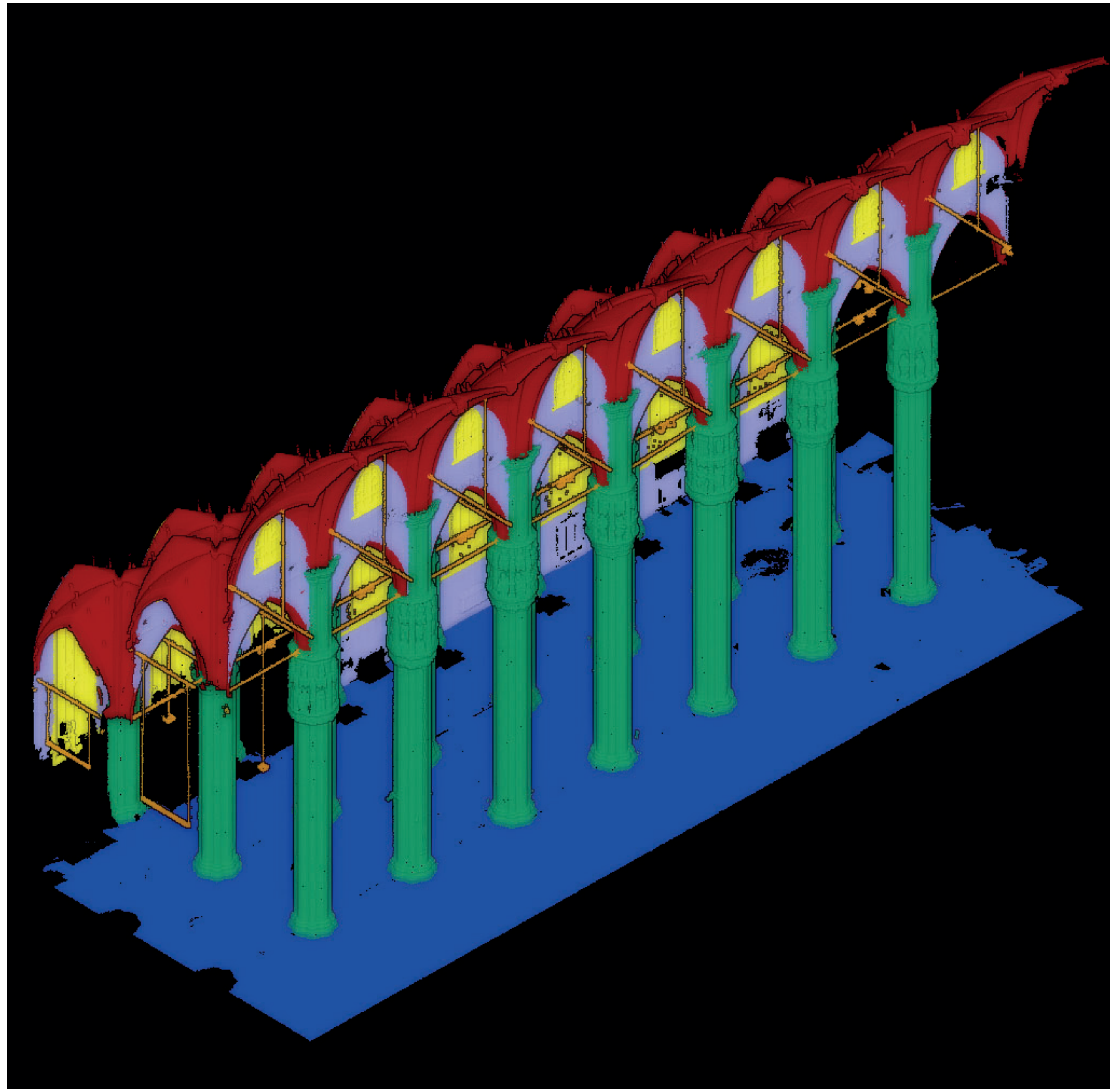

fig.36

fig.37 Floor class isolated from the final result

fig.38 Pillar class isolated from the final result

fig.39 Wall class isolated from the final result

fig.40 Winsow class isolated from the final result

fig.41 Vault class isolated from the final result

fig.42 Chain class isolated from the final result

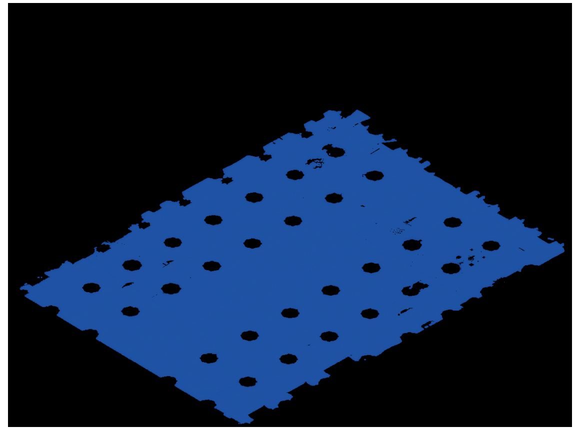

fig.37

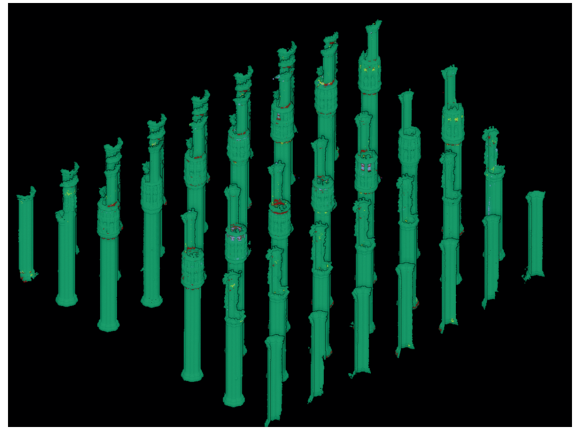

fig.38

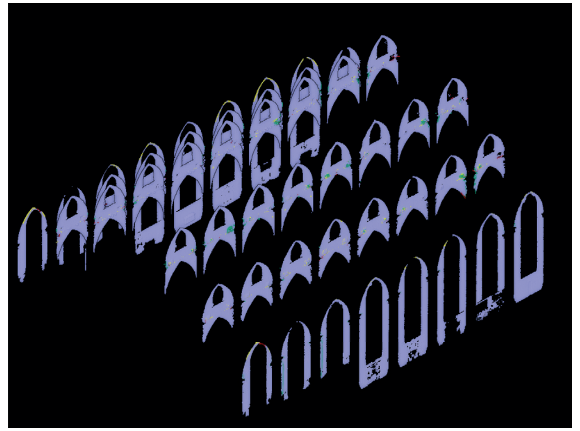

fig.39

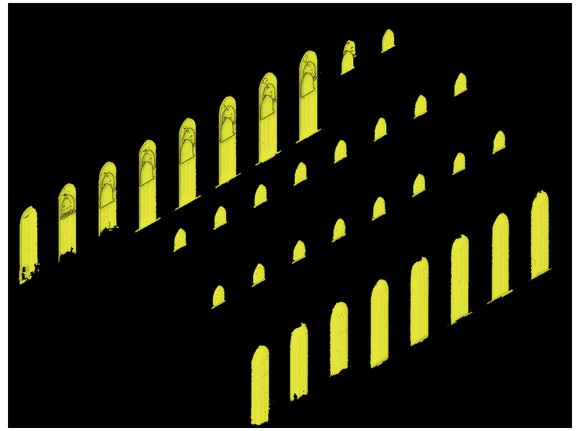

fig.40

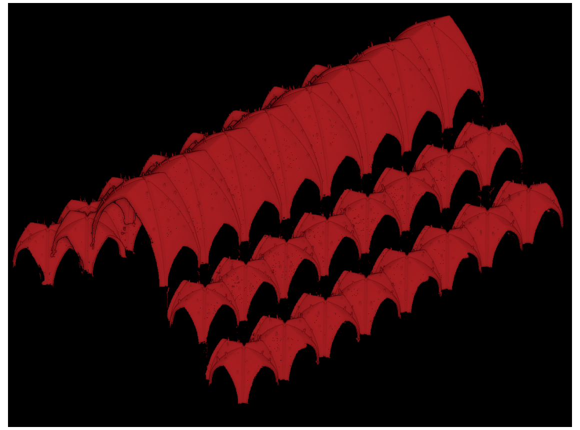

fig.41

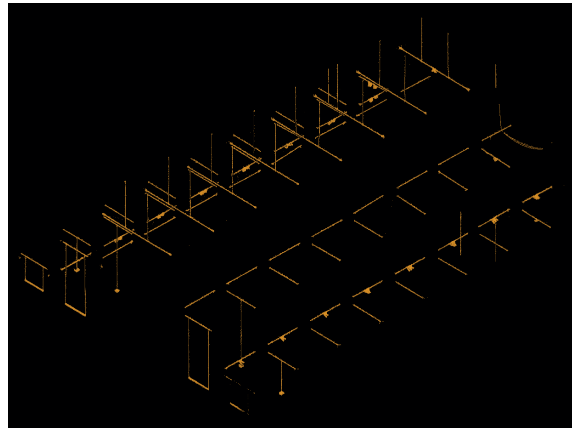

fig.42

CLASSIFICATION LEVEL 2

PILLAR

This second level of classification involves the identification of the base, shaft and capital, using a 20 mm resolution cloud.

This stage of work began by identifying the presence of three main types of columns that are repeated at the division between one nave and another. In order of increasing size, we find the first (in green in the diagram below) which is a parastra located on the perimeter wall, the second (in red in the diagram below) which divides the two side aisles, and the third (in blue in the diagram below) located between the central aisle and the side aisle adjacent to it. With the help of the following section, I will attempt to expose the compositional characteristics of the different pillars that can potentially disturb the recognition of the three subcategories. The first important distinction regards the two inner columns and the one juxtaposed to the outer wall, which is never shown as a true column but as a parasta. This at an architectural level may not be a very relevant difference, but for our work it is a substantial difference in order to identify the geometrical characteristics of the relationships between the points. The main reason why it is necessary to examine these situations is that the parasta, exiting only for part of its section from the wall, never presents the continuity of points along the diameter that we find in traditional columns. Similarly, we can note the fact that the two central columns do not end with the capital as usual, but continue along the upper wall in the form of parastas until they reach the last capital. The central column (highlighted in red) shows a further peculiarity, namely the fact that the capital at the end of the shaft does not crown the entire shaft of the column but only the half of it facing the nave, continuing instead in a single shaft all the way to the second capital on the side facing the center of the Cathedral.

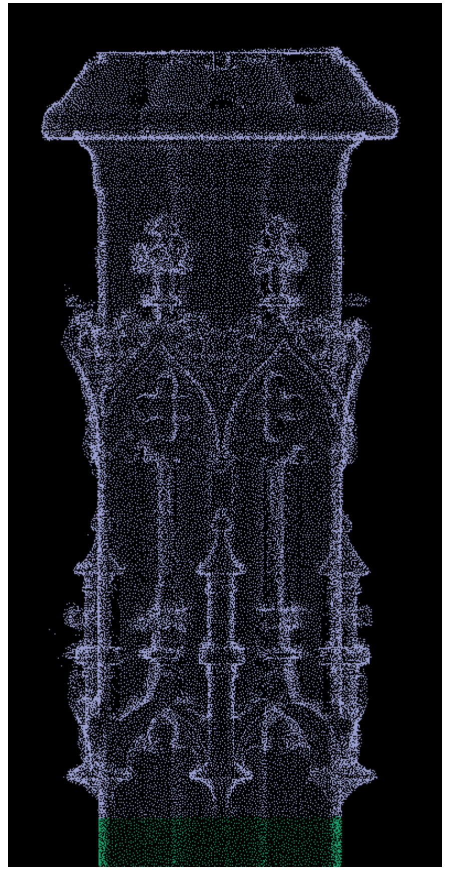

fig.43

fig.43 Detail of manual classification of the capital fig.44 Trainign set and Evaluation set for second level of classification of pillars

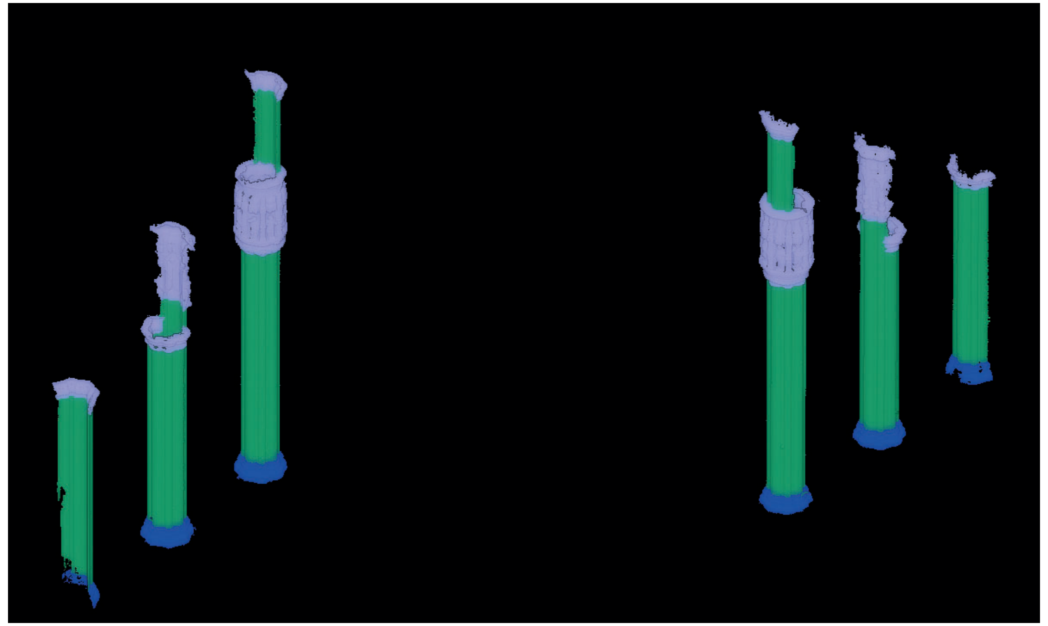

fig.44

A further point on which an operational strategy had to be elaborated concerns the composition of the tall capitals of the central columns. In fact, unlike all the other capitals in the Cathedral, they are not arranged as classical capitals with a base that is orthogonal and continuous with respect to the vertical development of the column, but (as can be seen in the second figure) begin with a jagged segment that divides them from the shaft. This characteristic led to the definition of a standardized approach to solve these problems in the construction of the study models. For the purposes of the research carried out in this work, it was decided to define an abstract limit at which the shaft stops, and the capital begins at the lower limit of its components. By acting in this way, it was possible to generalize the information regarding the characteristics of the capitals. For the classification of the pillars at this level, two sets were assigned composed of the three different types of pillars inside the Cathedral, one on the north side and one on the south side. It was considered that the repetitiveness of these three types of pillars was consistent throughout the interior of the Cathedral and that these patterns could be sufficiently representative for the identification of each class.

