heritage

MDPI

Article

Semantic Enrichment of Architectural Heritage Point Clouds Using Artificial Intelligence: The Palacio de Sástago in Zaragoza, Spain

Michele Buldo [1,*], Luis Agustín-Hernández [2] and Cesare Verdoscia [1]

[1] Department of Civil, Environmental, Land, Construction and Chemistry, Polytechnic University of Bari, 70125 Bari, Italy; cesare.verdoscia@poliba.it
[2] Department of Architecture, University of Zaragoza, 50018 Zaragoza, Spain; lagustin@unizar.es
* Correspondence: michele.buldo@poliba.it

Abstract: In the current landscape dominated by Artificial Intelligence, the integration of Machine Learning and Deep Learning within the realm of Cultural Heritage, particularly within architectural contexts, is paramount for the efficient processing and interpretation of point clouds. These advanced methods facilitate automated segmentation and classification, significantly improving both the clarity and practical use of data acquired from laser scanning and photogrammetry. The present study investigates the *Palacio de Sástago*—a prominent Renaissance palace in Zaragoza, Spain—and introduces a cutting-edge *modus operandi* for the automated recognition of architectural elements within the palace's inner courtyard. Employing the well-established Random Forest algorithm, implemented in a Python environment, the framework begins with a comprehensive evaluation of the geometric features identified in the LiDAR point cloud. This process employs the Mean Decrease in Impurity metric to evaluate the relevance of each variable. To boost the accuracy and efficiency of the final classifications, the features are refined post-assessment, enhancing both the training phase and the algorithm's later evaluation. The research's findings demonstrate significant potential, supporting advancements in CAD systems and HBIM that will enable more precise, automated modelling of architectural elements, thereby enhancing the accuracy of digital reconstructions and improving conservation planning for heritage sites.

Keywords: architectural heritage; point cloud; semantic enrichment; segmentation; classification; artificial intelligence; machine learning; random forest; geometric features; mean decrease impurity

Citation: Buldo, M.; Agustín-Hernández, L.; Verdoscia, C. Semantic Enrichment of Architectural Heritage Point Clouds Using Artificial Intelligence: The Palacio de Sástago in Zaragoza, Spain. *Heritage* **2024**, *7*, 6938–6965. https://doi.org/10.3390/heritage7120321

Academic Editor: Giovanni Castellazzi

Received: 11 November 2024
Revised: 2 December 2024
Accepted: 4 December 2024
Published: 9 December 2024

1. Introduction and State of the Art

The digital documentation of Cultural Heritage (CH), more specifically Architectural Heritage (AH), is crucial for its preservation, providing key benefits in terms of storage, accessibility, and resilience against degradation and anthropogenic impacts [1,2].

In this context, architectural surveying, supported by advanced techniques and technologies such as Structure from Motion (SfM) and Light Detection and Ranging (LiDAR), plays a crucial role in enhancing the acquisition of complex structures. Their application facilitates the transition from 2D to 3D representations, enabling the integration and refinement of traditional surveying methods [3–7].

A key outcome of these cutting-edge advancements is the generation of point clouds, which underpin further steps like parametric 3D modelling and comprehensive historical site documentation. Point clouds, derived from hybrid approaches that combine image-based (photogrammetry) and range-based (laser scanning) methods [8–10], enable accurate analysis through spatial coordinates and colourimetric attributes.

Recent advances in Neural Radiance Field (NeRF) technology have revolutionised the preservation of CH by employing Deep Learning and Computer Vision to replicate historical buildings from 2D photos by utilising Artificial Neural Networks (ANN) [11,12].

Despite advances in point cloud processing techniques, their integration into HBIM workflows remains constrained compared to Computer-Aided Design (CAD) systems, which effectively utilise segmented point clouds in Scan-to-CAD workflows to enhance design accuracy, streamline modelling processes, and improve stakeholder collaboration through precise architectural representations [13,14]. Modelling processes still largely rely on manual intervention, which limits both the efficiency and automation of 3D data processing and management. In order to tackle this issue, the Scan-to-BIM technique [15] presents a viable solution by using point clouds as a guide to create precise 'as-is' and 'as-built' BIM models, especially by obtaining exact building parameters from a geometric viewpoint [16–19].

These strategies promote the creation of complete digital information models that enable integration and application across a range of interdisciplinary domains and offer a strong basis for operational planning [20–22].

In the early stages, the point cloud data can be decimated and further processed through automated algorithms, optimising the workflow towards the final model. Several studies have outlined the most widely adopted approaches for segmenting and classifying point clouds, such as 'edge-based', 'region-based', 'model-based', 'attribute-based', and 'graph-based' methods, each utilising algorithms that cluster data features extracted from the point clouds [23]. These techniques focus on different aspects of the point cloud data, such as geometric boundaries, spatial continuity, object models, data attributes, and relational structures, allowing for a more comprehensive and context-sensitive segmentation and classification process. When combined in a hybrid approach, they can further enhance performance [24,25].

The intricate geometric topology of cultural and architectural environments poses considerable challenges for the automated application of these algorithmic techniques. In this regard, Artificial Intelligence (AI) proves to be a powerful ally, significantly enhancing the precision, efficiency, and scalability of heritage management. By providing advanced tools for documentation, restoration, and analysis, AI offers unprecedented support in overcoming these challenges and optimising preservation practices [26,27].

However, ethical concerns also arise, particularly regarding the risk of data bias and the potential loss of human agency in decision-making, which could compromise the authenticity of preservation efforts [28]. Therefore, while AI can significantly optimise workflows and improve conservation practices, it is crucial to strike a balance between technological progress and ethical considerations to avoid unintended consequences [29,30].

Moreover, a key limitation lies in the difficulty in obtaining uniformly structured training datasets, which are crucial for ensuring the accuracy of AI algorithms [31,32]. Consequently, automated approaches frequently require human oversight to ensure proper implementation and verification, underscoring the necessity of an in-depth understanding of the context to achieve reliable outcomes.

When supported by well-organised training datasets, AI techniques allow for automated segmentation and classification based on user-defined criteria and properties [33].

In the field of AH, leveraging Machine Learning (ML) and Deep Learning (DL) make it easier to group data points or patches with comparable colorimetric and/or geometric properties into homogenous subsets [34–36]. In this context, until a few years ago, the use of ML was restricted because the majority of the work relied on statistical toolboxes applied to small datasets that are normally not publicly available [37].

Recent developments have resulted in the definition of a number of classification techniques for 3D models [38–41]: These include segmenting ortho-rectified images [42–44], classifying textures from 2D images processed through UV mapping and reprojected

onto 3D models, segmenting point clouds using colorimetric and geometric feature extraction [45–48], and classifying multi-level and multi-resolution images [49].

The incorporation of DL approaches has led to notable advancements in semantic annotation in HBIM procedures [50,51]. The segmentation and classification of architectural features such as arcs, columns, walls, windows, and more has been made possible by notable approaches like PointNet, Pointnet++, and Dynamic Graph Convolutional Neural Network (DGCNN) [52–55] which demonstrate the field's dynamic character and the ongoing development of techniques for historic buildings.

1.1. Aims and Overview of the Research

This study delves into the segmentation and classification of point clouds using Machine Learning techniques, arising from an initial bibliographic review of the state of the art (see Section 1). Specifically, it addresses the semantic enrichment of a LiDAR point cloud of the Palacio de Sástago in Zaragoza (Spain), employing the Random Forest (RF) algorithm within the PyCharm environment of Python.

The primary aim is to enhance the point cloud's semantic content by breaking it down into its architectural components, thereby establishing a foundation for future reconstructions within the framework of Scan-to-BIM.

A target bibliographic review was first carried out to investigate the best strategies for classification and segmentation. A bibliometric analysis was performed using the 'bibliometrix' package in RStudio and its web interface 'biblioshiny' to create a network map (Figure 1). This graph, based on search terms such as 'Cultural Heritage', 'Point Cloud', 'Segmentation', and 'BIM', highlights key concepts related to semantic enrichment. It was constructed by analysing 470 English-language papers from Scopus and Web of Science published between 2012 and 2024.

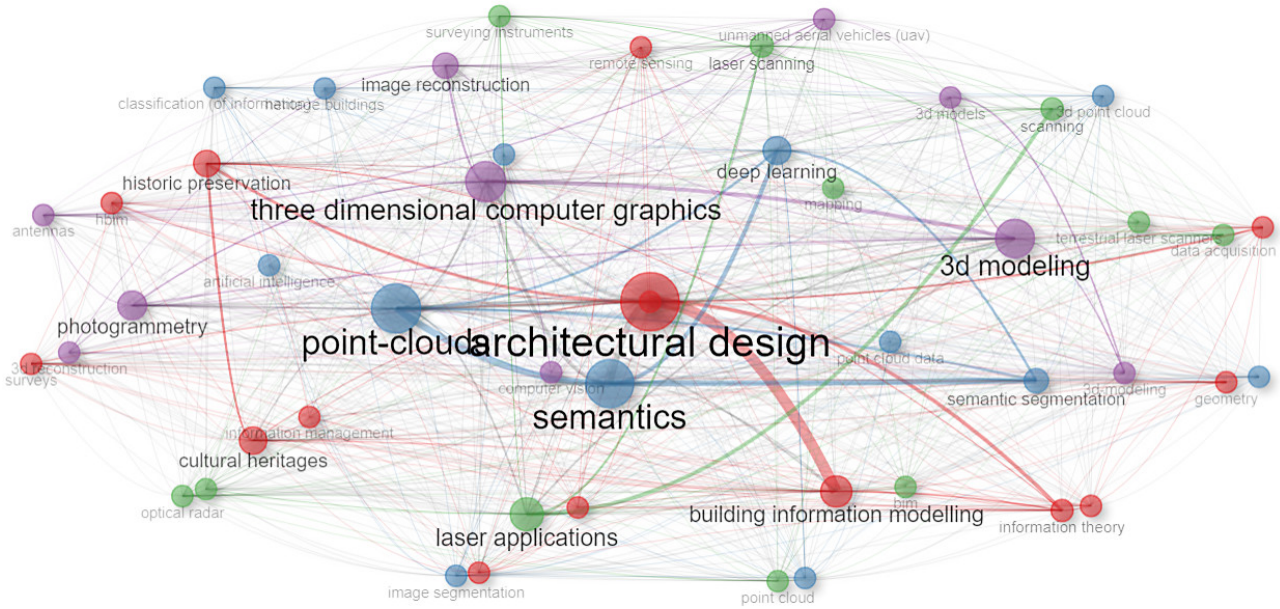

Figure 1. Network map illustrating the semantic enrichment of Architectural Heritage, generated using the R programming language.

Afterward, the RF algorithm was implemented, to initially identify the most relevant geometric features related to the point cloud and subsequently to refine the training of the algorithm, aiming to achieve the highest possible accuracy in the semantic data output.

More specifically, RF is an ensemble learning technique that uses random subsets of the training data and feature set to build multiple decision trees. Through majority voting, it aggregates these trees' predictions and chooses the class with the highest number of

votes overall as the final classification result [56], thereby reducing overfitting and improving both reliability and accuracy.

After evaluating the best features to use through the 'Mean Decrease in Impurity' (MDI) measure—which quantifies the importance of variables in the model based on the reduction in impurity during node splitting—the algorithm was first trained and then validated. This was performed by manually segmenting different portions of the point cloud for each phase.

In addition to the best features, normal features, height, and those related to the architectural element's dimension were also included. In the final phase of the project, an evaluation of the algorithm's metrics was conducted to gain a clearer understanding of the diverse combinations of features. This assessment is facilitated by the discretisation of the point clouds, which enhances visual observation, as will be detailed in the next sections.