n the segmentation of these categories of pillar components, the correct relationship between the resolution of the cloud, the choice of feature radiuses and the realization of the training set and the evaluation set played a key role. The correct preparation and handling of the preliminary data made it possible to obtain an acceptable result right from the start, which took only a few minutes to clean up and was ready for the subsequent work steps. In this particular case, the architectural composition and the vertical scanning mode of the elements also made the operations, both manual and automatic, easier and more effective.

fig.45 Final result of automatic classification of pillars

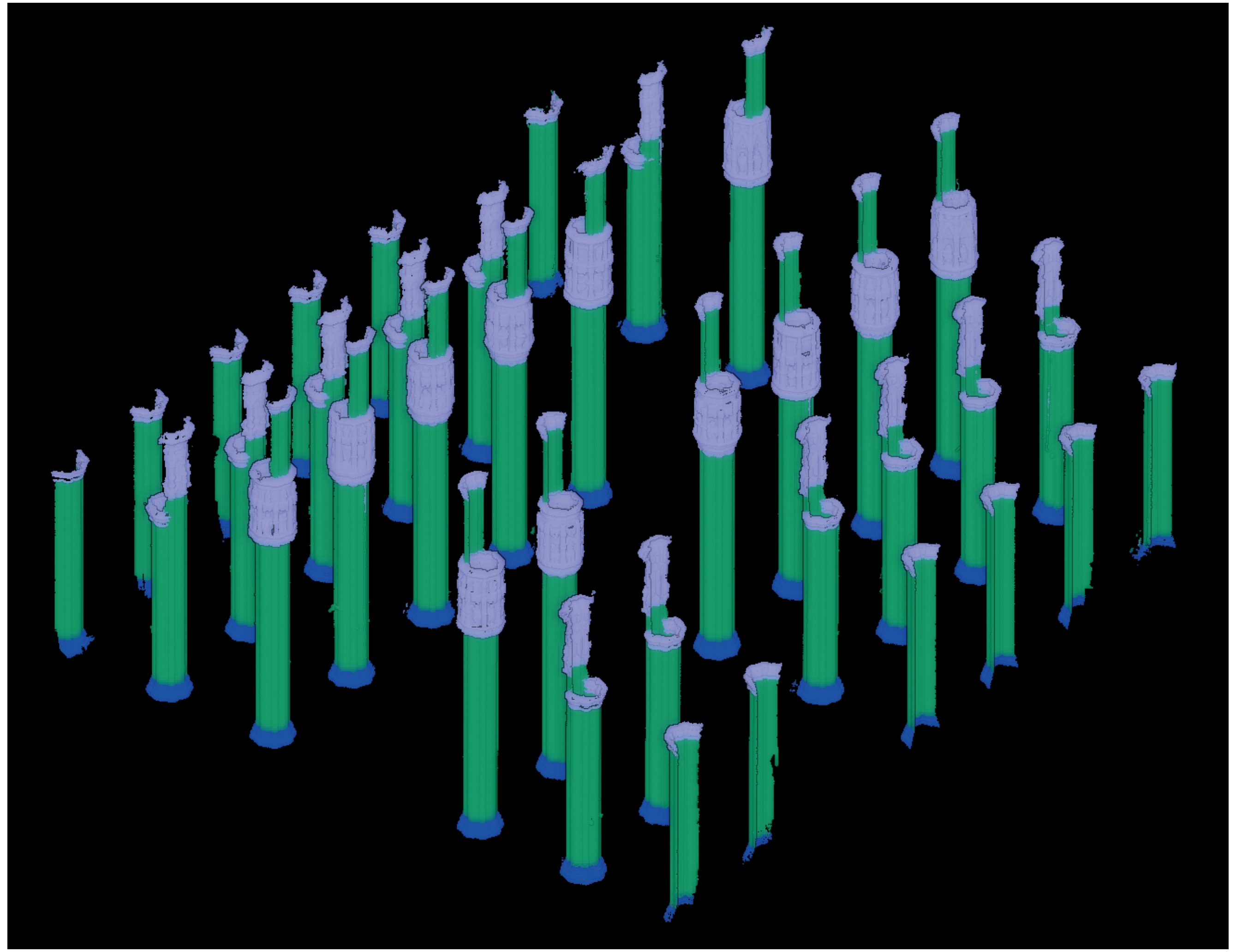

fig.45

Table 7
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	12,532,597
Train set points	1,062,930
Evalu set points	1,025,246

Table 7

Table 8
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	35 min
Time processing	650 sec.
Manual reassignment	10 min

Table 8

Table 9
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Base	0.97	0.96	0.96
Shaft	0.99	0.97	0.98
Capital	0.95	0.97	0.96
Accuracy avg	0.97		
Macro avg	0.97	0.97	0.97
Weighted avg	0.97	0.97	0.97

Table 9

VAULT

The interior vaults of the Cathedral constitute, from an architectural point of view, some of the purest elements we can find. They are located at different heights from the ground depending on the portion of the Cathedral they cover, but this should not affect the automatic classification of its subcategories too much since, being completely separate from each other, they should be recognized and analyzed as distinct entities. The aforementioned purity mainly concerns the form and surfaces that make up the arches, ribs and vaults; these are presented almost as they are, without decoration, simplifying their recognizability both during the creation of the study models by the operator and in the automatic classification phase by the algorithm. Furthermore, the similarity of the vaults within the cathedral, both in the nave and the transept, makes it possible to work using the same study models to generalize as much as possible the classification of all internal data regarding these elements.

For the second classification level of the vaults, given the proportions of the architectural elements and the elements of detail that can be compared with those of the pillars, the same basic principles seen in the previous case were used: point cloud with a density of 20 mm and the same radiuses to investigate the geometric characteristics. The main peculiarity of the second level of classification of these elements is the composition and development of the different parts that compose the vault. Until now, in fact, classification has mainly concerned elements with vertical development, such as pillars, where the 'z' coordinate has played a fundamental role in the algorithm's recognition of the different element categories. The classification of vaults, with the development of its components, leads for the first time to the consideration of curved elements for which the previous approach might not be sufficient. In particular, it is useful to highlight the substantial difference between arches and ribs which, from an architectural point of view, may be assimilated as structural elements with a linear development, but that, considered as a whole, in the same model to be classified, show two different trends. The arches, in fact, assuming a reference system x and y parallel to a vault seen in plan, develop along the x and y axes in a linear manner, while the ribs of the same portion of the vault assume an inclined course with respect to these axes. One of the most discussed points at this level of classification concerned the keystones. At first, they were intended for a fourth

category but, due to the problems that will be explained below, it was decided to reduce the number of categories and assign the keystones points to one of the three existing categories.
From the following image (left), with the cloud set at 0.05 m verticality, one can observe the above-mentioned problem related to the z-coordinate. The increase in verticality can be recognized by the increase in the hue of the cold colors, from which one can appreciate the fact that all elements assume similar behavior as a function of this variable.
For the definition of the study models to be manually classified, a single vault span was selected between two pillars, which was in turn divided in half to separate the training and evaluation sets.

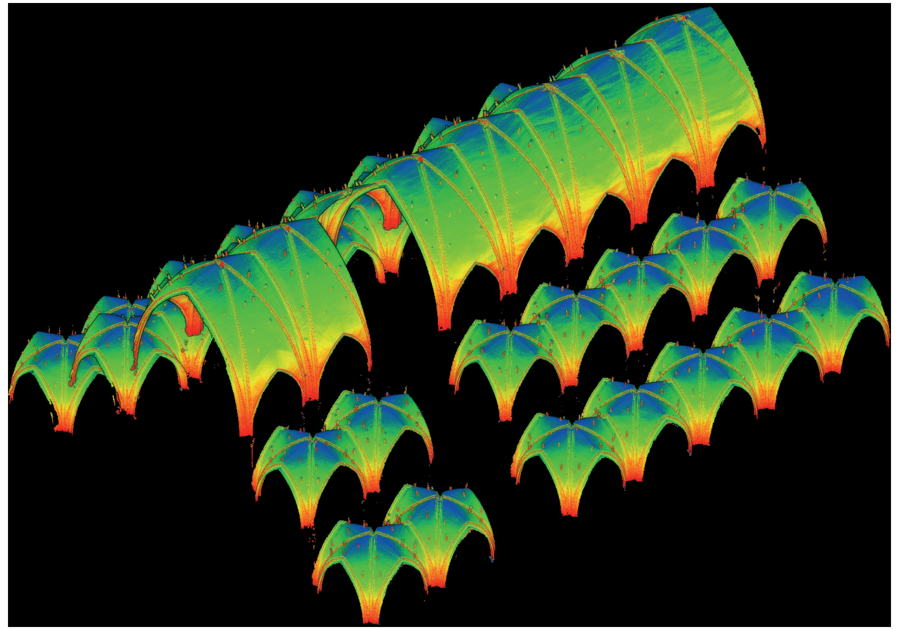
fig.46

fig.46 Vault dataset with feature verticality 0.05m


fig.47

fig.47 Training set ed Evaluation set for second leveò of classification of vault

The definition of the initial semantic categories involved the identification of arches, ribs, sails and keystones. It was assumed that the latter could be identified through the contribution of the z-characteristic, given their position at the top of each vault. The main problem with this strategy lies in the fact that the keystones are located at the point of intersection of ribs and sails, but do not develop in height in relation to them, but rather below them. This composition of the elements was in itself, at first, a limitation in the process of recognition of the elements by the operator in the study models. This condition is caused by the instrumental limitation and by the resolution of the data used for this level of classification. The first is related to the fact that these clouds were detected by positioning the instrument below the vaults in order to beat the points of greatest interest; the second relates to the subsampling of the original data, which reduces the amount of points related to the connection between the ribs and the keystones.
These circumstances showed limitations in the classification of these four categories from the first phase of the creation and evaluation of the training set and were then reflected in the automatic classification of the algorithm. It must be said that the problem, in this case, does not concern the general quality of the automatic classification (as visible in the following images), but rather the architectural composition of these elements, which makes them difficult to identify both automatically and manually. Considering the level of resolution of the cloud, the processing time becomes very long and comparable to the time that manual classification would require, going against the purpose of this work. The main limitation of this phase lies in the fact that, as mentioned at the beginning, the laser scanner technique only surveys the surfaces of objects, without any information on the overall structural composition of the elements. These are therefore situations in which, for this type of work and with this surveying instrumentation, there are no viable solutions.

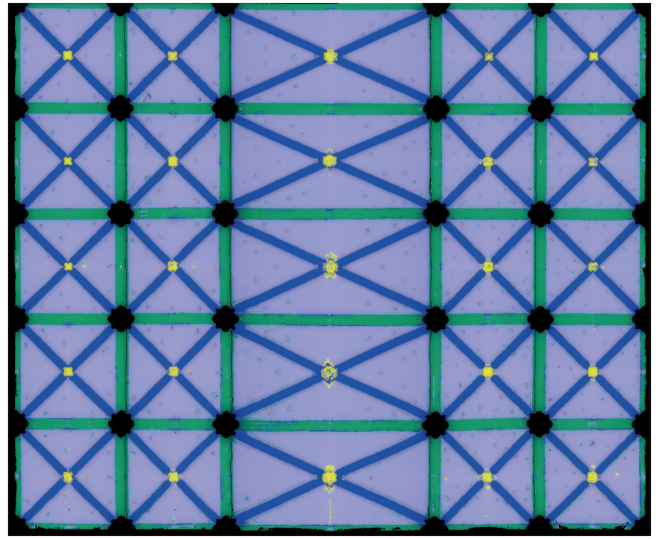
fig.48

fig.48 Plan of the result of automatic vault calssification with key's classes

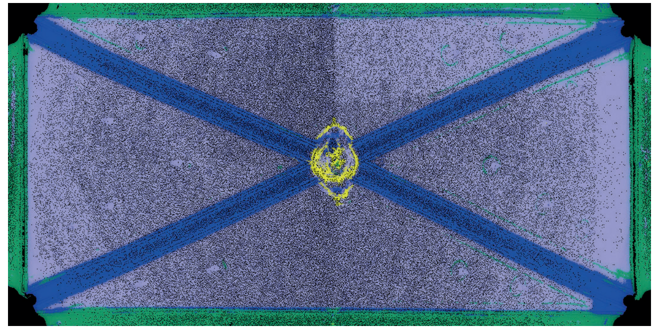
fig.49

fig.49 Detail of the result of automatic vault calssification with key's classes fig.50 Frontal detail of arches and ribs detail with features Surface variation 0.05m

fig.51 Plan detail of arches and ribs detail with features Surface variation 0.05m

A further problem encountered is the similarity between the edges of the arches and ribs; these are almost 'pure' elements, except for the only decoration we can identify, which are the moldings. Arches and ribs, in fact, at the point where they join the vaulting sails, have regular moldings that define their limits. The main disturbance of this element is that it is present on both elements in equal proportions, creating difficulties for the code in deciding which category to assign to the points that constitute these decorations. In the following images, the cloud has been set to the 'Surface Variation' feature with a perceptual field of 0.05 m, in order to capture the geometric similarity of the ends of the arches and ribs. Although these elements are semantically and structurally distinct, it must be underlined that this type of data does not have the capacity to contain semantic or structural information to identify such distinctions.