1.2. The Case Study

This study focuses on the Palace of the Counts of Sástago (Palacio de Sástago), located in the old city (Casco Antiguo) of Zaragoza, standing as a significant testament to the socio-political and cultural developments of the 16th century (Figure 2).

Figure 2. Inner courtyard of the *Palacio de Sástago* in Zaragoza, Spain.

Commissioned by Don Artal de Alagón y Luna, the third count of Sástago, the palace has served various roles throughout history, reflecting the evolution of architectural styles and societal values in Aragon. Constructed during this transformative period, it symbolises the formidable influence of its patron, having been a crucial venue for notable events, including hosting Spanish monarchs and serving as a military headquarters during pivotal conflicts, most notably the Siege of Zaragoza.

The building was eventually repurposed after suffering severe damage during the burning of the adjoining Convent of San Francisco.

The palace has served as both a police headquarters and the residence of the Captaincy General of Aragon for centuries, demonstrating its versatility and significance in Zaragoza's urban landscape. Renowned for its eclectic neo-baroque architecture, the palace includes the distinctive 'Casa Zorraquino' on its ground floor.

Its brick façade, shaped by the region's limited access to stone, highlights an innovative use of local materials that blend seamlessly with the surrounding environment. Inside, the palace opens to a grand Renaissance-style courtyard, leading into an entrance hall and a noble staircase, along with prominent rooms like the Throne Room.

The inner courtyard exemplifies a refined integration of classical architecture and elaborate ornamentation. Its symmetrical layout features partially fluted columns and graceful arches that create a balanced and elegant ambiance. These columns are often decorated with carved motifs and floral patterns, adding depth to the visual experience, while the courtyard walls are embellished with intricate friezes, delicately sculpted Ionic capitals, and decorative windows, all of which contribute to the building's rich historical essence.

Despite limited public access, the Palace of the Counts of Sástago has evolved into a vibrant cultural hub; it includes a modernist library, and frequently hosts exhibitions showcasing works by renowned artists like Pablo Picasso, Salvador Dalí, and Manuel Viola. This transformation reflects the palace's ongoing relevance in contemporary society and its role in fostering artistic dialogue.

2. Materials and Methods

With regard to the building's inner courtyard, the adopted approach (Figure 3) seeks to automatically identify recurrent architectural elements in the point cloud. Subsequently, these parts are classified and catalogued within a Scan-to-BIM framework, simplifying upcoming modelling processes and boosting efficiency.

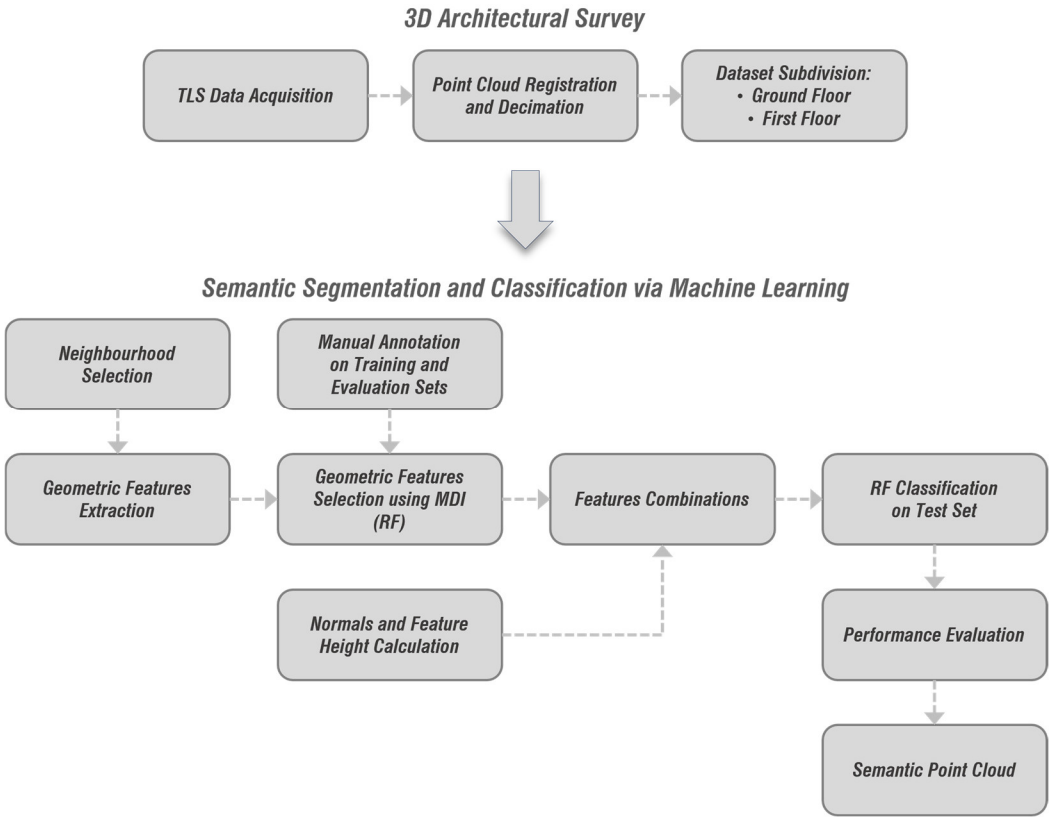

Figure 3. Methodology workflow.

A structured processing workflow was developed to handle raw point cloud data collected via laser scanning. This process involved identifying and extracting the most critical geometric features, which, along with carefully curated manual annotations, were used to train a Random Forest Machine Learning algorithm. The purpose was to empower the algorithm to autonomously classify the remaining point cloud data with high accuracy. To streamline the training process and reduce redundancy in error assessment, the study area was segmented into two main subsets: the ground floor and the first floor. This division was intentionally designed to leverage the unique architectural characteristics specific to each floor, enhancing the model's performance and precision in classifying diverse architectural elements.

2.1. Three-Dimensional Architectural Survey

The surveying operations were carried out using the CAM2® FARO Focus M 70 laser scanner, which features a vertical field of view of 300° and a horizontal one of 360°. With a ranging error of ±3 mm and an unambiguity interval of 614 m, this scanner can capture up to 0.5 million points per second. A total of 32 indoor scans were completed: 21 scans on the ground floor and 11 on the first floor. Each scan was performed within a radius of 10 m, achieving a resolution of 11 million points, with a net scan duration of 5 min and 11 s, resulting in a scan size of 5156 × 2134 points.

Point cloud registration was conducted using Autodesk Recap software, employing both automated and manual alignment techniques. The manual process involved identifying three corresponding points between pairs of scans to ensure accurate alignment. The initial registration yielded an accuracy of 6 mm, which was further refined through a noise reduction process using Statistical Outlier Removal (SOR) within CloudCompare software. Additionally, subsampling was performed at 2 cm intervals to optimise the computational file size. As a result, the entire dataset, covering an area of approximately 12.20 m × 8.90 m, was divided into two groups: one for the ground floor of the inner courtyard and the other one for the first floor (Figure 4).

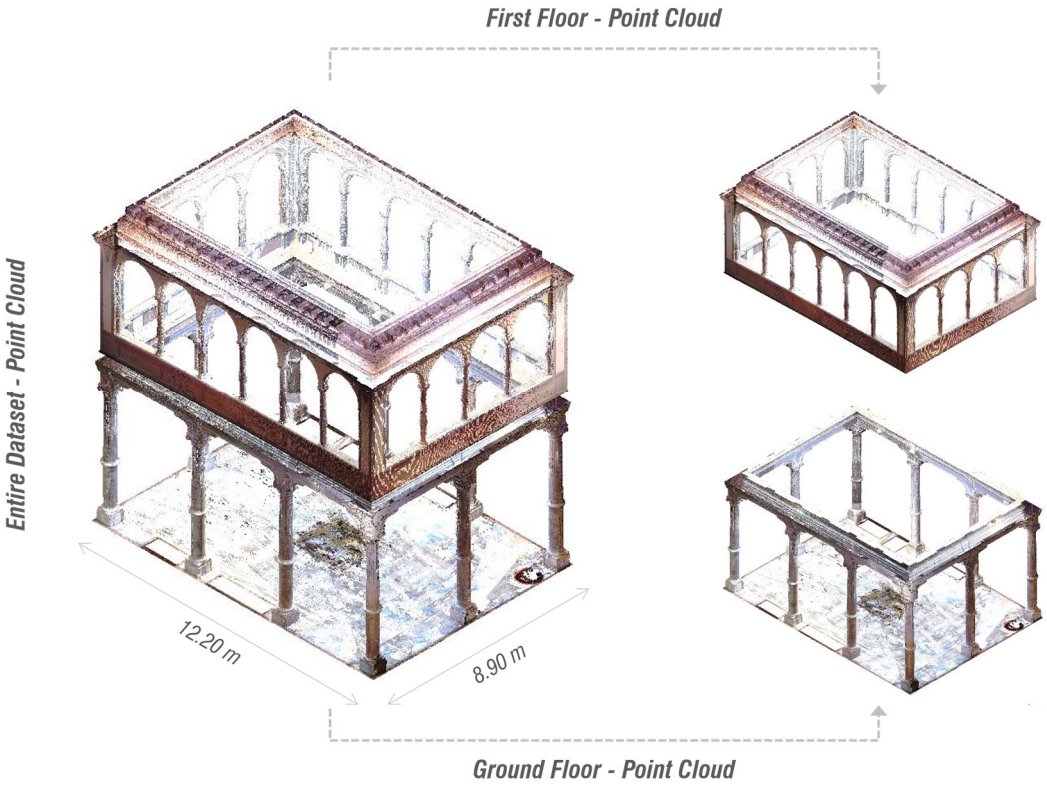

Figure 4. Subdivision of the complete point cloud into two groups: ground floor and first floor.

Subsequent processing steps were carried out concurrently for both levels. For the training phase, a specific subset of points was identified, while another subset was set aside for model evaluation. The remaining data for each floor were classified as the test set for the final classification. A comprehensive summary of the dataset is provided in Table 1.

Table 1. Full dataset specification.

	No. Points	Point Spacing
Point cloud_ground floor	5,323,756	6 mm
Subsampled point cloud	427,068	2 cm
Training	53,713	2 cm
Evaluation	54,432	2 cm
Test	318,923	2 cm
Point cloud_first floor	4,579,345	6 mm
Subsampled point cloud	440,695	2 cm
Training	56,032	2 cm
Evaluation	54,914	2 cm
Test	329,749	2 cm

2.2. Semantic Segmentation and Classification via Machine Learning

The subsampled and enhanced point cloud for each floor is employed as input for the semantic segmentation process utilising Machine Learning techniques.

To implement this, a sequence of phases is executed, as detailed in the workflow and summarised below: (a) neighbourhood selection; (b) extraction of the geometric features; (c) manual annotation for training and evaluation sets; (d) selection of the geometric features using Mean Decrease in Impurity; (e) calculation of the normals and feature height; (f) combinations of the features; (g) Random Forest classification on the test set and performance evaluation; and (h) semantic point cloud.

2.2.1. Neighbourhood Selection

A detailed series of covariance features were extracted from the two subsets of point clouds using CloudCompare software, with feature selection primarily guided by varying local neighbourhood radii. The term 'local neighbourhood radius' defines the spatial area surrounding each point in the point cloud that is considered when extracting geometric features [57]. This parameter specifies the maximum distance within which adjacent points are included in the computation of the geometric attributes for a central point.

The core principle underlying this approach is that the geometric features of a point are greatly influenced by its neighbouring points. By adjusting the local neighbourhood radius, the extent of space considered in the analysis of each point's features can be precisely controlled. To achieve this, a systematic radius increment was established, ranging from 10 to 20 cm for the analysis of smaller elements, while larger architectural components were assessed with radii approaching one metre.

2.2.2. Extraction of the Geometric Features

The term 'covariance features' denotes geometric characteristics derived from the covariance matrix of a set of points in three-dimensional space, often known to as the 3D structure tensor [58–60], which is calculated using the spatial information of points within a defined local neighbourhood. The covariance matrix is computed for each point in the point cloud relative to the centroid, revealing variance values along its diagonal.

Through a Principal Component Analysis (PCA), eigenvectors and eigenvalues are derived, which describe the local geometry of the point cloud, with the eigenvalues ordered as $\lambda_1 \geq \lambda_2 \geq \lambda_3$, representing the primary components of the spatial distribution.

Using these eigenvalues, several 3D properties, referred to as dimensional features, are calculated: Linearity [Equation (1)], Planarity [Equation (2)], and Sphericity [Equation (3)]. Additionally, these eigenvalues serve as the foundation for further measurements, including Omnivariance [Equation (4)], Anisotropy [Equation (5)], Eigenentropy [Equation (6)], Sum of Eigenvalues [Equation (7)], and Surface Variation, also known as the Change of Curvature [Equation (8)]. Verticality [Equation (9)], which is categorised as a 'normal-based feature', and Height [Equation (10)], considered as a 'height-based feature', are additional features that can be obtained.

The different types of features, along with their corresponding equations, are summarised in Table 2.

Table 2. Geometric features explanation.

Feature Typology	Name	Equation	
Covariance	Linearity	$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$	(1)
	Planarity	$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}$	(2)
	Sphericity	$S_\lambda = \frac{\lambda_3}{\lambda_1}$	(3)
	Omnivariance	$O_\lambda = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$	(4)
	Anisotropy	$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}$	(5)
	Eigenentropy	$E_\lambda = -\sum_{i=1}^{3} \lambda_i \ln(\lambda_i)$	(6)
	Sum of Eigenvalues	$\Sigma_\lambda = \sum_{i=1}^{3} \lambda_i$	(7)
	Surface Variation	$C_\lambda = \frac{\lambda_3}{\Sigma_\lambda}$	(8)
Normal-based	Verticality	$V_\lambda = 1 - \lvert([0\ 0\ 1], e_3)\rvert$	(9)
Height-based	Height	Z Coordinate	(10)

An in-depth examination of the 3D point cloud is made easier by these features, which clarify many aspects of local geometry structures. The scale at which these geometric features are assessed inside the point cloud shifts when the neighbourhood radius is modified.

Specifically, Linearity measures the alignment of points along a straight line, with higher values indicating elongated structures. Planarity evaluates how closely points resemble a flat surface, while Sphericity describes their distribution in a spherical manner, with higher values suggesting rounder forms. Omnivarance captures overall spatial variability, and Anisotropy quantifies directional disparities in point distribution. Eigenentropy reflects the disorder in the point cloud, with higher values indicating more complexity.

The Sum of Eigenvalues indicates spatial dispersion, and Surface Variation assesses curvature variability. Verticality measures the alignment along the vertical axis, while Height identifies elevation differences within the point cloud.

The graphics in Figure 5 showcase examples of 'ad hoc' features, which are further examined in the context of the geometric dimensions, specifically the radius and diameter, of architectural elements.

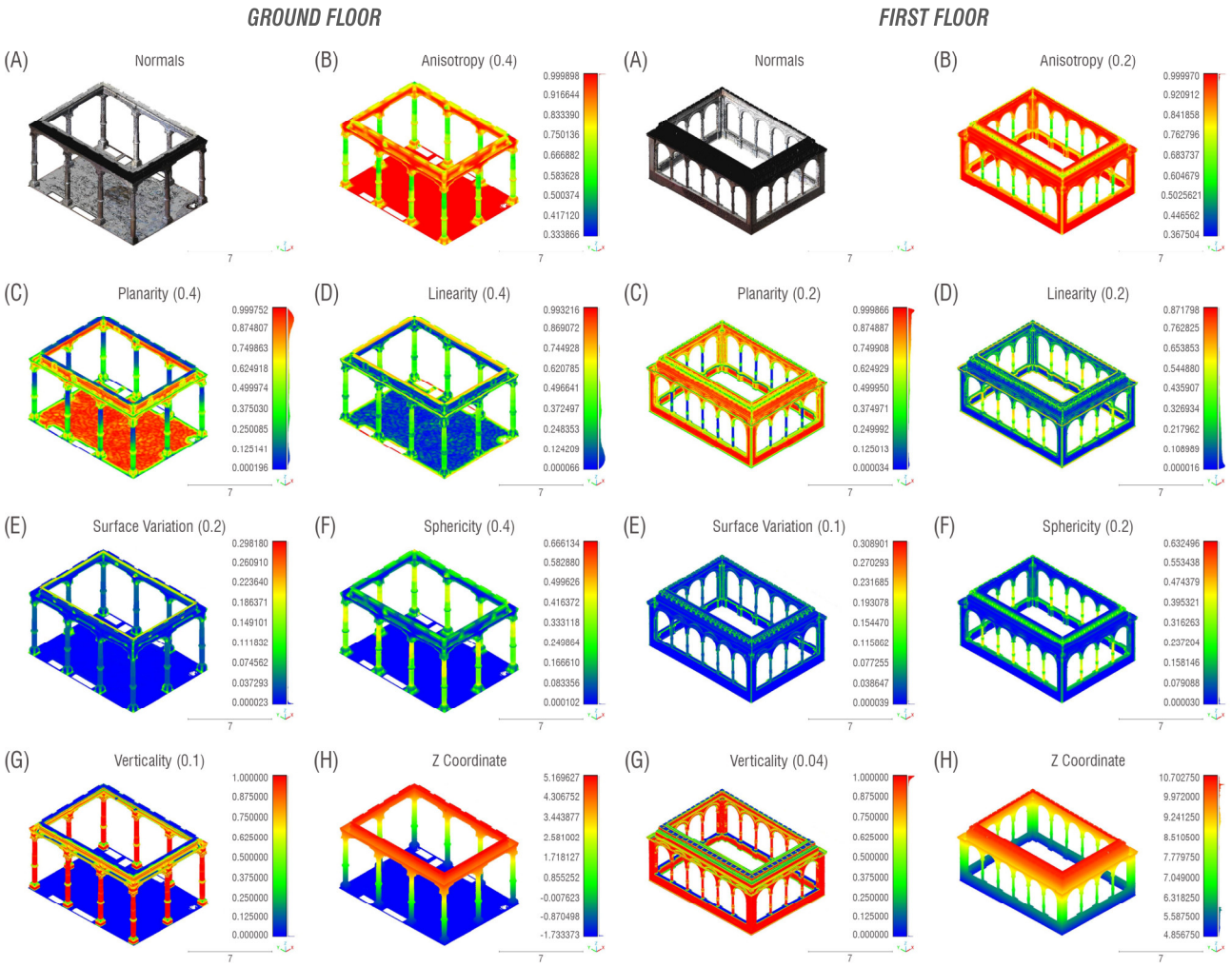

Figure 5. Example of geometric features used during the training phase, pertaining to the point clouds of the ground and first floors: (**A**) Normals; (**B**) Anisotropy; (**C**) Planarity; (**D**) Linearity; (**E**) Surface Variation; (**F**) Sphericity; (**G**) Verticality; (**H**) Z coordinate.

Calculating normals can improve the classification of complex scenes, though their necessity is debated [61]. Normals are vectors perpendicular to surfaces at each point, representing surface orientation. Each normal is a 3D vector (*Nx*, *Ny*, *Nz*), where *Nx*, *Ny*, and *Nz* are the components along the *x*, *y*, and *z* axes, respectively.

At this point, the focus has been narrowed to the extraction of normal-based and covariance features. After the Machine Learning algorithm automatically identifies the most significant features, normals and height features will be computed. These will be essential for testing various combinations that the algorithm can use to automatically partition the test set.

2.2.3. Manual Annotation for Training and Evaluation Sets

The ground floor and first floor point clouds were manually segmented in order to provide an appropriate training and evaluation set for identifying architectural elements in the remaining section of the point cloud, which will act as the test set. It is possible to integrate certain geometric features, point coordinates, surface normals, and colour qualities (such as RGB values) with these segmented data, which have been designated as classes.

Based on the categorisation of architectural elements using the ontological taxonomy supplied by the Getty Research Institute's Art & Architecture Thesaurus© (AAT), the point cloud was semantically split [62]. In terms of glossary set-up, this method adheres

to classical architectural principles and falls in line with the 'Information Standards in Practice' (ISO and NISO). For the ground floor, the segmentation follows a hierarchical structure, from bottom to top, as follows: Pavement, Base, Lower Shaft, Ring, Upper Shaft, Capital, Corbel, Architrave, Frieze, and Cornice.

For the first floor, the segmentation includes: Parapet, Base, Lower Shaft, Ring, Upper Shaft, Capital, Arch, Stringcourse, Cornice, and Corner Column (Figure 6). For architectural analysis, ten classes were manually created from each floor.

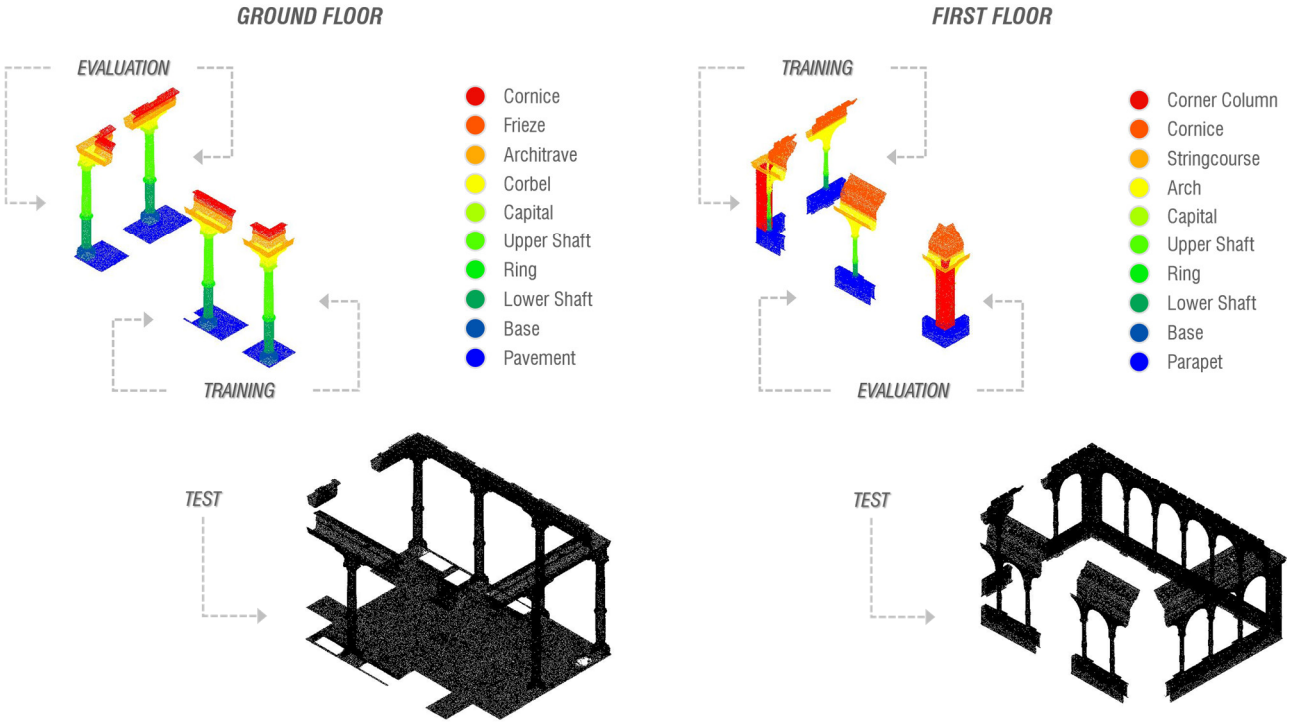

Figure 6. Manual annotation of the training and evaluation sets for the automatic identification of the test set classes using Machine Learning on the ground and first floors.