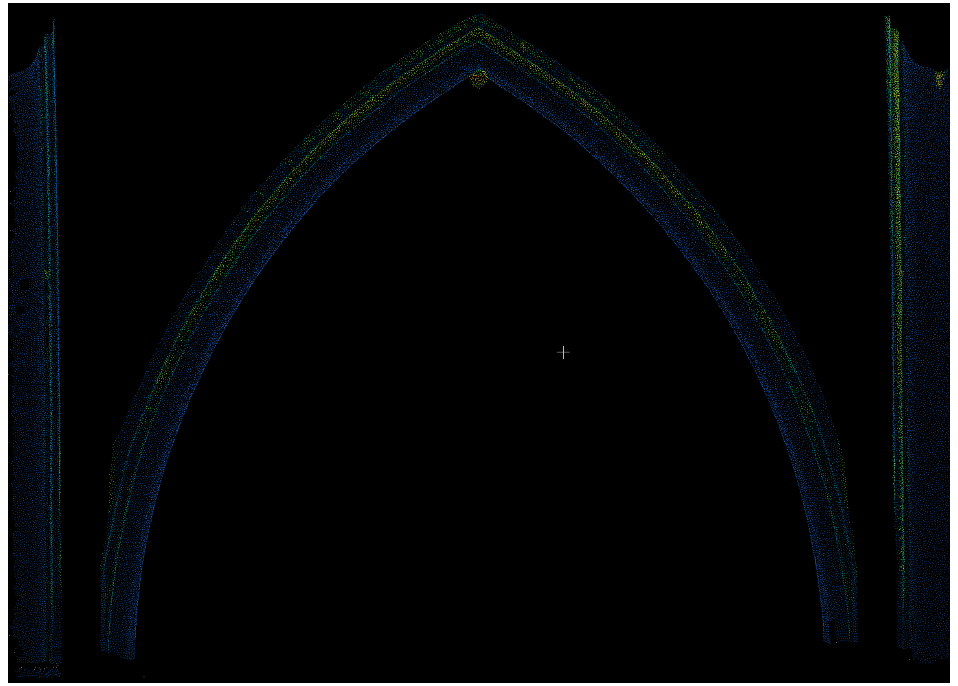

fig.50

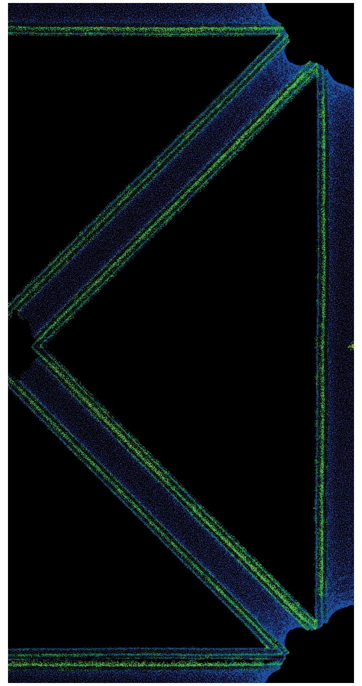

fig.51

Regarding the difficulty in classifying the vaulting keys, after a few attempts, it was realized that no matter how much care was taken in creating the study models to define the correct points belonging to ribs and vaulting keys, during the automatic segmentation phase the algorithm was unable to identify and separate them correctly. Changing the survey radiuses or removing certain features also did not lead to a significant improvement but did not compromise the identification of the remaining classes. In view of the previous considerations, the keystones were incorporated into the ribs category, considering them as direct extensions of these elements, since from a structural point of view they are also closely connected and can be considered as the last indivisible compositional element. The result of this was an improvement in the performance of the automatic classification and, consequently, a reduction in the time required for manual cleaning and reassignment of the points to the correct categories. The eventual subdivision of the keystones from the ribs could, however, later be taken from a more specific datum, as it lacked arches and sails.
This would result in a cloud with higher resolution and the consequent computation of new perceptual fields better suited to the scale. In this way, it would be easier to distinguish keystones from ribs, both in the modelling phase and in the data processing phase of the code.

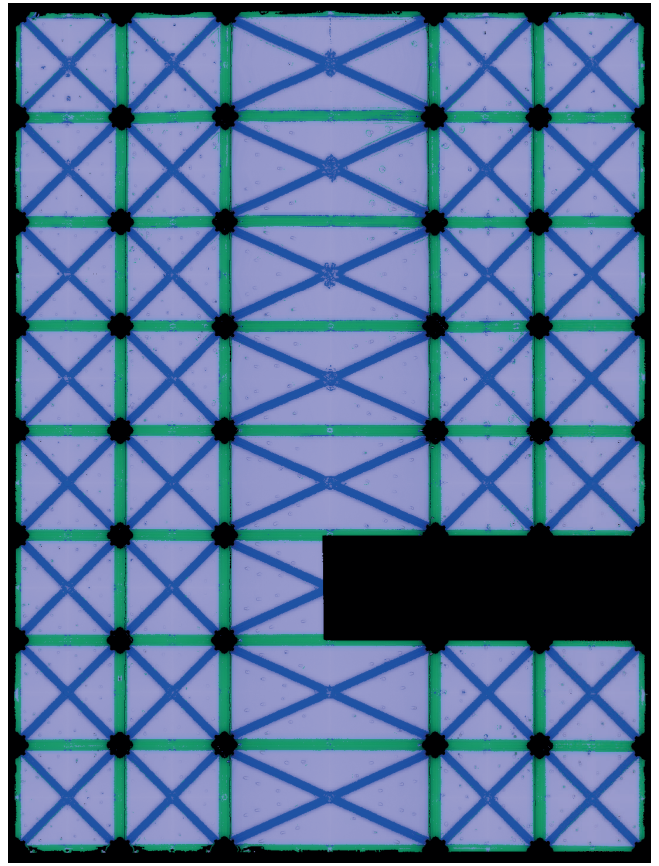

fig.52

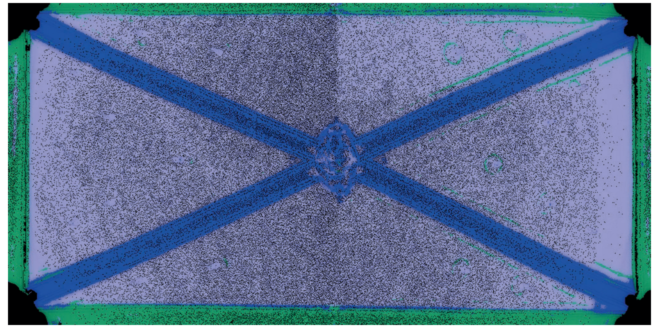

fig.53

fig.52 Plan of the result of automatic vault calssification without key's classes

fig.53 Detail of the result of automatic vault calssification without key's classes fig.54 Final result of automatic classification of vaults

From the following image one can appreciate the good quality of the result of the automatic classification by the algorithm, in which only two problems persist: that of the moldings of arches and ribs and that of certain points of the keystones classified as sails instead of ribs. With regard to the first, as already mentioned, it was decided to resolve it during the manual cleaning. As for the second, the situation was simplified in view of the cleaning operations, since the main problem was distinguishing keystone points from ribs. Now that these two classes have been unified, despite the persistence of a similar number of incorrectly classified points, manual operations will simply reassign the keystone points classified as sails to the correct category of ribs.

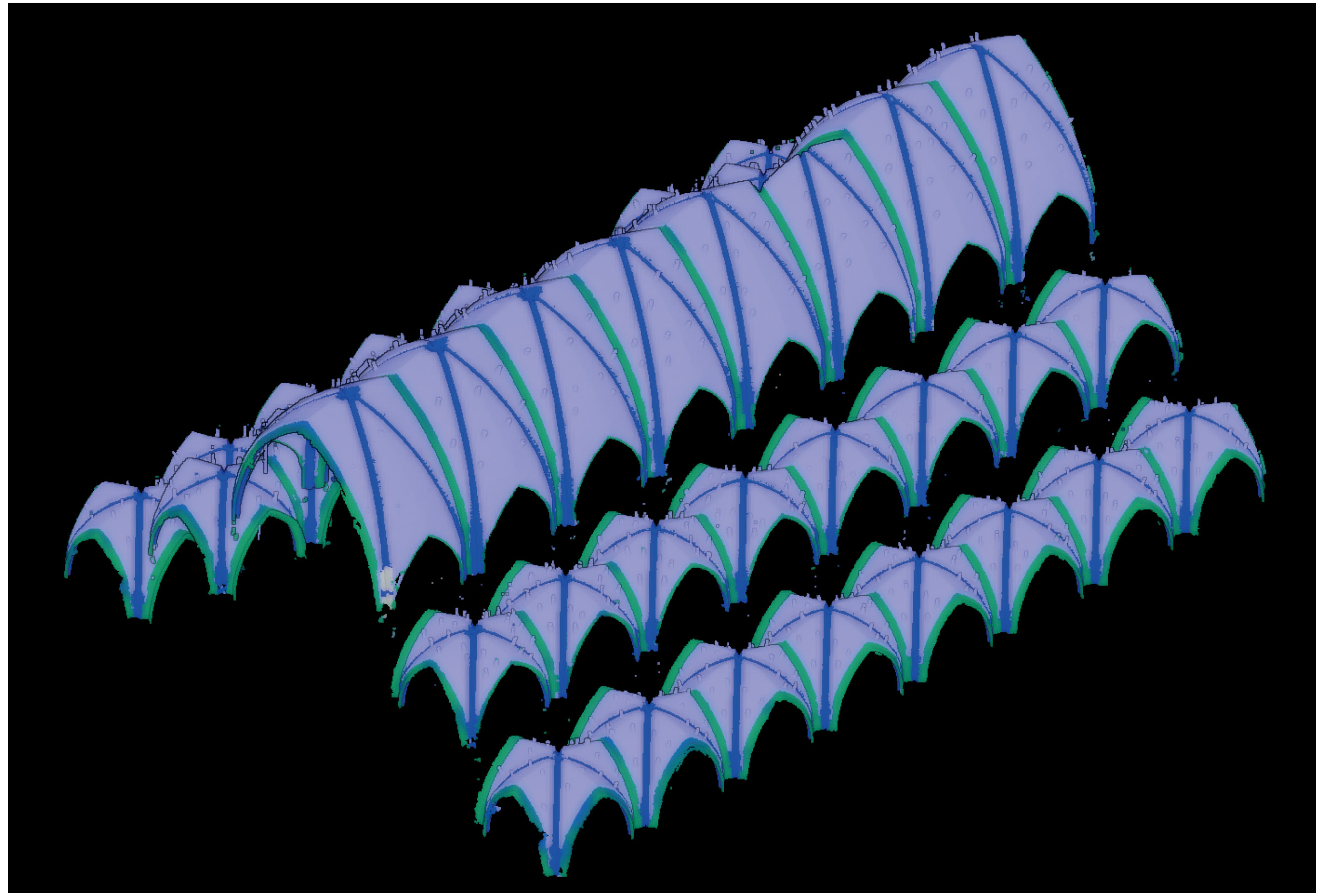

fig.54

fig.55 Rib class isolated from the final result

fig.56 Arch class isolated from the final result

fig.57 Sail class isolated from the final result

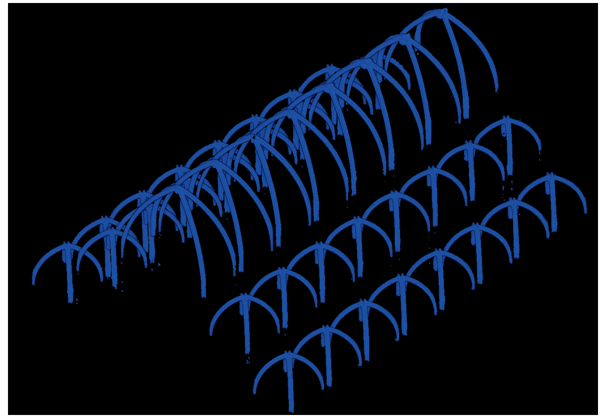

fig.55

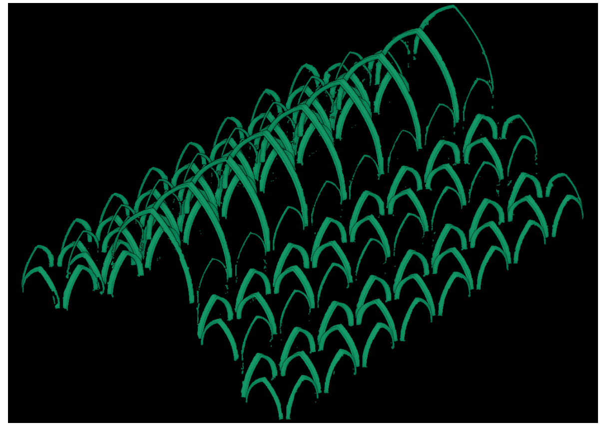

fig.56

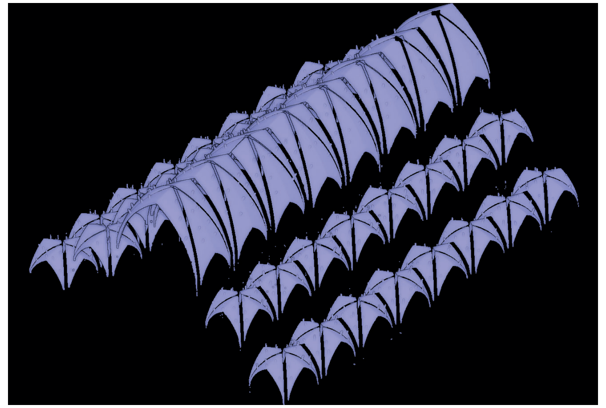

fig.57

Table 10
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	15,235,708
Train set points	512,290
Evalu set points	509,502

Table 10

Table 11
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	20 min
Time processing	900 sec.
Manual reassignment	120 min

Table 11

Table 12
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Rib	0.97	0.98	0.98
Arch	0.97	0.97	0.97
Sail	0.99	0.98	0.99
Accuracy avg	0.98		
Macro avg	0.98	0.98	0.98
Weighted avg	0.98	0.98	0.98

Table 12

GENERALIZATION FOR INTERNAL CLASSIFICATION

PILLAR CLASSIFICATION LEVEL 2

For the second-level classification of the transept elements, an attempt was made to maximize the concept of generalization set out above. With regard to pillars and vaults, given the repetitiveness of the shapes and composition of these elements located in the transept, an attempt was made to use study models made manually in the main nave for their classification.
With regard to the pillars, the repetition of the same types of bases, shafts and capitals is evident. The only aspect that varies is the general arrangement in space, but since each column is at least 5 metres apart, these are relationships outside the perceptual fields used for feature analysis, i.e. outside what the code uses for learning. In this case, the classification provided encouraging results, producing automatically segmented data with good accuracy. The only cases in which the code encountered problems in assigning classes to points were at some of the lights of the lighting system present only on some of the pillars of the transept. Since these are non-architectural elements, they did not affect the overall assessment of the result.
A final consideration concerns the peducci, vertical elements continuing the four corner columns on which the pendentives supporting the dome of the Cathedral rest. As these are peculiar elements within the Cathedral, it was decided to disregard them and assign them to a new category, as they are not part of the canonical elements that make up a pillar. The classification of the architectural parts that make up the pillars was seen as a success considering that the source study model, from which the information for the code was extracted, belonged to another area of the cathedral.

fig.58 Detail of pillar classification with "peducci"