2.2.4. Selection of the Geometric Features Using Mean Decrease in Impurity

A preliminary multiscale analysis of feature relevance was carried out using the Random Forest algorithm in a Python environment (PyCharm) in order to determine which geometrical features were most pertinent for the training stage. In the beginning, 63 features were retrieved for the ground floor and 72 features for the first floor. A standard metric for 'feature importance' in Random Forests [63,64] was used: the 'Mean Decrease in Impurity' (MDI). This metric calculates the contribution of each feature to the training process of lowering impurity in the decision tree nodes. Features that improve data separation by significantly reducing impurity are deemed more significant.

The following procedures were used to calculate the feature importance ranking:

- Data Preparation: After class labels (dependent variables) and features (independent variables) were added, the dataset was divided into training and test sets.
- Random Forest Training: The RF classifier was trained on the point cloud data using the Python 'scikit-learn' library. The model determined feature importance automatically throughout training.
- Feature Importance Extraction: Following the training phase, feature importance scores were obtained using the model's 'feature_importances_' attribute.
- Feature Ranking: Features were ranked to create a clear visual representation of their importance. These rankings are displayed in Figure 7, with the local neighbourhood radius represented on the x-axis and the normalised feature importance scores shown on the y-axis.

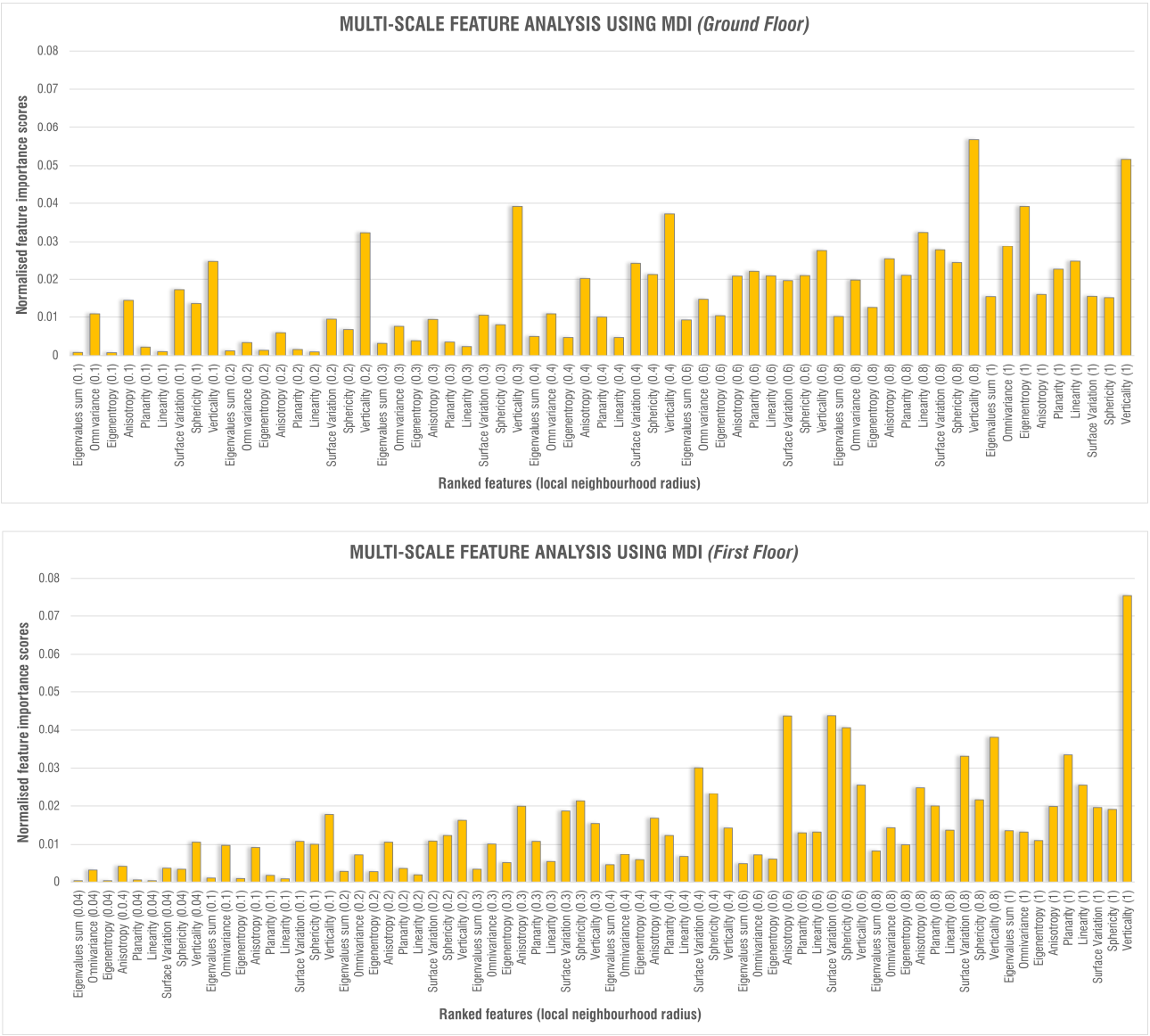

Figure 7. Feature importance ranking for the ground floor (on the **top**) and first floor (on the **bottom**), derived from the Mean Decrease in Impurity (MDI) analysis.

At this point, the 15 most important features were carefully chosen from the 63 features on the ground floor and the 72 features on the first floor, respectively, to maximise training time and assess the prediction model's overall performance. An additional set of 6 features was subsequently selected in an iterative process that used the same strategies. In order to minimise overfitting errors, this was carried out to see how they affected the model's behaviour when the overall number of features was progressively decreased.

According to Table 3, Verticality is a significant feature on both floors, with 'Verticality (1)' being more prominent on the first floor. While additional features like Eigenentropy and Omnivariance are emphasised on the ground floor, Surface Variation and Anisotropy are also crucial on the first floor.

Table 3. Feature importance scores of the 15 features for the ground and first floors.

Ground Floor		First Floor	
Feature (Radius)	**Score**	**Feature (Radius)**	**Score**
Verticality (0.8)	0.05674217	Verticality (1)	0.07552335
Verticality (1)	0.0516716	Surface Var. (0.6)	0.0438021

Verticality (0.3)	0.03921137	Anisotropy (0.6)	0.04372397
Eigenentropy (1)	0.03920605	Sphericity (0.6)	0.04066766
Verticality (0.4)	0.03728326	Verticality (0.8)	0.03792759
Linearity (0.8)	0.03241092	Planarity (1)	0.03345083
Verticality (0.2)	0.03235743	Surface Var. (0.8)	0.03309696
Omnivariance (1)	0.02869746	Surface Var. (0.4)	0.03010814
Surface Var. (0.8)	0.02767854	Linearity (1)	0.02567088
Verticality (0.6)	0.02750381	Verticality (0.6)	0.02564896
Anisotropy (0.8)	0.02535267	Anisotropy (0.8)	0.02492746
Linearity (1)	0.02473062	Sphericity (0.4)	0.0232871
Verticality (0.1)	0.02466849	Sphericity (0.8)	0.02164398
Sphericity (0.8)	0.02439382	Sphericity (0.3)	0.02137714
Surface Var. (0.4)	0.02417008	Planarity (0.8)	0.02003774

Verticality stands out as a noteworthy feature in both floors, as shown in Table 4, but other specific features also make a substantial contribution to the overall score, giving each floor a unique profile.

Table 4. Feature importance scores of the 6 features for the ground and first floors.

Ground Floor		**First Floor**	
Feature (Radius)	**Score**	**Feature (Radius)**	**Score**
Eigenentropy (1)	0.20324868	Verticality (1)	0.23757974
Verticality (0.8)	0.19481669	Planarity (1)	0.22981696
Verticality (1)	0.19275127	Surface Var. (0.6)	0.16260546
Linearity (0.8)	0.18289954	Verticality (0.8)	0.1468888
Verticality (0.4)	0.12677354	Sphericity (0.6)	0.11651344
Verticality (0.3)	0.09951028	Anisotropy (0.6)	0.10659559

Then, as previously illustrated in Figure 5, a customised set of 6 'ad hoc' features was added to each level. These features were picked because they closely matched the manually classified classes and could effectively depict distinct architectural components.

Features with a radius of 0.4 m to match the diameter of the column, including Anisotropy, Planarity, Linearity, and Sphericity, were carefully chosen for the ground floor. Verticality was given a radius of 0.1 m, which is half the radius of the column, and Surface Variation was chosen with a radius of 0.2 m, which is the same as the radius of the column.

The first floor's chosen features were Sphericity, Linearity, Planarity, and Anisotropy, each having a radius of 0.2 m to correspond with the diameter of the column. The radius for Surface Variation was set at 0.1 m, which is the same as the radius of the column, and the radius for Verticality was set at 0.04 m, which is half of the radius of the column. Certain features were purposefully left out on both levels, including Sum of Eigenvalues, Eigenentropy, and Omnivariance. Unlike those chosen using the MDI method, these features, at least visually, gave results that did not match expectations.

2.2.5. Calculation of the Normals and Feature Height

When combined with other geometric features, parameters like normals and feature height (Z coordinate) can greatly aid in the segmentation and classification of point clouds. The model's capacity to distinguish between various entities is improved by these extra factors, which also enrich surface descriptions. As previously stated, normals provide crucial information about the orientation of surfaces within the point cloud, representing the perpendicular direction to each surface at every point. This function is essential for capturing the intricate geometry of surfaces and is especially helpful for

segmenting objects such as floors, walls, sloping structures, and other patterns that would be tricky to identify with just spatial coordinates.

At the same time, the Z coordinate is quite useful for investigating the environment's three-dimensional structure. It makes it easier to spot changes in vertical space, which improves comprehension of the building's general layout. In order to achieve more precise and thorough segmentation, this is particularly helpful for identifying key architectural detailing, such as height variations between floors, balconies, or other raised features. In light of these benefits, the Z coordinate and normals were incorporated into the model to evaluate their performance throughout the training and classification phases.

2.2.6. Combinations of the Features

Three alternative combinations were carried out for each iteration with a different set of features (63–15–6–6 'ad hoc', specifically for the ground floor; 72–15–6–6 'ad hoc', for the first floor), giving a total of 12 combinations per floor, as shown below.

Ground floor:

- 63 features; 63 features + Z coordinate; 63 features + Z coordinate + normals;
- 15 features; 15 features + Z coordinate; 15 features + Z coordinate + normals;
- 6 features; 6 features + Z coordinate; 6 features + Z coordinate + normals;
- 6 'ad hoc' features; 6 'ad hoc' features + Z coordinate; 6 'ad hoc' features + Z coordinate + normals.

First floor:

- 72 features; 72 features + Z coordinate; 72 features + Z coordinate + normals;
- 15 features; 15 features + Z coordinate; 15 features + Z coordinate + normals;
- 6 features; 6 features + Z coordinate; 6 features + Z coordinate + normals;
- 6 'ad hoc' features; 6 'ad hoc' features + Z coordinate; 6 'ad hoc' features + Z coordinate + normals.