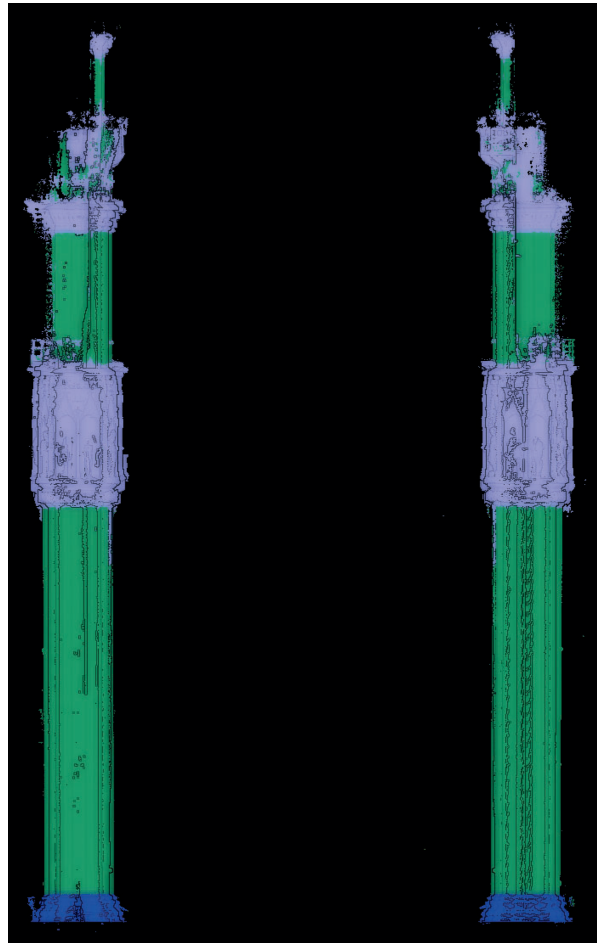

fig.58

fig.59 Result of the automatic classification of the transept pilalrs by generalising the study model of the nave

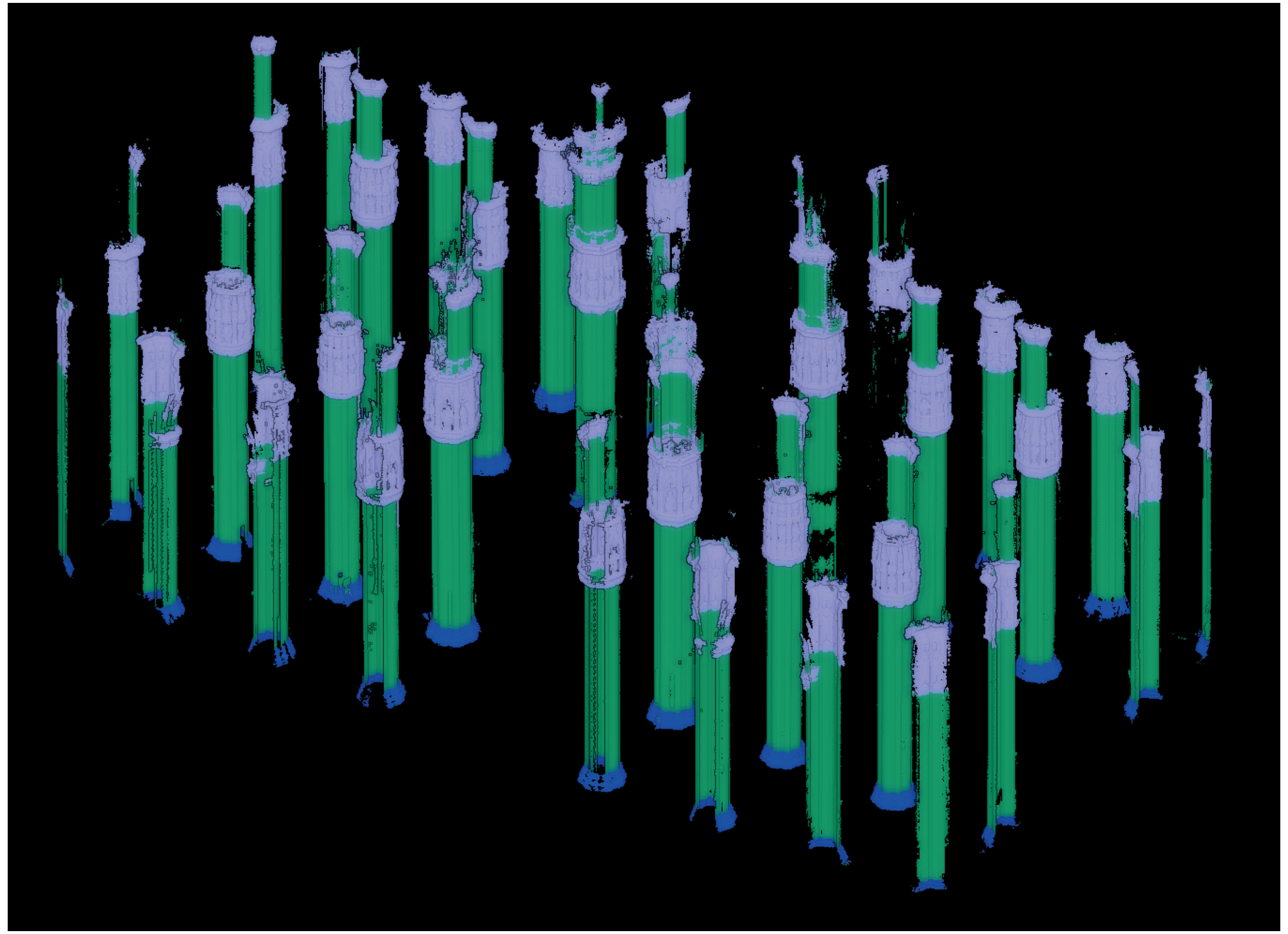

fig.59

Table 1
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	11,962,193
Train set points	0*
Evalu set points	0*

Table 13

Table 2
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	0* min
Time processing	450 sec.
Manual reassignment	45 min

Table 14

Table 3
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Base	0.96	0.96	0.96
Shaft	0.98	0.97	0.98
Capital	0.95	0.97	0.96
Accuracy avg	0.97		
Macro avg	0.96	0.97	0.97
Weighted avg	0.97	0.97	0.97

Table 15

* Used training and evaluation set realized for pillar classification in the central nave

fig.60 Result of the automatic classification of the transept vaults by generalising the study model of the nave

VAULT CLASSIFICATION LEVEL 2

For the interior vaults of the transept, the same approach was applied as for the pillars: provide the code with the training set created for the nave from which it learns the information for classification. At the compositional level, no particular differences are noticeable between the vaults of the nave and those of the transept, in both spaces they follow the compositional system formed by arches, ribs (joined at the meeting point by a keystone) and sails.
The only peculiar and non-repeatable elements are: the vaults of the two side chapels located in the center of the north and south walls of the transept, the four pendentives resting on the peducci and supporting the dome and, finally, the central lantern at the top of the dome itself. The first two listed, given their different structural purpose and compositional structure within the organism of the vaults, were not taken into consideration for the evaluation of the final result, as they are unique elements both in form and composition. As far as the lantern is concerned, given its structural function similar to a normal keystone and connection between the eight ribs that make up the dome, it was decided to keep it within the rib category.

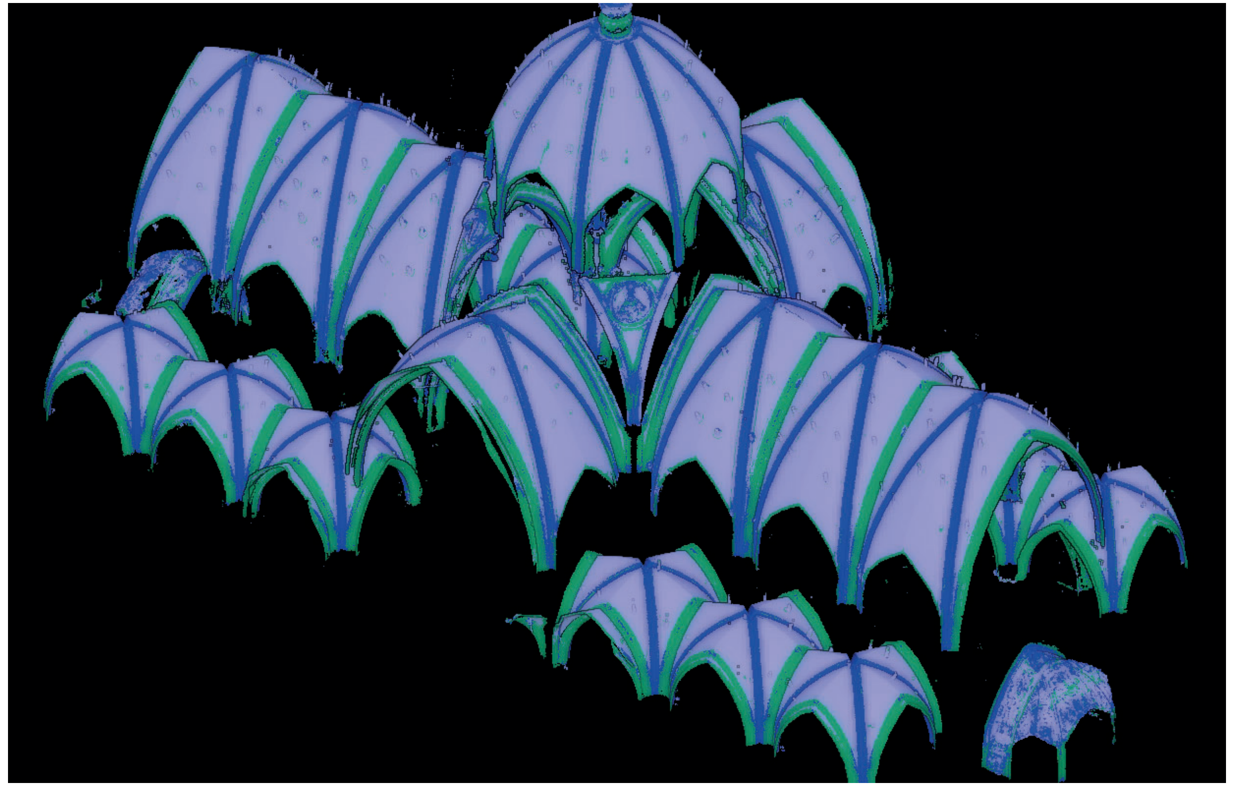

fig.60

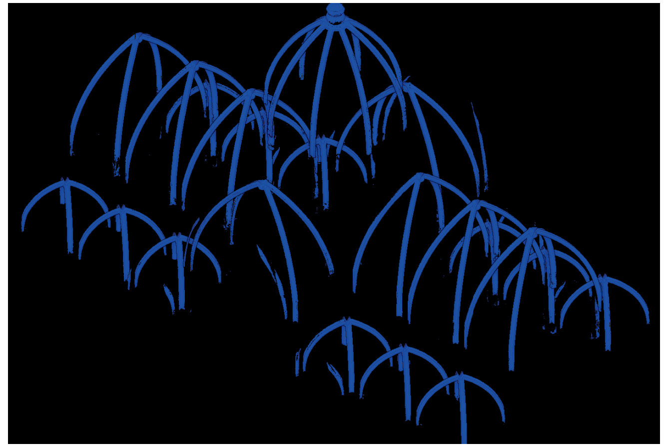

fig.61

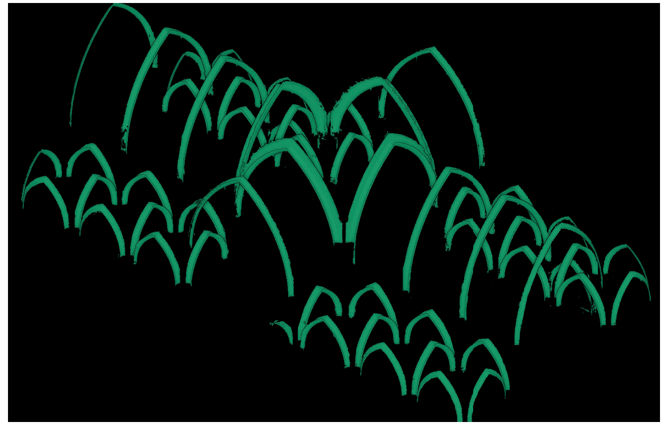

fig.62

fig.63

fig.61 Rib class isolated from the final result

fig.62 Arch class isolated from the final result

fig.63 Sail class isolated from the final result

Table 16
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	13,267,056
Train set points	0*
Evalu set points	0*

Table 16

Table 17
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	0* min
Time processing	850 sec.
Manual reassignment	150 min

Table 17

Table 18
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Rib	0.97	0.98	0.98
Arch	0.98	0.97	0.97
Sail	0.99	0.98	0.99
Accuracy avg	0.98		
Macro avg	0.98	0.98	0.98
Weighted avg	0.98	0.98	0.98

Table 18

* Used training and evaluation set realized for vault classification in the central nave

CLASSIFICATION OF EXTERIOR

fig.64 Scheme of the pipeline for external classification

Exterior Classification Pipeline

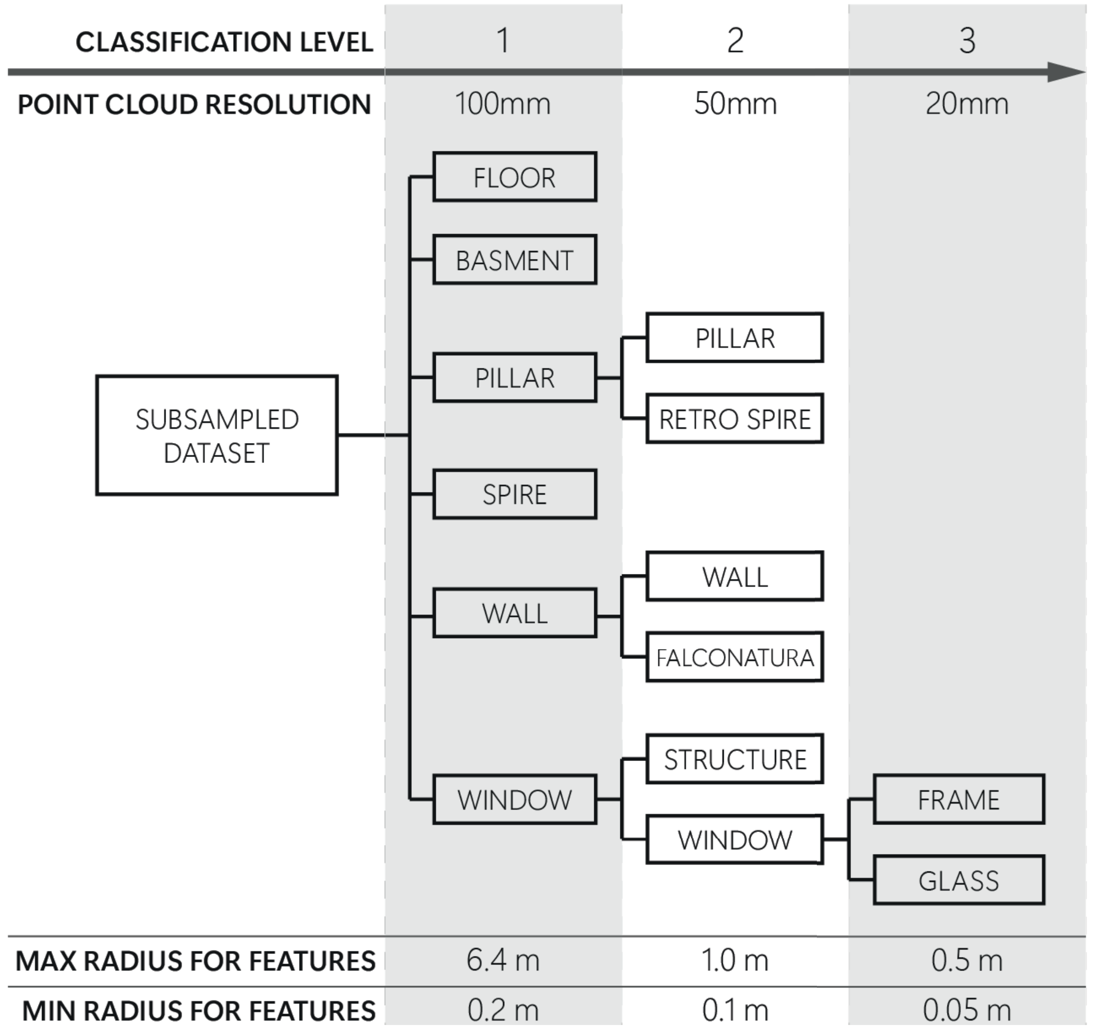

fig.64

fig.65 50mm resolution external cloud with RGB information

The starting data provided for the classification of the exteriors consisted, as for the interiors, of a cloud with a resolution of 5 mm, but unlike the former, this was implemented with RGB information derived from interpolation with photogrammetric data. This additional information makes it possible to create a point cloud with real color data of the revealed surfaces; the only limitation of this tool is the color difference between the areas illuminated by direct sunlight and the ones in shadow.
The possibility of using this information improves both manual processing by the operator and automatic processing by the code. In the first case, due to the presence of the actual colors of the elements, it is easier for those who know the semantics of the elements to identify and divide them. In the second case, the RGB data is added to the list of information that the code uses to learn and subsequently classify the individual points. From the following images, one can appreciate the visual clarity of windows compared to walls provided by the color information of the points.

fig.65

fig.66

fig.66 Detail of exteriors cloud at 50mm resolution with RGB information fig.67 Inizial training set and evaluation set of classification level 1 of exterior

fig.68 Final training set and evaluation set of classification level 1 of exterior

CLASSIFICATION LEVEL 1

For the first exterior classification level, the same approach was used as for the corresponding interior level, considering similar the proportions of the elements to be classified: Floor, Basement, Pillars, Walls and Windows. In fact, the source cloud at the 5 mm definition was subsampled to create a cloud with 50 mm resolution on which features were calculated for perceptual intervals from 0.1 m to 1 m. The decision to limit the classification to only four macro-elements arose from the complexity and quantity of decorative elements present on the external surfaces. One of the main obstacles that affected operations at this stage was the incomplete external survey data. In fact, for the reasons stated above, the survey of the north transept was not available. The complete absence of one of the two transepts caused a significant gap for the classification of the exteriors, since certain elements, such as the corner pillars or the protuberances that internally correspond to the altars in the center of the transepts, are exclusive to these portions of the cathedrals and are not repeated in any other area. At first, an attempt was made to minimize manual classification even for quite unique portions (such as the transept) in order to test, even in non-optimal conditions, the possibility of generalizing the data provided for automatic segmentation. As the first image below shows, initially a portion of the walls of the nave and a relatively small portion of the southern transept were used as training set and evaluation set. Not obtaining acceptable results in terms of accuracy, the portions included in the training set were expanded, resulting in the manual classification of half of the transept and of a good portion of the apse.

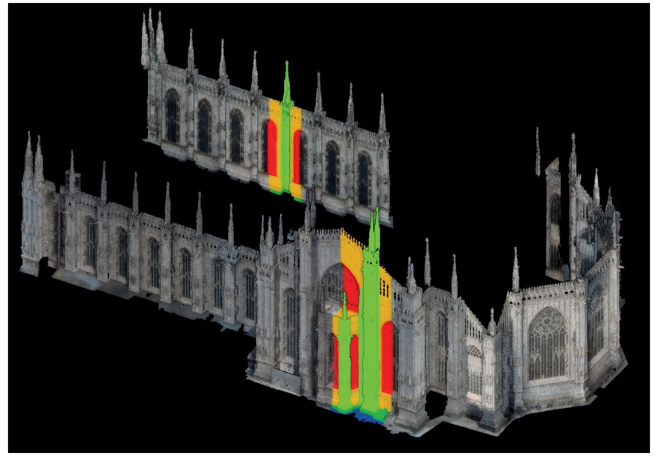

fig.67

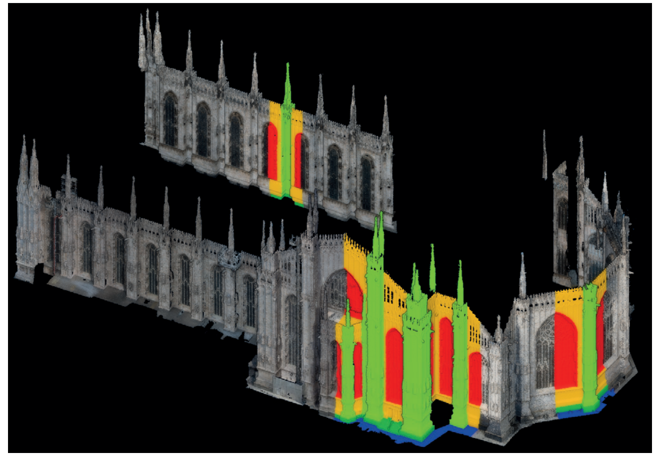

fig.68

Despite the attempt to provide more information to the code through the study models, the results obtained with this data resolution were not considered satisfactory. As can be appreciated from the image below, the macro-categories were correctly identified allowing them to be distinguished from one another, but they presented a large amount of noise due to the incorrect segmentation of the decorative elements on most of the surfaces of the main elements to be classified. In particular, from the second image below, it is possible to appreciate that the main problem is related to the confusion caused by the decorative elements, such as frames and statues, present on the surfaces of the walls and along their boundaries with pillars and windows.

fig.69 Result of automatic classification with 50mm point cloud resolution

fig.70 Detail of the result of automatic classification with 50mm point cloud resolution

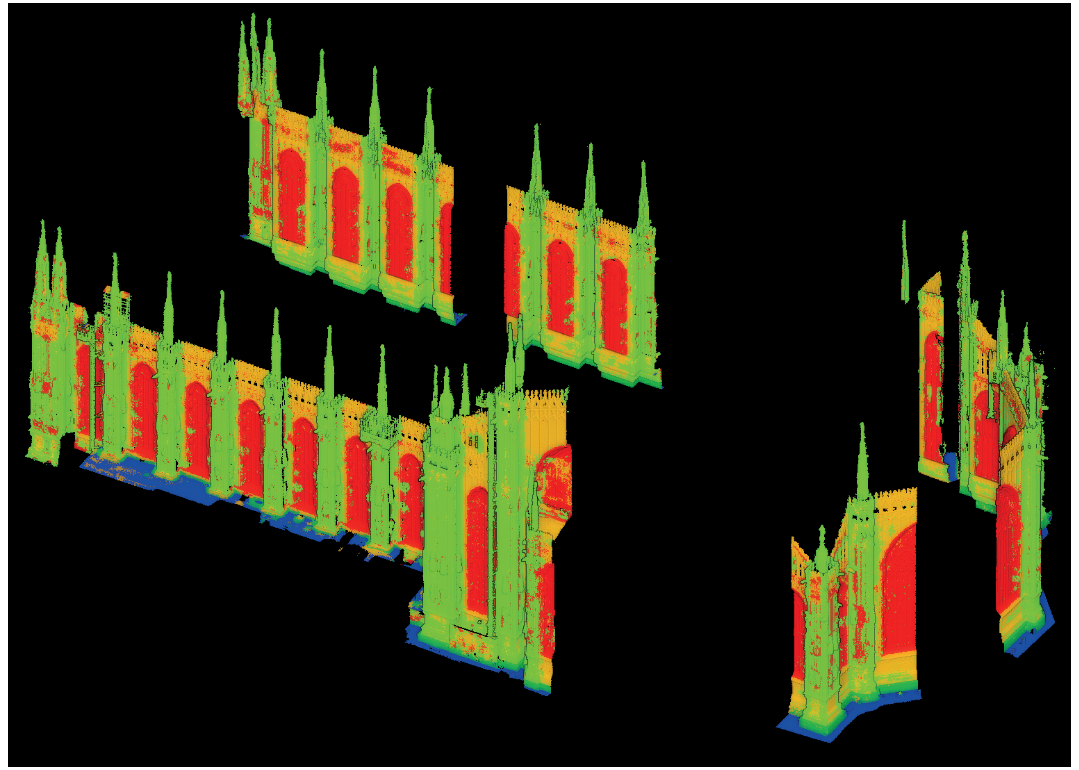

fig.69

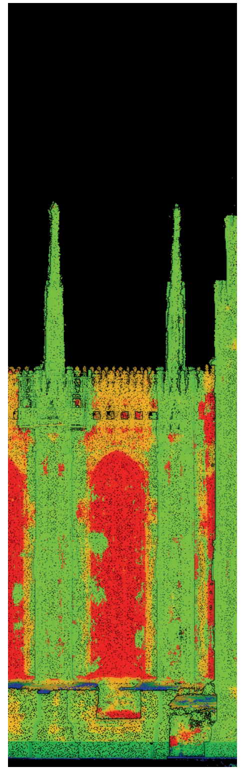

fig.70

Table 19
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	4,618,141
Train set points	79,197
Evalu set points	16,221

Table 19

Table 20
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	45 min
Time processing	240 sec.
Manual reassignment	Undefine

Table 20

Table 21
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Floor	1.00	0.97	0.99
Basment	0.97	0.98	0.97
Pillar	0.87	0.95	0.91
Wall	0.91	0.76	0.83
Window	0.94	0.94	0.94
Accuracy avg	0.96		
Macro avg	0.95	0.92	0.93
Weighted avg	0.96	0.96	0.96

Table 21

In light of the results obtained, it became evident that the main cause of the classification errors was the incorrect resolution set for the starting cloud. In this case, in fact, with a cloud resolution of 50 mm and the presence of decorative elements of significant size, we obtain a cloud in which the shapes of these decorations are defined by too many points for them to be ignored by the code during the learning and classification phase. As can be seen from the following images, with this cloud resolution the decorative elements, such as the statues along the window splay or the frames on the wall surface, are represented with too many points that define their boundaries too clearly. From a practical point of view, we can say that depending on the resolution of the cloud, the code 'sees' the elements as our eyes see them: if the cloud we observe has enough resolution to allow us to recognize the shapes of certain elements, then the code will have enough points on which to investigate geometric relationships to recognize the same elements. We exploit sight while the code exploits the investigation of geometric relationships between points, so we use different tools to perceive the same things.

fig.71 Section of 50mm resolution point cloud that frame a window between two external pillars

fig.72 Frontal view of 50mm point cloud resolution

fig.71

fig.72

fig.73 Result of automatic classification with 100mm point cloud resolution

fig.74 Detail of the result of automatic classification with 100mm point cloud resolution

Once the initial resolution error of the cloud was recognized as the main problem, we further subsampled the cloud to a resolution of 100 mm. This was considered the most suitable solution to minimize the errors caused by the decorations in the first classification level. We moved from a cloud of 6,244,137 points to one of 1,243,909 points. This sub-sampling made it possible to obtain a cloud with a smaller number of points, in which it was no longer possible to clearly identify the shapes of the detail elements that previously had a negative influence on the classification. In addition to that, this subsampling allowed a considerable reduction in the size of the data and the resulting processing times, both manual and automatic. Comparing the subsequent images with the previous ones, it is possible to notice the lower resolution of the cloud and, consequently, the lower definition of the detail elements, both in section and elevation.
These sub-sampling operations, as well as improving the accuracy of the automatic classification by the code, allowed to increase the number of classes to be segmented directly at the first level. At this stage, some tools provided by Cloud Compare made it possible to use the manually segmented study models on the 50 mm cloud as a source to transfer the point class information to the 100 mm resolution cloud without having to perform the manual training set creation and evaluation operations again. The "Guglia" category was added to the classes already included in the previous attempt. From the very first tries with this new data handling, the results have been encouraging and have confirmed the importance of choosing the correct data resolution in relation not only to

fig.73

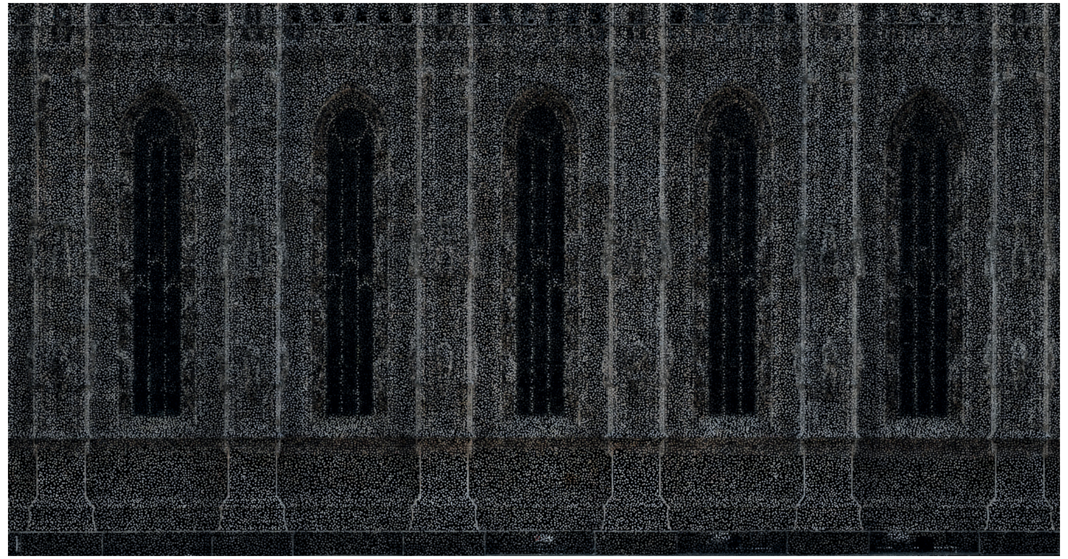

fig.74

the proportion of macro-elements to be classified, but also to the minor disturbing elements that are to be made negligible by the code. Comparing the following images (result of the classification with the 100-mm cloud) with those shown previously (result of the classification with the 50-mm cloud), it is easy to notice the improvements obtained with the new data. Noise due to decorative elements has been almost completely eliminated, with the exception of small portions at the intersection of walls and pillars. The result obtained is an excellent starting point for the subsequent manual cleaning and reassignment of the incorrectly classified points. The problems and related solutions adopted in this phase of the classification work brought to attention a fundamental value within the MLMR approach, as they highlighted the fact that more information does not always correspond to a better quality of the final result. Indeed, in this specific case, it was necessary to take a step back from what had been planned, as the complex uniqueness of the decorations on the cathedral's exterior walls constituted an insuperable obstacle by keeping the cloud resolution at 50 mm. It was only through a change of scale and thus a 'loss' of information that the correct execution of this level of classification was achieved.