The Random Forest approach was used to test the 12 combinations, and the weighted F1-score and model training time were the two main metrics used for comparison. When working with imbalanced datasets, the weighted F1-score—calculated using the following formula [Equation (11)]—is a valuable performance evaluation metric for classification models since it takes into account variations in the number of samples in each class:

$$F1_{weighted} = \frac{\sum_i w_i * F1_i}{\sum_i w_i} \quad (11)$$

where

$F1_i$ = F1-score for class i;

w_i = weight associated with class i.

All classes are used for the summation. The standard F1-score formula [Equation (12)] is used to determine the F1-score for a single class:

$$F1_i = \frac{2\ Precision_i * Recall_i}{Precision_i + Recall_i} \quad (12)$$

where

$Precision_i$ = precision for class i;

$Recall_i$ = recall for class i.

The weights w_i are determined by calculating the ratio of the number of samples in class i to the overall number of samples in the dataset.

3. Results and Discussion

The following section presents the various combinations of features, along with the outcomes of the Machine Learning classification. These results are accompanied by evaluation metrics carefully selected to serve the intended purpose, offering a clear assessment of the model's performance.

3.1. Results of the Feature Combinations

In addition to the weighted F1-score, the hyperparameter that defines the number of decision trees in the forest, known as the number of estimators, is also considered. Each decision tree contributes to the final prediction, and the number of trees can significantly influence the model's performance and generalisation capability. Training time is another parameter used for feature comparison.

The following four combinations can be used to analyse the results of the ground floor analysis, as seen in Figure 8 and Table 5:

(a) 63 features; 63 features + Z coordinate; 63 features + Z coordinate + normals:
- For 63 features, the model achieved an impressive weighted F1-score of 0.9736 using 200 estimators, with a training time of 40.9824 s.
- For 63 features + Z coordinate, introducing the Z coordinate resulted in a slight decline in the weighted F1-score to 0.9606 with 100 estimators, while the training time increased to 44.2787 s.
- For 63 features + Z coordinate + normals, adding normals led to a further minor decrease in the weighted F1-score to 0.9598 with 100 estimators, and the training time remained comparable, at 43.9770 s.

(b) 15 features; 15 features + Z coordinate; 15 features + Z coordinate + normals:
- For 15 features, the model used 200 estimators and a training duration of 19.1061 s to reach a high weighted F1-score of 0.9710.
- For 15 features + Z coordinate, the weighted F1-score dropped to 0.9373 with 200 estimators when the Z coordinate was added, and the training time went up to 22.3032 s.
- For 15 features + Z coordinate + normals, the weighted F1-score increased slightly to 0.9439 with 50 estimators when normals were added, but the training time remained the same, at 22.7215 s.

(c) 6 features; 6 features + Z coordinate; 6 features + Z coordinate + normals:
- For 6 features, with 200 estimators and a training duration of 15.9416 s, the model obtained a weighted F1-score of 0.9296.
- For 6 features + Z coordinate, the weighted F1-score was reduced to 0.9148 with 50 estimators when the Z coordinate was included, and the training time dropped to 13.9675 s.
- For 6 features + Z coordinate + normals, the weighted F1-score increased slightly to 0.9170 with 150 estimators when normals were added, while the training time increased to 16.6206 s.

(d) 6 'ad hoc' features; 6 'ad hoc' features + Z coordinate; 6 'ad hoc' features + Z coordinate + normals:
- For 6 'ad hoc' features, the model's weighted F1-score was 0.7930 after training for 19.5866 s with 200 estimators.
- For 6 'ad hoc' features + Z coordinate, while the training duration was lowered to 15.7060 s, the weighted F1-score increased significantly to 0.9180 with 200 estimators when the Z coordinate was added.
- For 6 'ad hoc' features + Z coordinate + normals, the weighted F1-score increased slightly to 0.9200 with 100 estimators when normals were added, although the training time remained the same, at 17.0059 s.

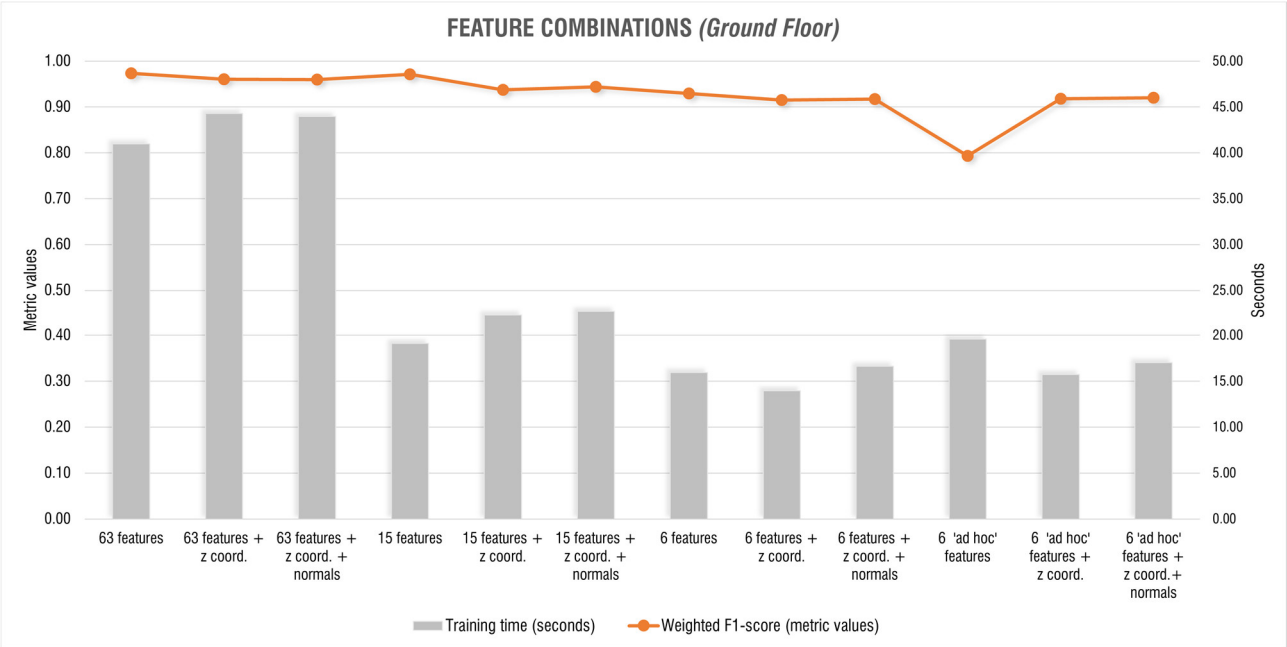

Figure 8. Analysis of the feature combinations for the ground floor: training time and weighted F1-score.

Table 5. Overview of the feature combinations for the ground floor.

Feature Combinations	No. Estimators	Training Time (s)	Weighted F1-Score
63 features	200	40.9824	0.9736
63 features + z coord.	100	44.2787	0.9606
63 features + z coord. + normals	100	43.9770	0.9598
15 features	200	19.1061	0.9710
15 features + z coord.	200	22.3032	0.9373
15 features + z coord. + normals	50	22.7215	0.9439
6 features	200	15.9416	0.9296
6 features + z coord.	50	13.9675	0.9148
6 features + z coord. + normals	150	16.6206	0.9170
6 'ad hoc' features	200	19.5866	0.7930
6 'ad hoc' features + z coord.	200	15.7060	0.9180
6 'ad hoc' features + z coord. + normals	100	17.0059	0.9200

The optimal performance for each of the four combinations is as follows when concentrating only on the F1-score and ignoring the training time, which stayed comparatively low throughout all configurations:

a. With a score of 0.9736, the 63 features combination yielded the greatest F1-score.
b. The configuration that uses only 15 features provided the greatest F1-score of all the combinations with 15 features, reaching 0.9710.
c. The greatest F1-score, 0.9296, was obtained by combining the 6 features.
d. For the 6 'ad hoc' features, the inclusion of both the Z coordinate and normals yielded the highest F1-score of 0.9200.

According to this analysis, when all features were considered together, the ground floor achieved its best result. However, as the following images demonstrate, a very good result can still be obtained by reducing the number of features, particularly when integrating 'ad hoc' features tied to the geometric size of architectural components.

The results for the first floor's designated feature combinations, as presented in Figure 9 and Table 6, are as follows:

(a) 72 features; 72 features + Z coordinate; 72 features + Z coordinate + normals:

- For 72 features, the model obtained a high weighted F1-score of 0.9772 after training for 43.5884 s using 100 estimators.
- For 72 features + Z coordinate, including the Z coordinate lowered the training time to 41.4688 s and produced a marginally lower F1-score of 0.9700 with 200 estimators.
- For 72 features + Z coordinate + normals, the F1-score increased slightly to 0.9712 with 100 estimators when normals were added, while the training time was similar at 40.1178 s.

(b) 15 features; 15 features + Z coordinate; 15 features + Z coordinate + normals:

- For 15 features, with 200 estimators and a training duration of 21.1950 s, the model obtained an acceptable weighted F1-score of 0.9092 for 15 features.
- For 15 features + Z coordinate, the training time was maintained at 21.2527 s; however, adding the Z coordinate improved the F1-score to 0.9248 with 200 estimators.
- For 15 features + Z coordinate + normals, the training time decreased to 19.7888 s and the F1-score increased slightly to 0.9230 with 100 estimators after adding normals.

(c) *6* features; 6 features + Z coordinate; 6 features + Z coordinate + normals:

- For 6 features, with 200 estimators and a training time of 18.0612 s, the model obtained a moderate F1-score of 0.8521.
- For 6 features + Z coordinate, including the Z coordinate reduced the training time to 14.1282 s and improved the F1-score to 0.8970 with 100 estimators.
- For 6 features + Z coordinate + normals, the F1-score decreased slightly to 0.8934 with 50 estimators when normals were included, while the training time was similar, at 14.5784 s.

(d) 6 'ad hoc' features; 6 'ad hoc' features + Z coordinate; 6 'ad hoc' features + Z coordinate + normals:

- For 6 'ad hoc' features, with 200 estimators and a training time of 22.2948 s, the model obtained a lower F1-score of 0.6749.
- For 6 'ad hoc' features + Z coordinate, the F1-score improved significantly to 0.9018 with 200 estimators when the Z coordinate was incorporated, and the training time dropped to 14.5593 s.
- For 6 'ad hoc' features + Z coordinate + normals, the F1-score decreased marginally to 0.8815 with 150 estimators when normals were added, while the training time went up to 17.6194 s.