fig.69 Result of automatic classification with 100mm point cloud resolution

fig.70 Detail of the result of automatic classification with 100mm point cloud resolution

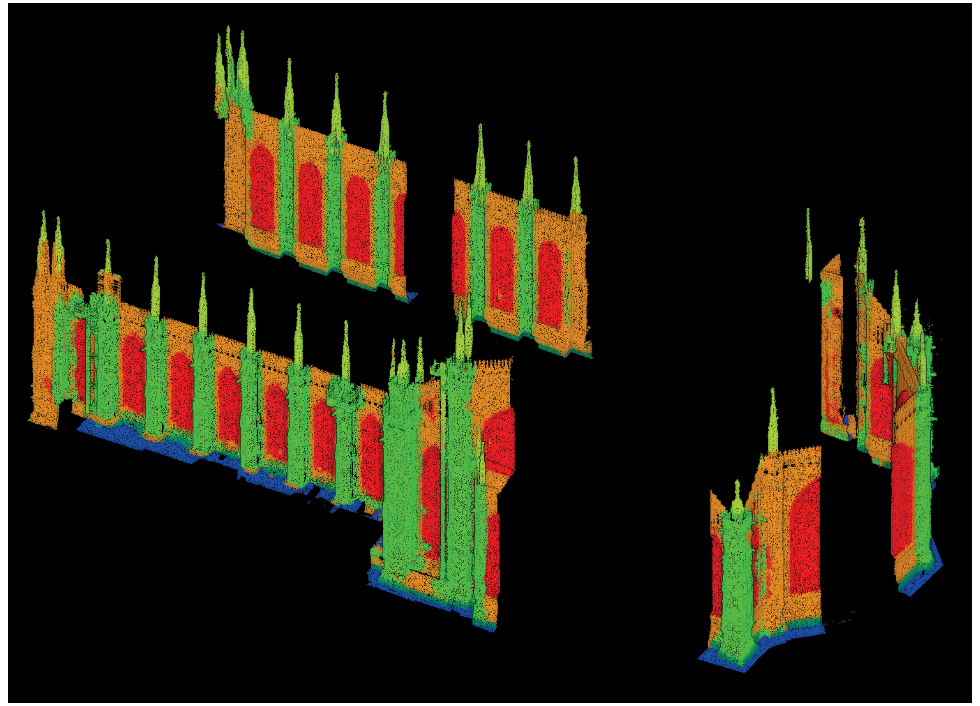

fig.75

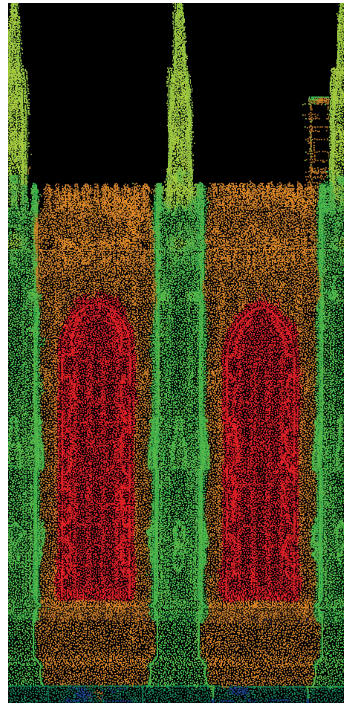

fig.76

Table 22
Summry of 3D point cloud number of points

Table 23
Summary of time including manual and automatic operation

Table 24
Accuracy of automatic classification

	N. POINTS
Points to clasify	926,216
Train set points	0*
Evalu set points	0*

Table 22

OPERATION	TIME
Manual annotation	0* min
Time processing	120 sec.
Manual reassignment	20 min

Table 23

	PRECISION	RECALL	F1-SCORE
Floor	1.00	0.96	0.98
Basment	0.93	0.95	0.94
Pillar	0.99	1.00	0.99
Guglia	1.00	1.00	1.00
Wall	0.99	0.96	0.97
Window	0.96	1.00	0.98
Accuracy avg	0.98		
Macro avg	0.98	0.98	0.98
Weighted avg	0.98	0.98	0.98

Table 24

* Used training and evaluation set realized for exterior classification with 50mm point cloud resolution

CLASSIFICATION LEVEL 2

This level of classification requires that from the classes of walls and pillars are segmented respectively the sub-categories of “falconature” and retro spires. The former constitutes the last order of decoration of the exterior walls and also act as parapets, while the latter constitute the continuation of the former and are elements that cover the base of the spires. In both cases, these subcategories are physically represented as crowning bands that follow the development of walls and pillars. The realization of the study models revealed a remarkable simplicity in the manual segmentation of the categories at this level. Consideration was given to the fact that the time required for code processing and subsequent cleaning of the obtained data would have been longer than the time required to manually classify the entire data. For this reason, it was decided not to pursue the investigation of the classes.

fig.77 Training set and evaluation set to segment pillar and retro spire

fig.78 Training set amd evaluation set to segment wall and falconature

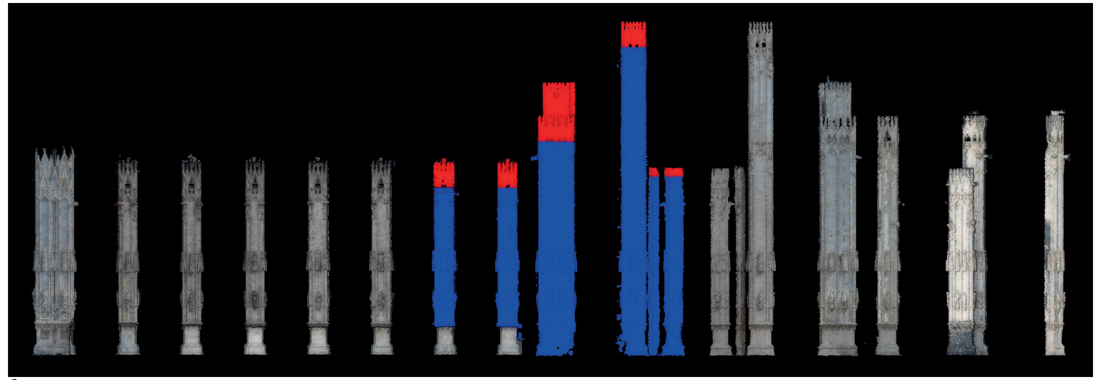

fig.77

fig.78

Table 25
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	1,920,203
Train set points	227,893
Evalu set points	227,458

Table 25

Table 26
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	5 min
Time processing	240 sec.
Manual reassignment	5 min

Table 26

fig.79 Result of automatic classification of pillar and retro spire

fig.80 Result of automatic classification of wall and falconature

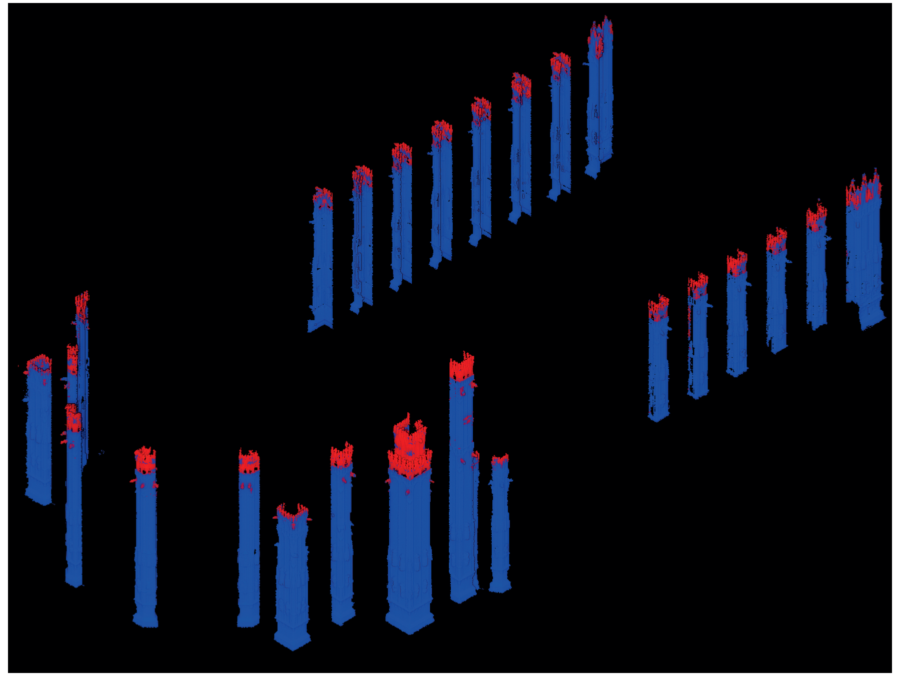

fig.79

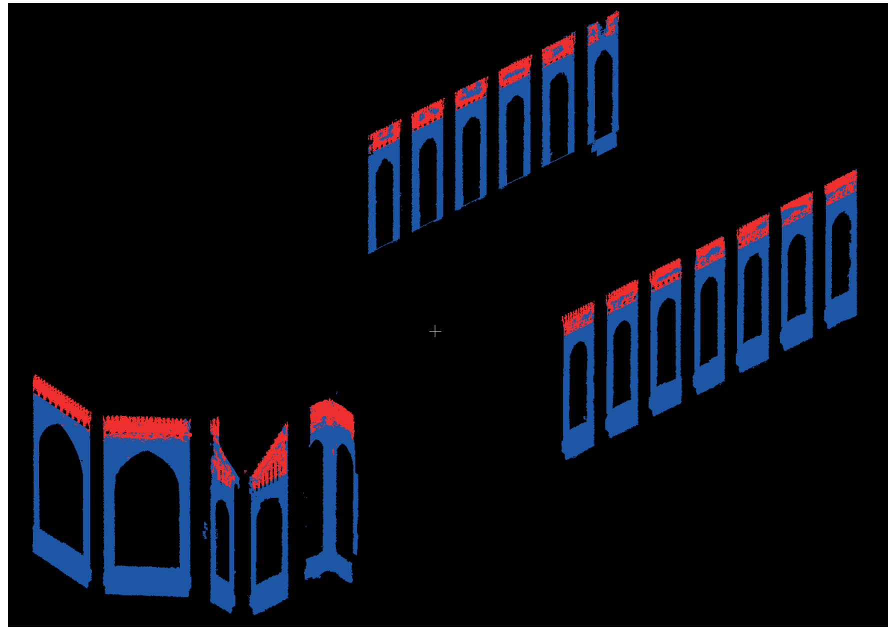

fig.80

Table 27
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	857,804
Train set points	297,166
Evalu set points	376,703

Table 27

Table 28
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	5 min
Time processing	240 sec.
Manual reassignment	5 min

Table 28

CLASSIFICATION LEVEL 3

WINDOWS

For the third and final level of window classification, it was required to identify and segment the frame (represented by the mullions and marble decorations) and the glass, the two ultimate and indivisible elements that make up the windows. From the previous classification level, a point cloud was obtained that includes all the exterior openings of the Cathedral. From an initial analysis, it is clear that there are very heterogeneous shapes and compositions. In fact, we have different types of windows depending on their position along the exterior façade of the Cathedral, one type along the main nave, different types in the transept and, finally, the three large openings located in the apse.
In contrast to the interiors, for the third classification level of the exteriors, the elements to be identified were larger in size and, as a logical consequence, the starting point cloud was set with a density of 20 mm instead of 5 mm. The aim was to best isolate these three features using the right combination of feature detection radiuses.

fig.81 Training set and evaluation set to segment structure, frame and glass

fig.82 Detail of the definition of training set amd evaluation set to segment structure, frame and glass

fig.81

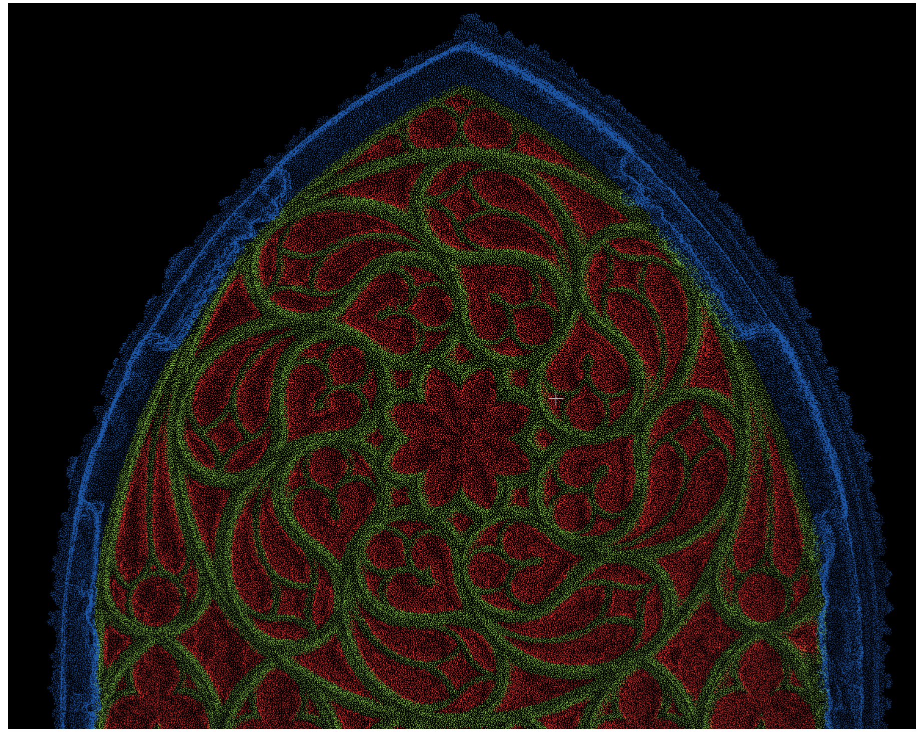

fig.82

fig.83 Detail of classification errors inside the window

The main difficulty lies in clearly distinguishing the outer cornice (splay) from the inner window due to the presence of large statues along the cornice. As can be seen from the detail above (fig. n), part of the points of the outer cornice and the window frame overlap each other. The first analysis arising from this problem led to the hypothesis that the algorithm had difficulties in identifying the cornice as a continuous entity due to the presence of the statues, which significantly altered the classical shape of the outer cornice. In cases like this, where elementary architectural elements are enriched and, consequently, partially covered by numerous decorations, it is difficult to identify architectural forms. Due to their uniqueness in terms of geometric forms, decorations, whether statues or ornaments, most often assume the role of obstacles to the full understanding and identification of the elementary forms of architectural elements.