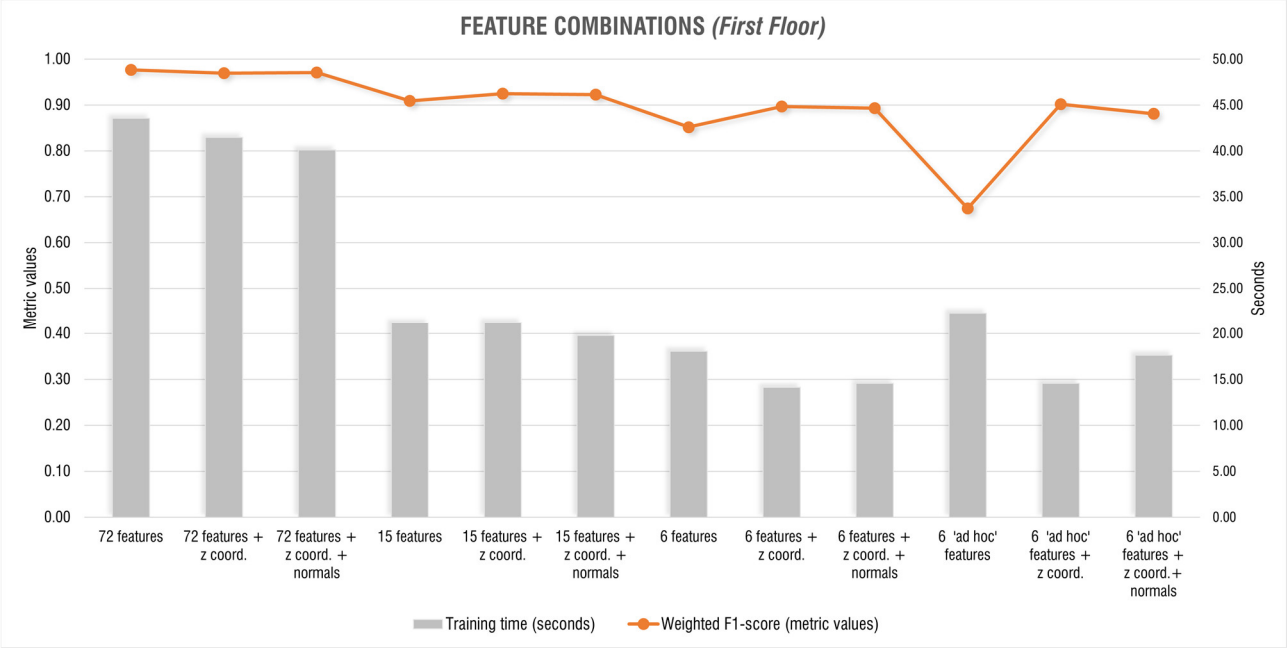

Figure 9. Analysis of the feature combinations for the first floor: training time and weighted F1-score.

Table 6. Overview of the feature combinations for the first floor.

Feature Combinations	No. Estimators	Training Time (s)	Weighted F1-Score
72 features	100	43.5884	0.9772
72 features + z coord.	200	41.4688	0.9700
72 features + z coord. + normals	100	40.1178	0.9712
15 features	200	21.1950	0.9092
15 features + z coord.	200	21.2527	0.9248
15 features + z coord. + normals	100	19.7888	0.9230
6 features	200	18.0612	0.8521
6 features + z coord.	100	14.1282	0.8970
6 features + z coord. + normals	50	14.5784	0.8934
6 'ad hoc' features	200	22.2948	0.6749
6 'ad hoc' features + z coord.	200	14.5593	0.9018
6 'ad hoc' features + z coord. + normals	150	17.6194	0.8815

The following are the best outcomes for each of the four combinations in terms of F1-score, excluding training time:

a. With a score of 0.9772, the '72 features' combination yields the highest F1-score.
b. The '15 features' combination provides the best F1-score of 0.9248.
c. The combination of '6 features + Z coordinate' produces the greatest F1-score, 0.8970.
d. The '6 ad hoc features + Z coordinate' combination brings about the best F1-score of 0.9018.

In conclusion, the first floor's results seem encouraging. The incorporation of 'ad hoc' features pertaining to the geometric dimensions of architectural components nevertheless produces an exceptionally high score, even if the combination that includes every conceivable feature produces the maximum F1-score.

3.2. Random Forest Classification on the Test Set and Performance Evaluation

Two primary modules were used for segmentation in the PyCharm environment: the first module is in charge of automatically segmenting point clouds by training a Random Forest (RF) model, the second one uses a pre-trained model to categorise point clouds, enabling effective segmentation after the model has been tuned. To guarantee a distinct division of functions, the training module's code is arranged in a modular structure.

The 'load_features_and_class' function loads feature indices and class labels from a designated text file, while the 'read_data' function is dedicated to loading labelled point cloud data. Model training is handled by the 'train_model' function, which uses a RF classifier with parameters defined by the user. The classification module, on the other hand, uses a pre-trained model to categorise point clouds from the test set. The 'load_features' function is designed to import feature indices from a specified text file, while the 'read_model' function loads a pre-trained RF model from a binary .pkl file. The primary classification process is coordinated by the main function, which integrates feature loading, model retrieval, and test data processing to perform classification. The results are subsequently saved to an output file for further analysis.

To evaluate the performance of the ML model, the 'Cross-Validation' technique was employed as a rigorous method for hyperparameter tuning. This technique was instrumental in selecting the optimal number of trees and maximum depth for the RF classifier. The training set was partitioned into five folds ($k = 5$), a choice made to mitigate overfitting during hyperparameter selection. Through this process, the model underwent five training and evaluation cycles, with each fold alternately serving as the validation set, while the remaining folds supported training.

This iterative approach provided a more reliable estimate of key evaluation metrics, including Precision, Recall, F1-Score, and Overall Accuracy. These metrics, derived from confusion matrices, offered insights into the model's predictive performance across the four most promising feature combinations per floor, as presented in the following figures.

In the confusion matrix (Table 7), the rows correspond to the actual (true) classes—manually annotated—and the columns represent the model's predicted classes.

Table 7. Example of a confusion matrix.

Class	**Actual Values**	
Predicted values	TP	FP
	FN	TN

where

TP (True Positive): instances accurately identified as positive.
FP (False Positive): instances mistakenly identified as positive rather than negative.
TN (True Negative): instances accurately identified as negative.
FN (False Negative): instances mistakenly identified as negative rather than positive.

Based on the values in the confusion matrix, the metrics Precision [Equation (13)], Recall [Equation (14)], F1-Score [Equation (15)], and Overall Accuracy [Equation (16)] are computed. The following lists their matching formulas:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{13}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{14}$$

$$\text{F1-Score} = \frac{2\ \text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{15}$$

$$\text{Overall Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{16}$$

The following images display the confusion matrices for the best four feature combinations of each floor: Figure 10 shows the results for the ground floor, while Figure 11 pertains to the first floor. Corresponding metric evaluations are summarised in Tables 8 and 9 for the ground floor, and Tables 10 and 11 for the first floor, providing detailed performance insights for each configuration.

Notable outcomes for overall accuracy were noted on the ground floor, especially when 63 features were combined, which produced an astounding accuracy of 97.38. The combination of just 15 features comes in second, with an impressive accuracy of 97.11, while the combination of 6 features achieved a reasonable accuracy of 92.98. With a score of 91.97, the combination of 6 'ad hoc' features, the Z coordinate, and the normals has the lowest accuracy, which is still sufficient for the scope.

Moving on to the first floor, the combination of 72 features produced the best accuracy, achieving a remarkable 97.73. On the other hand, the combination of 15 features and the Z coordinate showed a lower accuracy of 92.15. The combination of 6 'ad hoc' features and the Z coordinate produced the highest accuracy (89.94), whereas the combination of 6 features and the Z coordinate produced the lowest accuracy (89.37).

CLASS		*Actual values*									
		Pavement	Base	Lower Shaft	Ring	Upper Shaft	Capital	Corbel	Architrave	Frieze	Cornice
Predicted values	Pavement	11,354	61	0	0	0	0	0	1	0	0
	Base	13	4409	21	0	0	0	15	16	0	0
	Lower shaft	0	14	5846	24	395	0	0	0	0	0
	Ring	0	0	0	1240	7	0	0	0	0	0
	Upper shaft	0	0	0	19	10,991	31	0	0	0	0
	Capital	0	0	0	0	17	3522	23	0	0	0
	Corbel	0	0	0	0	0	157	6871	70	0	0
	Architrave	0	0	0	0	0	0	236	7746	55	0
	Frieze	0	0	0	0	0	0	0	367	1570	21
	Cornice	0	0	0	0	0	0	0	0	30	5731

63 Features *(GROUND FLOOR)*

CLASS		*Actual values*									
		Pavement	Base	Lower Shaft	Ring	Upper Shaft	Capital	Corbel	Architrave	Frieze	Cornice
Predicted values	Pavement	11,376	40	0	0	0	0	0	0	0	0
	Base	72	4380	19	0	0	2	1	0	0	0
	Lower shaft	0	20	5828	16	401	14	0	0	0	0
	Ring	0	0	0	1200	47	0	0	0	0	0
	Upper shaft	0	0	0	23	10,975	43	0	0	0	0
	Capital	0	0	0	0	20	3474	68	0	0	0
	Corbel	0	4	0	0	0	120	6836	138	0	0
	Architrave	0	18	0	0	0	0	279	7553	187	0
	Frieze	0	0	0	0	0	0	0	164	1757	37
	Cornice	0	0	0	0	0	0	0	0	27	5734

15 Features *(GROUND FLOOR)*

CLASS		Actual values									
		Pavement	Base	Lower Shaft	Ring	Upper Shaft	Capital	Corbel	Architrave	Frieze	Cornice
Predicted values	Pavement	11,385	31	0	0	0	0	0	0	0	0
	Base	139	4197	7	0	0	31	15	62	0	23
	Lower shaft	0	83	5021	0	1100	74	0	0	0	1
	Ring	0	0	0	1018	229	0	0	0	0	0
	Upper shaft	0	0	198	36	10,750	57	0	0	0	0
	Capital	0	37	0	0	27	3420	70	8	0	0
	Corbel	0	41	1	0	0	289	6587	180	0	0
	Architrave	0	22	0	0	0	2	421	7011	569	12
	Frieze	0	0	0	0	0	0	0	303	1607	48
	Cornice	0	2	0	0	0	0	0	15	143	5601

6 Features *(GROUND FLOOR)*

CLASS		Actual values									
		Pavement	Base	Lower Shaft	Ring	Upper Shaft	Capital	Corbel	Architrave	Frieze	Cornice
Predicted values	Pavement	11,387	29	0	0	0	0	0	0	0	0
	Base	72	4395	7	0	0	0	0	0	0	0
	Lower shaft	0	39	6240	0	0	0	0	0	0	0
	Ring	0	0	47	1200	0	0	0	0	0	0
	Upper shaft	0	0	0	38	11,003	0	0	0	0	0
	Capital	0	0	0	0	85	3477	0	0	0	0
	Corbel	0	0	0	0	0	799	6270	29	0	0
	Architrave	0	0	0	0	0	0	1822	6215	0	0
	Frieze	0	0	0	0	0	0	0	1228	730	0
	Cornice	0	0	0	0	0	0	0	0	693	5068

6 'ad hoc' Features + Z Coordinate + Normals *(GROUND FLOOR)*

Figure 10. Confusion matrix with the best combinations (63 features, 15 features, 6 features, and 6 'ad hoc' features + Z coordinate + normals) for the ground floor.

CLASS		Actual values									
		Parapet	Base	Lower Shaft	Ring	Upper Shaft	Capital	Arch	Stringcourse	Cornice	Corner Column
Predicted values	Parapet	16,013	1	0	0	0	0	59	3	0	74
	Base	65	143	6	0	0	0	0	0	0	0
	Lower shaft	0	21	1109	21	0	0	3	0	0	13
	Ring	0	0	20	162	30	0	2	0	0	13
	Upper shaft	0	0	0	25	2024	11	2	0	0	90
	Capital	0	0	0	0	10	535	64	0	0	18
	Arch	3	0	0	0	0	41	8176	78	0	24
	Stringcourse	0	0	0	0	0	0	172	4076	37	26
	Cornice	0	0	0	0	0	0	0	85	12,127	0
	Corner column	44	4	36	5	30	6	83	21	0	9289

72 Features *(FIRST FLOOR)*

CLASS		Actual values									
		Parapet	Base	Lower Shaft	Ring	Upper Shaft	Capital	Arch	Stringcourse	Cornice	Corner Column
Predicted values	Parapet	14,618	551	351	0	0	0	40	0	0	590
	Base	0	150	63	0	0	0	0	0	0	1
	Lower shaft	0	32	1096	23	1	0	0	0	0	15
	Ring	0	0	28	123	28	0	0	0	0	48
	Upper shaft	0	0	0	18	1650	76	0	0	0	408
	Capital	0	0	0	0	0	341	270	0	0	16
	Arch	0	0	0	0	0	0	7492	807	0	23
	Stringcourse	0	0	0	0	0	0	20	3996	290	5
	Cornice	0	0	0	0	0	0	0	0	12,212	0
	Corner column	0	7	70	5	114	0	146	261	0	8915

15 Features + Z Coordinate *(FIRST FLOOR)*

CLASS		Actual values									
		Parapet	Base	Lower Shaft	Ring	Upper Shaft	Capital	Arch	Stringcourse	Cornice	Corner Column
Predicted values	Parapet	13,922	329	508	0	0	0	0	0	0	1391
	Base	0	81	63	0	0	0	0	0	0	70
	Lower shaft	0	17	1056	39	0	0	0	0	0	55
	Ring	0	0	18	121	55	0	0	0	0	33
	Upper shaft	0	0	0	3	1821	116	0	0	0	212
	Capital	0	0	0	0	0	342	259	0	0	26
	Arch	0	0	0	0	0	0	6963	1258	0	101
	Stringcourse	0	0	0	0	0	0	42	3994	275	0
	Cornice	0	0	0	0	0	0	0	0	12,212	0
	Corner column	0	0	70	31	209	63	299	292	0	8554

6 Features + Z Coordinate *(FIRST FLOOR)*

CLASS		Actual values									
		Parapet	Base	Lower Shaft	Ring	Upper Shaft	Capital	Arch	Stringcourse	Cornice	Corner Column
Predicted values	Parapet	14,959	528	83	2	2	90	2	0	0	484
	Base	0	138	68	0	0	0	0	0	0	8
	Lower shaft	0	0	849	55	37	0	0	0	0	226
	Ring	0	0	1	87	116	0	0	0	0	23
	Upper shaft	0	0	0	0	1687	57	0	0	0	408
	Capital	0	0	0	0	4	345	242	0	0	36
	Arch	0	0	0	0	0	1	7151	713	0	457
	Stringcourse	0	0	0	0	0	0	433	3470	341	67
	Cornice	0	0	0	0	0	0	0	0	12,212	0
	Corner column	0	9	85	2	138	10	636	157	0	8481

6 'ad hoc' Features + Z Coordinate *(FIRST FLOOR)*

Figure 11. Confusion matrix with the best combinations (72 features, 15 features + Z coordinate, 6 features + Z coordinate, and 6 'ad hoc' features + Z coordinate) for the first floor.

Table 8. Performance metrics for the ground floor using 63, and 15 features.

PERFORMANCE METRICS (63 Features)				PERFORMANCE METRICS (15 Features)			
Precision	**Recall**	**F1-Score**		**Precision**	**Recall**	**F1-Score**	
99.46	99.89	99.67		99.65	99.37	99.51	
98.55	98.33	98.44		97.90	98.16	98.03	
93.10	99.64	96.26		92.82	99.68	96.12	
99.44	96.65	98.02		96.23	96.85	96.54	
99.55	96.33	97.91		99.40	95.91	97.62	
98.88	94.93	96.86		97.53	95.10	96.30	
96.80	96.17	96.48		96.31	95.16	95.73	
96.38	94.46	95.41		93.98	96.16	95.05	
80.18	94.86	86.91		89.73	89.14	89.44	
99.48	99.63	99.56		99.53	99.36	99.45	
Simple Average			Overall Accuracy	Simple Average			Overall Accuracy
96.18	97.09	96.55	97.38	96.31	96.49	96.38	97.11

Table 9. Performance metrics for the ground floor using 6 features and 6 'ad hoc' features + Z coordinate + normals.

PERFORMANCE METRICS (6 Features)				PERFORMANCE METRICS (6 'Ad Hoc' Features + Z Coordinate + Normals)			
Precision	**Recall**	**F1-Score**		**Precision**	**Recall**	**F1-Score**	
99.73	98.79	99.26		99.75	99.37	99.56	
93.81	95.11	94.45		98.23	98.48	98.36	
79.96	96.06	87.28		99.38	99.14	99.26	
81.64	96.58	88.48		96.23	96.93	96.58	
97.36	88.80	92.88		99.66	99.23	99.44	
96.01	88.30	92.00		97.61	81.31	88.72	
92.80	92.87	92.83		88.33	77.48	82.55	
87.23	92.51	89.79		77.33	83.18	80.15	
82.07	69.30	75.15		37.28	51.30	43.18	
97.22	98.52	97.87		87.97	100.00	93.60	
Simple Average			Overall Accuracy	Simple Average			Overall Accuracy
90.78	91.68	91.00	92.98	88.18	88.64	88.14	91.97

Table 10. Performance metrics for the first floor using 72 features and 15 features + Z coordinate.

PERFORMANCE METRICS (72 Features)				PERFORMANCE METRICS (15 Features + Z Coordinate)			
Precision	**Recall**	**F1-Score**		**Precision**	**Recall**	**F1-Score**	
99.15	99.31	99.23		90.51	100.00	95.02	
66.82	84.62	74.67		70.09	20.27	31.45	
95.03	94.71	94.87		93.92	68.16	78.99	
71.37	76.06	73.64		54.19	72.78	62.12	
94.05	96.66	95.34		76.67	92.02	83.65	
85.33	90.22	87.70		54.39	81.77	65.33	
98.25	95.50	96.85		90.03	94.03	91.98	
94.55	95.61	95.08		92.69	78.91	85.25	
99.30	99.70	99.50		100.00	97.68	98.83	
97.59	97.30	97.45		93.66	88.96	91.25	
Simple Average			Overall Accuracy	Simple Average			Overall Accuracy
90.14	92.97	91.43	97.73	81.62	79.46	78.39	92.15

Table 11. Performance metrics for the first floor using 6 features + Z coordinate and 6 'ad hoc' features + Z coordinate.

PERFORMANCE METRICS (6 Features + Z Coordinate)			PERFORMANCE METRICS (6 'Ad Hoc' Features + Z Coordinate)		
Precision	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**
86.20	100.00	92.59	92.63	100.00	96.17
37.85	18.97	25.27	64.49	20.44	31.05
90.49	61.57	73.28	72.75	78.18	75.37
53.30	62.37	57.48	38.33	59.59	46.65
84.62	87.34	85.96	78.39	85.03	81.58
54.55	65.64	59.58	55.02	68.59	61.06
83.67	92.07	87.67	85.93	84.49	85.20
92.65	72.04	81.06	80.49	79.95	80.22
100.00	97.80	98.89	100.00	97.28	98.62

89.87	81.92	85.71		89.10	83.23	86.07	
Simple Average			Overall Accuracy	Simple Average			Overall Accuracy
77.32	73.97	74.75	89.37	75.71	75.68	74.20	89.94

3.3. Semantic Point Cloud

The final point cloud rankings based on Machine Learning are shown in Figure 12, which represents the best four combinations of features for each floor.

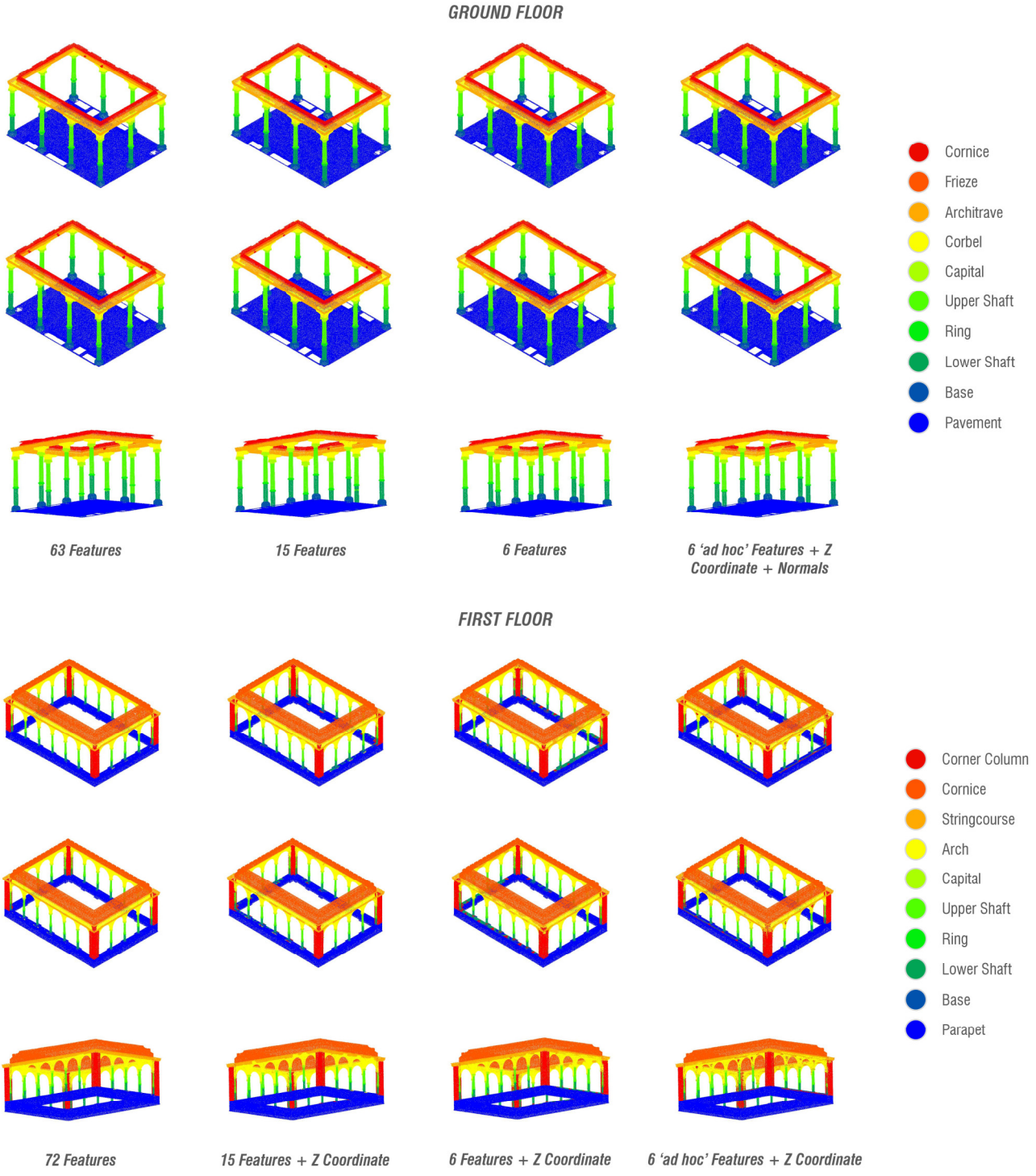

Figure 12. Semantic point cloud of the ground floor (on the **top**) and the first floor (on the **bottom**), with the best four combinations for each floor.

Applying each combination to the two different floors produces quite important outcomes. From a graphical standpoint, an interesting phenomenon that suggests a possible reversal of roles arises when prioritising combinations with more attributes based on the confusion matrix results.

In this regard, it appears visible that 'ad hoc' features, in conjunction with the Z coordinate and normals, provide almost flawless semantic discretisation for column-related components. The ring that divides the shaft's two portions may be clearly identified in this arrangement. However, there is some semantic dissonance in the upper sections, where the arch is occasionally confused with the first-floor corner column and the architrave with the ground-floor corbel. Combinations of 15 and 6 features produced intermediate findings, showing multiple semantic differentiating mistakes, mostly related to the ground-floor lower shaft and the first-floor parapet and stringcourse.