By changing the radius of investigation of the geometric relationships between points, the extent of the geometric characteristics that each individual point has in relation to the others contained within that radius is changed. As a result, the data provided to the code for the automatic classification of elements also increases. This operation therefore acts directly on the basic information contained in the model, without the need for manual modifications to the cloud.

Given this potential, the development of this method of data modification and management is fundamental, as it solves the problem of the simultaneous presence of decorations of very different sizes that hinder the identification of architectural classes.

O n a practical level, increasing the radius of the relief implies an increase in the study carried out between points close to each other. The presence of a statue with a section greater than 1 m, and the simultaneous absence of radiuses greater than this value, do not allow the algorithm to understand that it is a finite element belonging to the frame.

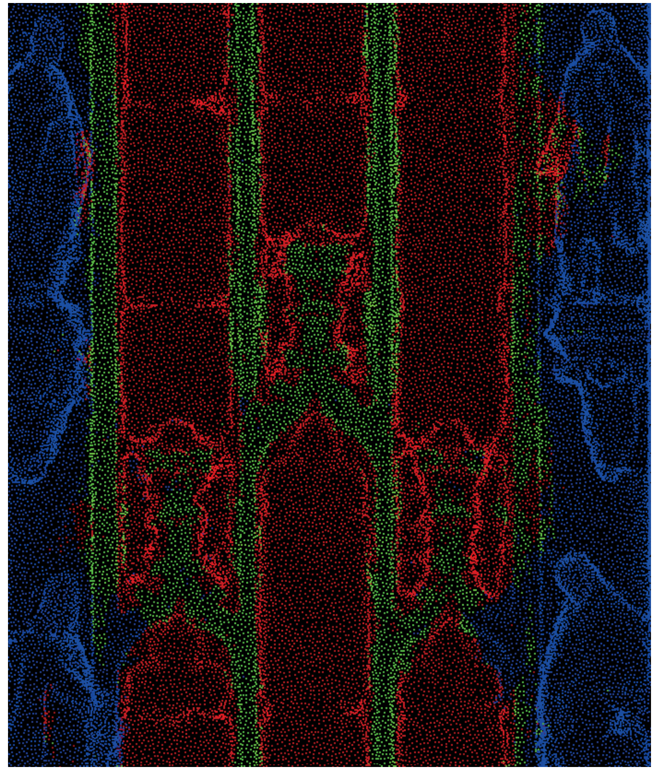

fig.83

fig.84 Result of third level of classification of window

After a few attempts, it became clear that at the third classification level (cloud intensity 20mm) it is inconsistent to identify small elements such as statues, due to the composition of the small radiuses (of the order of 0.05mm).

This causes an information overload that confuses the code instead of helping it to improve. It was therefore found that the main problem lied in the relationship between the scale of the classes to be identified, the intensity of the cloud and the radiuses for calculating the geometric characteristics. Working with a point cloud of 20 mm intensity at the base, the difficulty of identifying features at such different scales emerged: the window frame (10-15 cm) and the statues lying along the frame (approximately 1 m).

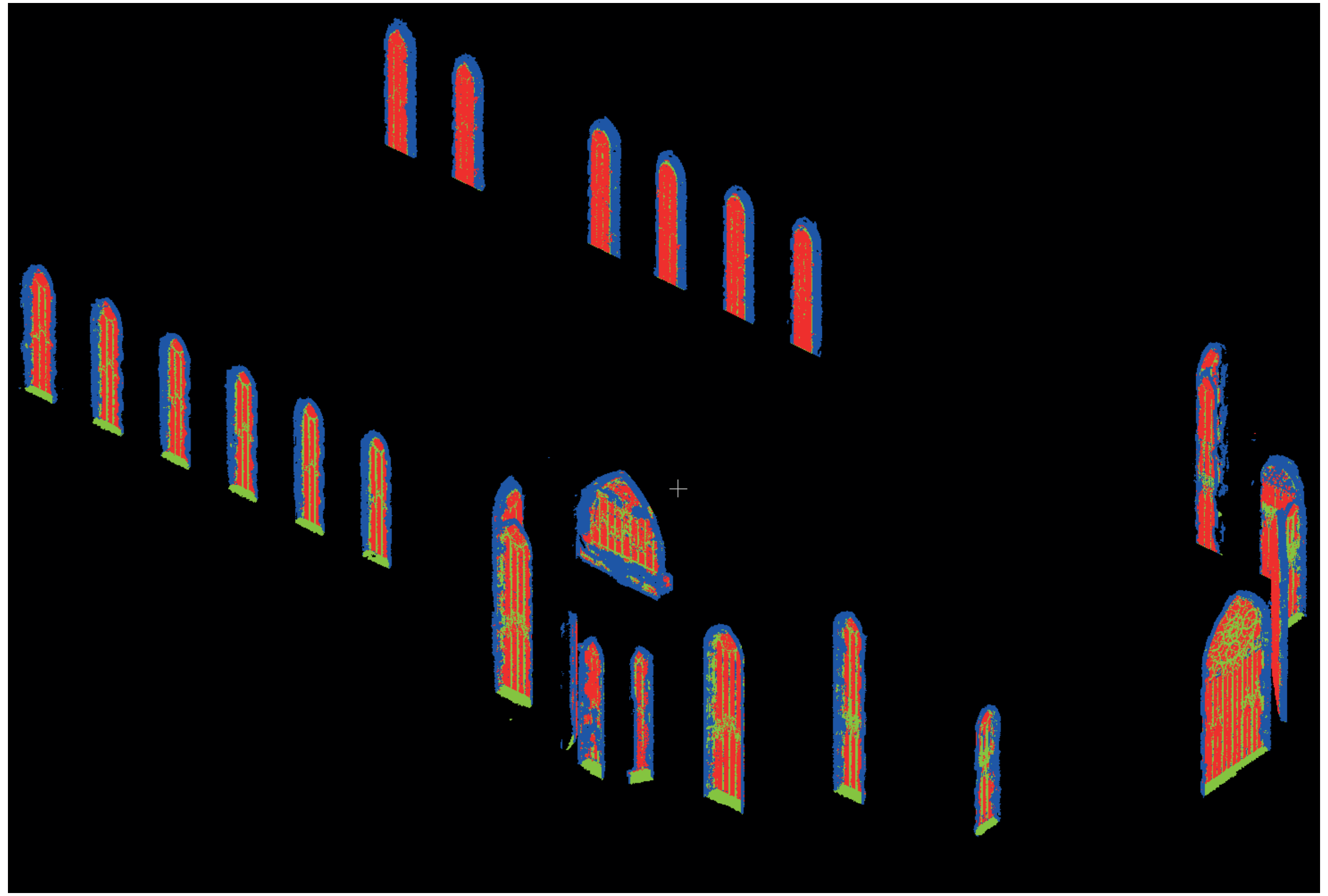

fig.84

From the figures above, one can notice the result of the automatic classification by the algorithm. Observing in general (fig. no. left) the segmentation of the points into the various classes, it may appear that the code has succeeded in identifying the different elements acceptably. The main problem with this result (fig. right) is the fact that no matter how small the classification errors may be, since we are dealing with such detailed and small elements, even a small error requires a very long cleaning phase and reassignment of the points to the correct classes. For the level of detail required by these elements, such cleaning would translate into an almost total manual re-classification by the operator, which would be in stark contrast to the objectives of this paper.

The solution adopted in this case was to take a step back and return to the second classification level, reducing the intensity of the cloud to 50 mm. Consequently, the number of elements to be identified also decreased, which became the outer frame and the architectural window hole, where the frame and glass are joined.
As a result, the geometric features were recalculated on the basis of different radiuses more coherent with the intensity of the cloud. We then returned to the third level of classification, with cloud intensity and radiuses of the initial characteristics, limiting the analysis to the architectural hole, in which two classes were distinguished: frame and glass.

Table 29
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	4,011,785
Train set points	1,082,574
Evalu set points	293,938

Table 29

Table 30
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	90 min
Time processing	360 sec.
Manual reassignment	Undefined

Table 30

Table 31
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Frame	0.90	0.84	0.87
Glass	0.87	0.92	0.90
Accuracy avg	0.89		
Macro avg	0.89	0.88	0.88
Weighted avg	0.89	0.88	0.88

Table 31

fig.85 Parts defined as training set and evaluation set

fig.86 Detail of defintiion of training set and evaluation set in a single window

BACK CALSSIFICATION AT LEVEL 2

WINDOWS

Going back to the 50mm resolution, it was possible to divide the window into two categories: the splay (which belongs to the wall as an architectural element) and the window itself.

Thanks to the creation of the study models that occurred in the attempt to classify all window elements on the third level, it was sufficient to interpolate the category information with a point cloud at a lower resolution to transfer this data. However, wanting to segment the data into only two categories instead of the previous three, the two classes within the window, i.e. frame and glass, were merged in order to isolate this research to only the two classes concerned. The decrease in resolution provided for a consistent adjustment of the survey radius between 0.1 m and 1 m every 10 cm.

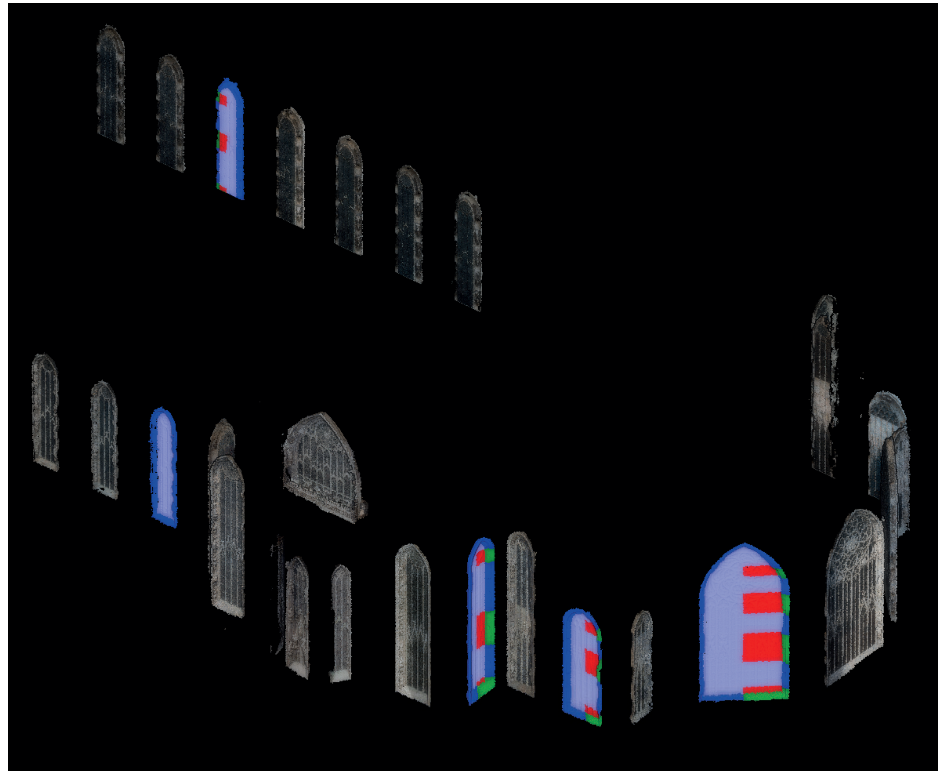

fig.85

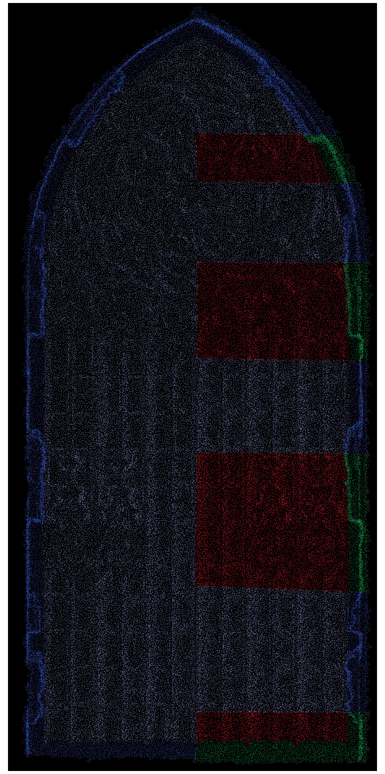

fig.86

Fro rom the following figure (fig. n), which represents the direct result of the automatic classification, we can apreciate the good quality of the segmentation performed by the algorithm.

The simplification of the objects derived from the reduced resolution of the cloud proved to be the most suitable solution to deal with the classification of these elements.

fig.87 Result of automatic classification of splay and windows

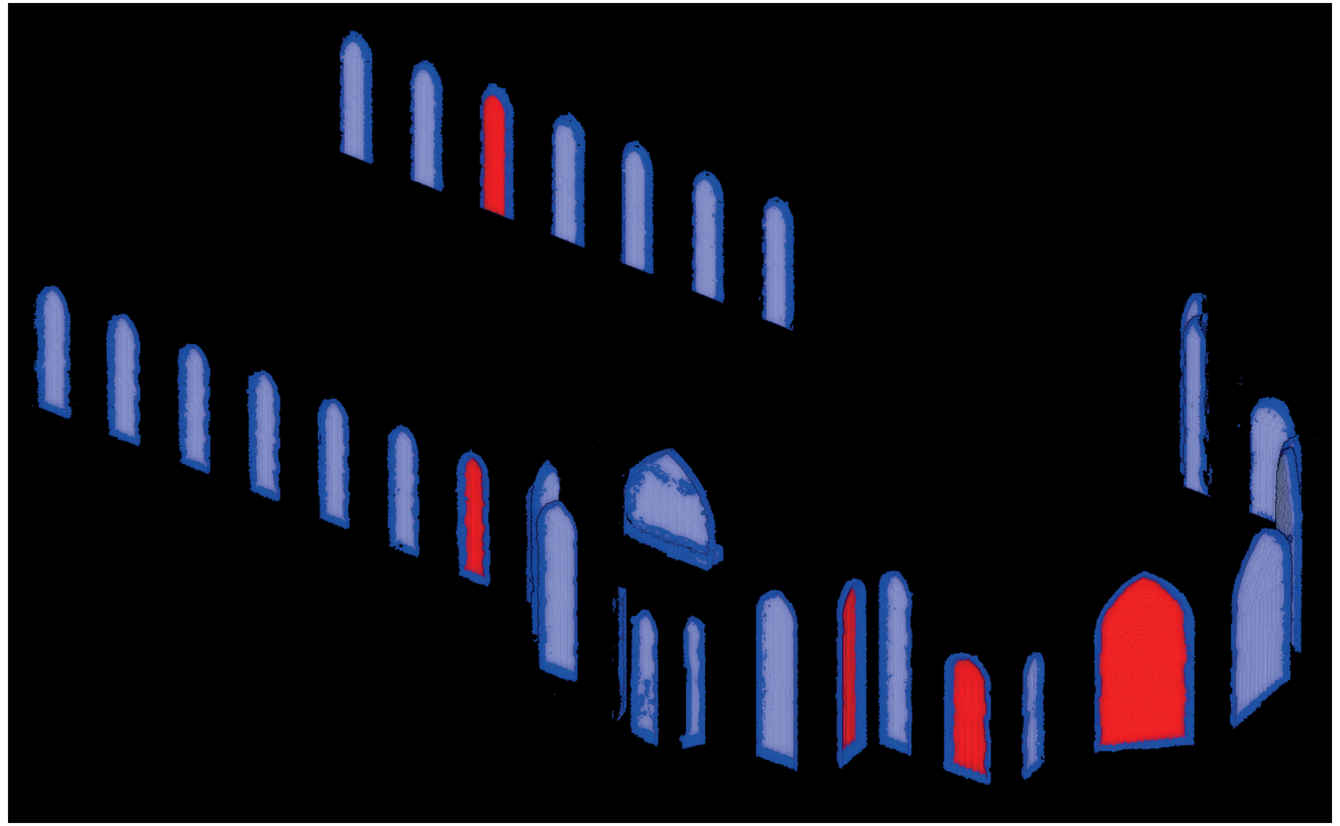

fig.87

Table 32
Summary of 3D point cloud number of points

	N. POINTS
Points to clasify	857,804
Train set points	203,126
Evalu set points	47,033

Table 32

Table 33
Summary of time including manual and automatic operation

OPERATION	TIME
Manual annotation	15 min
Time processing	180 sec.
Manual reassignment	40 min

Table 33

Table 34
Accuracy of automatic classification

	PRECISION	RECALL	F1-SCORE
Frame	0.96	0.94	0.95
Glass	0.97	0.98	0.97
Accuracy avg	0.97		
Macro avg	0.96	0.96	0.96
Weighted avg	0.97	0.97	0.97

Table 34

CRITICAL CONCLUSION

The objective that was intended to be achieved with this work can be considered fully accomplished. The possibility of addressing the subject of automatic classification on a Gothic architecture of this size made it possible to implement research into which factors most influence the algorithms for classification. The use of the MLMR approach was a success considering the level of detail that could be identified with automatic segmentation. Thanks to the tests performed during this work, the close relationship between the resolution of the point cloud and the measurement of the radiuses of the perceptual fields became evident. These two factors must necessarily be established in a coherent manner with the size and shape of the elements to be classified. Thanks to the problems and their elaborated solutions, it is possible to establish a hierarchy of the three main factors that influenced the results of automatic classification: the resolution of the data, the computation of the features and the realization of the study models. The first is the correct resolution to be used for the starting data, the second is the radiuses of the perceptual fields of the point features and, finally, the third is the realization of more or less complete study models. In situations where unacceptable results were obtained, the only useful solution turned to be to change the starting resolution of the datum. This is because setting an unsuitable resolution also affects the computation of the features and the subsequent realization of the training set and evaluation set due to the density of points that is not consistent with the shapes of the features to be identified. In the case where there are not enough points to locate the features, it is clear that there is a default error in the resolution. In the case where there is an excess of points, at first it will be more difficult to clearly perceive the division of one element from another, complicating the manual modelling phase. While, at a second moment the code will find itself examining an exaggerated amount of points compared to the size of the elements it has to distinguish, unnecessarily slowing down the processing time and, as seen in some cases, without obtaining acceptable classification results. For this reason, it was considered that the modification of features and the implementation of study models with other portions of the initial data are variables too dependent on the resolution of the starting data. In fact, in all the cases addressed in the course of this work, the definition of a resolution consistent with the elements to be classified was always decisive in order to obtain acceptable results. On the other hand, as far as generalization is concerned, it is considered to have provided encouraging results in the situations in which it was implemented. This strategy has proved to be an important resource in the context of the second-level classification of architectural elements such as pillars and vaults.
It is not excluded that this generalization could also be applied for the classification of pillars and vaults within other heritage buildings.
In summary, the goal of experimenting with automatic classification on large point clouds with very complex shapes can be considered to have been achieved. The experiences derived from the development of this work fit into the

national and international research landscape, contributing to the development and implementation of new systems for the management and enrichment of point clouds of historical heritage.

BIBLIOGRAPHY

[1] Achille, C., Fassi, F., Mandelli, A., Perfetti, L., Rechichi, F., Teruggi, S.: 'From a Traditional to a Digital Site: 2008-2019. The History of Milan Cathedral Surveys." In: Daniotti, B., Gianinetto, M., Della Torre, S. (eds) Digital Transformation of the Design, Construction and Management Processes of the Built Environment. Research for Development. Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-33570- 0_30

[2] Bariya, P., Novatnack, J. and Schwartz, G.: '3D Geometric Scale Variability in Range Images: Features and Descriptors', Int J Comput Vis 99, 232-255, 2012. https://doi.org/10.1007/s11263-012-0526-7.

[3] Ben-Shabat, Y., Lindenbaum, M. and Fischer, A.: '3DmFV: Three-Dimensional Point Cloud Classification in Real-Time Using Convolutional Neural Networks', in IEEE Robotics and Automation Letters, vol. 3, no. 4,
pp. 3145-3152, 2018. https://doi.org/10.1109/LRA.2018.2850061

[4] Cao, Y., and Scaioni, M.: 'A Pre-Training Method For 3d Building Point Cloud Semantic Segmentation', ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-2-2022, 219-226, 2020. https://doi.org/10.5194/isprs-annals-V-2-2022-219-2022.

[5] Cao, Y. and Scaioni, M.: '3DLEB-Net: Label-Efficient Deep Learning-Based Semantic Segmentation of Building Point Clouds at LoD3 Level', Appl. Sci., 11, 8996, 2021. https://doi.org/10.3390/app11198996.

[6] Chehata, N., Guo, L. and Mallet, C.: 'Airborne Lidar Feature Selection for Urban Classification Using Random Forests', Laserscanning, Sep 2009, Paris, France.

[7] Duan, S., Huang, P., Chen, M., Wang, T., Sun, X., Chen, M., Dong, X., Jiang, Z., & Li, D.: "Semi-supervised classification of fundus images combined with CNN and GCN.", Journal of applied clinical medical physics, 23(12), e13746, 2022. https://doi.org/10.1002/acm2.13746

[8] Grilli, E., Menna, F., and Remondino, F.: 'A Review of Point Clouds Segmentation and Classification Algorithms', Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W3, 339-344, 2017. https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017.

[9] Grilli, E. and Remondino, F.: 'Classification of 3D Digital Heritage', Remote Sens., 11, 847, 2019. https://doi.org/10.3390/rs11070847

[10] Grilli, E. and Remondino, F.: 'Machine Learning Generalization across Different 3D Architectural Heritage', ISPRS Int. J. Geo-Inf., 9, 379,2020. https://doi.org/10.3390/ijgi9060379

[11] Grilli, E. and Remondino, F.: 'Generalising machine learning across different 3D architectural heritages', ISPRS Int. J. Geo-Inf., 9, 379, 2020. https://doi.org/10.3390/ijgi9060379

[12] Grilli, E., Farella, E. M., Torresani, A., and Remondino, F.: 'Geometric Features Analysis for The Classification of Cultural Heritage Point Clouds', Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII- 2/W15, 541-548, 2019. https://doi.org/10.5194/isprs-archives-XLII-2-W15-541-2019

[13] Grillo, P.: 'Birth of a Cathedral. 1386-1418: la fondazione del Duomo di Milano', Milan, Mondadori, 2017.

[14] Guo, Y., Bennamoun, M., Sohel, F.: 'A Comprehensive Performance Evaluation of 3D Local Feature Descriptors', Int J Comput Vis 116, 66-89, 2016. https://doi.org/10.1007/s11263-015-0824-y

[15] Gupta, V., Mishra, V. K., Singhal, P. and Kumar, A.: "An Overview of Supervised Machine Learning Algorithm", 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 87-92, https://doi.org/10.1109/SMART55829.2022.10047618

[16] Holz, D., Ichim, A.E., Tombari, F., Rusu, R.B. and Behnke, S.: 'Registration with the Point Cloud Library: A Modular Framework for Aligning in 3-D', in IEEE Robotics & Automation Magazine, vol. 22, no. 4, pp. 110- 124, 2015. https://doi.org/10.1109/MRA.2015.2432331

[17] Li, Y., Ma, L., Zhong, Z., Liu, F., Chapman, M. A., Cao, D., & Li, J.: "Deep learning for lidar point clouds in autonomous driving: A review." IEEE Transactions on Neural Networks and Learning Systems, 32(8), 3412-3432, 2020. https://doi.org/10.1109/TNNLS.2020.3015992

[18] Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E. S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A., and Landes, T.: 'A Benchmark for Large-Scale Heritage Point Cloud Semantic Segmentation', Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2020, 1419-1426, 2020. https://doi.org/10.5194/isprs- archives-XLIII-B2-2020-1419-2020

[19] Matrone, F., Paolanti, M., Felicetti, A., Martini, M., and Pierdicca, R.: 'BubblEX: An Explainable Deep Learning Framework for Point-Cloud Classification', in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 6571-6587, 2022. https://doi.org/10.1109/JSTARS.2022.3195200

[20] Matrone, F., Grilli, E., Martini, M., Paolanti, M., Pierdicca, R., Remondino, F.: 'Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation', ISPRS Int. J. Geo-Inf., 9, 535, 2020. https://doi.org/10.3390/ijgi9090535.

[21] Malinverni, E. S., Pierdicca, R., Paolanti, M., Martini, M., Morbidoni, C., Matrone, F., and Lingua, A.: 'Deep Learning for Semantic Segmentation Of 3d Point Cloud', Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W15, 735-742, 2019. https://doi.org/10.5194/isprs-archives-XLII-2-W15-735- 2019.

[22] Khoury, M., Zhou, Q. and Koltun, V.: 'Learning Compact Geometric Features', IEEE International Conference on Computer Vision (ICCV), pp. 153-161, 2017. https://api.semanticscholar.org/CorpusID:8170545

[23] Pauly, M., Keiser, R. and Gross, M.: 'Multi-scale Feature Extraction on Point-Sampled Surfaces', Computer Graphics Forum, vol. 22, no. 3, pp. 281–289, 2003, doi: https://doi.org/10.1111/1467-8659.00675.

[24] Povoledo, E.: "A High-Maintenance Relationship for 637 Years, but Milan's Duomo Is Still Adored", The New York Times, Feb. 19, 2023. Why Milan's Duomo Cathedral Is Still So Adored - The New York Times (nytimes.com)

[25] Ren, X., Bo, L. and Fox, D.: 'RGB-(D) scene labelling: Features and algorithms', IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, pp. 2759-2766, 2012. https://doi.org/10.1109/CVPR.2012.6247999

[26] Rusu, R. B., Blodow, N. and Beetz, M.: "Fast Point Feature Histograms (FPFH) for 3D registration," 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, pp. 3212-3217, 2009. https://doi.org/10.1109/ROBOT.2009.5152473

[27] Spettu, F., Teruggi, S., Canali, F., Achille, C., & Fassi, F.: 'A Hybrid Model for The Reverse Engineering of The Milan Cathedral. Challenges And Lesson Learnt,' in Proceedings ARQUEOLÓGICA 2.0 - 9th International Congress & 3rd GEORES - GEOmatics and pREServation, Valencia, ES, pp. 281-291, 2021. https://dx.doi.org/10.4995/arqueologica9.2021.12138

[28] Thomas, H., Qi, C. R., Deschaud, J. E., Marcotegui, B., Goulette, F., & Guibas, L. J.: 'Kpconv: Flexible and deformable convolution for point clouds', In Proceedings of the IEEE/CVF international conference on computer vision, pp. 6411-6420, 2019. https://doi.org/10.48550/arXiv.1904.08889

[29] Teruggi, S., Grilli, E., Russo, M., Fassi, F., Remondino, F.: 'A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification', Remote Sens., 12, 2598, 2020. https://doi.org/10.3390/rs12162598.

[30] Wang, Y., Sun, Y., Liu, Z., Sarma, E. S., Bronstein, M. M. and Solomon, J. M.: 'Dynamic Graph CNN for Learning on Point Clouds', ACM Trans. Graph. 38, 5, Article 146, pp. 1-12, 2019. https://doi.org/10.1145/3326362

[31] Zhang, K., Teruggi, S., Ding, Y. and Fassi, F.: "A Multilevel Multiresolution Machine Learning Classification Approach: A Generalization Test on Chinese Heritage Architecture." Heritage, 2022, 5, 3970- 3992. https://doi.org/10.3390/heritage5040204

See the digital result of the work developed in this paper