Aside from the relatively minor errors arising from the different feature combinations, the results from the discretisation of point cloud geometries clearly demonstrate the value of this approach. A thorough, preliminary study focused on optimising point clouds through semantic enrichment can significantly facilitate the 2D and 3D digital reconstruction of Cultural and Architectural Heritage. In this context, the segmentation of point clouds into distinct and easily identifiable architectural components enables the creation of detailed and accurate representations of the elements, which can be used in CAD environments for extracting construction profiles and generating sections, details, plans, and more.

Moreover, this approach proves particularly beneficial in the creation of 3D or parametric BIM models. The semantic point cloud offers a solid foundation for precise modelling, enabling the target selection of specific layers corresponding to distinct portions of the point cloud to be visualised during the modelling process. This is particularly pertinent in the context of digitisation for the conservation and enhancement of heritage, where large datasets present significant challenges, not only due to their computational demands but also because of the often irregular and complex geometries of buildings, which are inherently difficult to manage.

4. Conclusions

The integration of Artificial Intelligence (AI) within the realm of Architectural Heritage (AH) marks a transformative advancement in the automatic segmentation and classification of point clouds through Machine Learning, offering significant potential for the preservation, analysis, and understanding of historical structures.

This study has achieved meticulous and accurate representations of a distinguished Renaissance palace in Zaragoza, yielding remarkable results. The AI-driven methodology has enabled the differentiation of geometrically intricate architectural and structural components, facilitating a deeper understanding and digital documentation of minute details that are often overlooked by the human eye. Notably, the automatic classification of point cloud segments was accomplished in mere seconds, demonstrating a significant optimisation of processing times.

Despite these accomplishments, several notable challenges have been encountered, particularly the necessity for a comprehensive and representative training dataset, which presents a critical constraint. The manual segmentation of point clouds is inherently time-consuming, dependent on the operator's desired level of detail and precision, and consequently impacts the accuracy of the final classification.

Nevertheless, the various combinations of geometric features explored here have yielded promising results, even with a limited training sample. In this instance, 'ad hoc' features were selected based on the geometric proportions of the architectural elements designated for segmentation. While these results were satisfactory, they did not match the performance levels achieved with combinations that included more extensive training data. The incorporation of normal coordinates and the Zcoordinate has substantially

enhanced the overall segmentation analysis, significantly improving the accuracy and efficacy of the classification process.

It is essential to recognise that while the demonstrated automated methodology is generalisable within the framework of AH, the training data are specific to each study examined. This specificity could affect the broader applicability of the results. Addressing this gap may require the implementation of more sophisticated models; the integration of multi-sensor data; and the exploration of techniques, such as Segment Anything Model (SAM), Generative Adversarial Networks (GANs), Deep Learning, and more, to further augment system performance. Consequently, the development of more inclusive and diverse datasets will facilitate the management of a wider array of historical contexts and enhance the generalisation of models.

Innovation in this field holds considerable potential to advance the understanding of historical architectural works and to create new avenues for heritage preservation, particularly when integrated into virtualisation frameworks such as Computer-Aided Design (CAD) or Heritage Building Information Modelling (HBIM). This offers valuable opportunities for future developments, particularly in facilitating automation across the operational stages of three-dimensional modelling.

In this context, the present study functions as a foundational step, rather than an endpoint, providing a basis for further refinement. The segmentation of point clouds, in conjunction with the continuous evolution of segmentation and classification methodologies, can be extended to encompass semantic enrichment, with particular emphasis on HBIM contexts. Moreover, the integration of the Industry Foundation Classes (IFC) format enables the expansion of these methodologies, facilitating applications such as temporal analysis, while also ensuring the efficient management of data throughout the entire lifecycle of built heritage. Semantic point clouds can be seamlessly incorporated into BIM workflows, improving the precision of architectural modelling and facilitating key tasks such as verifying as-built models, detecting clashes, and managing assets. This integration not only optimises the design process but also significantly improves the maintenance and conservation of historical buildings.

Author Contributions: Conceptualisation, M.B., L.A.-H. and C.V.; methdology, M.B., L.A.-H. and C.V.; software, M.B.; validation, M.B.; formal analysis, M.B.; investigation, M.B., L.A.-H. and C.V.; resources, M.B., L.A.-H. and C.V.; data curation, M.B.; writing—original draft preparation, M.B.; writing—review and editing, M.B., L.A.-H. and C.V.; visualisation, M.B.; supervision, L.A.-H. and C.V.; project administration, L.A.-H. and C.V; funding acquisition, M.B., L.A.-H. and C.V. All authors have read and agreed to the published version of the manuscript.

Funding: The research, part of the corresponding author's PhD project, was conducted in collaboration between the Polytechnic University of Bari and the University of Zaragoza, with co-funding from the European Union under the POR Puglia FESR-FSE 2014/2020 programme.

Data Availability Statement: The datasets presented in this article are not readily available because they are part of an ongoing study.

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

References

1. Balzani, M.; Maietti, F.; Mugayar Kühl, B. Point cloud analysis for conservation and enhancement of modernist architecture. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2017**, *42*, 71–77. https://doi.org/10.5194/isprs-archives-XLII-2-W3-71-2017.
2. Paschalidou, E.; Fafet, C.; Milios, L. A Strong Sustainability Framework for Digital Preservation of Cultural Heritage: Introducing the Eco-Sufficiency Perspective. *Heritage* **2022**, *5*, 1066–1088. https://doi.org/10.3390/heritage5020058.
3. Tan, P. Image-Based Modeling. In *Computer Vision*; Springer: Cham, Switzerland, 2020; pp. 1–4, ISBN 978-3-030-03243-2.
4. Aicardi, I.; Chiabrando, F.; Lingua, A.; Noardo, F. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach. *J. Cult. Herit.* **2018**, *32*, 257–266. https://doi.org/10.1016/j.culher.2017.11.006.
5. El-Din Fawzy, H. 3D laser scanning and close-range photogrammetry for buildings documentation: A hybrid technique towards a better accuracy. *Alexandria Eng. J.* **2019**, *58*, 1191–1204. https://doi.org/10.1016/j.aej.2019.10.003.

6. Remondino, F. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sens.* **2011**, *3*, 1104–1138. https://doi.org/10.3390/rs3061104.
7. Hassan, A.T.; Fritsch, D. Integration of Laser Scanning and Photogrammetry in 3D/4D Cultural Heritage Preservation—A Review. *Int. J. Appl. Sci. Technol.* **2019**, *9*, 76–91. https://doi.org/10.30845/ijast.v9n4p9.
8. Yang, S.; Xu, S.; Huang, W. 3D Point Cloud for Cultural Heritage: A Scientometric Survey. *Remote Sens.* **2022**, *14*, 5542. https://doi.org/10.3390/rs14215542.
9. Beraldin, J.-A. Integration of laser scanning and close-range photogrammetry. The last decade and beyond. In Proceedings of the XXth ISPRS Congress, Commission VII, Istanbul, Turkey, 12–23 July 2004.
10. Pierdicca, R.; Gorgoglione, L.; Martuscelli, M.; Usmanov, S. Challenging architectures: An integrated and multipurpose survey for the complete mapping of the Emir Palace in Kogon (Uzbekistan). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 1225–1232. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-1225-2023.
11. Mazzacca, G.; Karami, A.; Rigon, S.; Farella, E.M.; Trybala, P.; Remondino, F. NeRF for Heritage 3D Reconstruction. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 1051–1058. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-1051-2023.
12. Croce, V.; Caroti, G.; De Luca, L.; Piemonte, A.; Veron, P. Neural Radiance Fields (NeRF): Review and potential applications to Digital Cultural Heritage. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 453–460. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-453-2023.
13. Fan, J.; Ma, L.; Zou, Z. A registration method of point cloud to CAD model based on edge matching. *Optik* **2020**, *219*, 165223. https://doi.org/10.1016/j.ijleo.2020.165223.
14. Denayer, M.; De Winter, J.; Bernardes, E.; Vanderborght, B.; Verstraten, T. Comparison of Point Cloud Registration Techniques on Scanned Physical Objects. *Sensors* **2024**, *24*, 2142. https://doi.org/10.3390/s24072142.
15. Hichri, N.; Stefani, C.; De Luca, L.; Veron, P.; Hamon, G. From point cloud to BIM: A survey of existing approaches. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *40*, 343–348. https://doi.org/10.5194/isprsarchives-XL-5-W2-343-2013, 2013.
16. Badenko, V.; Fedotov, A.; Zotov, D.; Lytkin, S.; Volgin, D.; Garg, R.D.; Min, L. Scan-to-BIM methodology adapted for different application. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2019**, *42*, 1–7. https://doi.org/10.5194/isprs-archives-XLII-5-W2-1-2019.
17. Anil, E.B.; Tang, P.; Akinci, B.; Huber, D. Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data. *Autom. Constr.* **2013**, *35*, 507–516. https://doi.org/10.1016/j.autcon.2013.06.003.
18. Antón, D.; Medjdoub, B.; Shrahily, R.; Moyano, J. Accuracy evaluation of the semi-automatic 3D modeling for historical building information models. *Int. J. Archit. Herit.* **2018**, *12*, 790–805. https://doi.org/10.1080/15583058.2017.1415391.
19. Radanovic, M.; Khoshelham, K.; Fraser, C. Geometric accuracy and semantic richness in heritage BIM: A review. *Digit. Appl. Archaeol. Cult. Herit.* **2020**, *19*, e00166. https://doi.org/10.1016/j.daach.2020.e00166.
20. Rocha, G.; Mateus, L.; Fernández, J.; Ferreira, V. A Scan-to-BIM Methodology Applied to Heritage Buildings. *Heritage* **2020**, *3*, 47–67. https://doi.org/10.3390/heritage3010004.
21. Alshawabkeh, Y.; Baik, A.; Miky, Y. Integration of Laser Scanner and Photogrammetry for Heritage BIM Enhancement. *ISPRS Int. J. Geo-Information* **2021**, *10*, 316. https://doi.org/10.3390/ijgi10050316.
22. Verdoscia, C.; Buldo, M.; Musicco, A.; Tavolare, R. Technological Paradigms for Cultural Heritage. A Scan To BIM Methodology for the Description of Historical Architecture. In *New Technologies in Building and Construction. Lecture Notes in Civil Engineering*; Bienvenido-Huertas, D., Moyano-Campos, J., Eds.; Springer: Singapore, 2022; pp. 187–205, ISBN 978-981-19-1893-3.
23. Xie, Y.; Tian, J.; Zhu, X.X. Linking Points with Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 38–59. https://doi.org/10.1109/MGRS.2019.2937630.
24. Vieira, M.; Shimada, K. Surface mesh segmentation and smooth surface extraction through region growing. *Comput. Aided Geom. Des.* **2005**, *22*, 771–792. https://doi.org/10.1016/j.cagd.2005.03.006.
25. Lavoué, G.; Dupont, F.; Baskurt, A. A new CAD mesh segmentation method, based on curvature tensor analysis. *Comput. Aided Des.* **2005**, *37*, 975–987. https://doi.org/10.1016/j.cad.2004.09.001.
26. Gaber, J.A.; Youssef, S.M.; Fathalla, K.M. The Role of Artificial Intelligence and Machine Learning in preserving Cultural Heritage and Art Works via Virtual Restoration. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *10*, 185–190. https://doi.org/10.5194/isprs-annals-X-1-W1-2023-185-2023.
27. Ge, C. The review of AI and cultural heritage protection-Taking the whole process of cultural heritage protection as an example. *Appl. Comput. Eng.* **2024**, *71*, 137–143.
28. Foka, A.; Griffin, G. AI, Cultural Heritage, and Bias: Some Key Queries That Arise from the Use of GenAI. *Heritage* **2024**, *7*, 6125–6136. https://doi.org/10.3390/heritage7110287.
29. Tiribelli, S.; Pansoni, S.; Frontoni, E.; Giovanola, B. Ethics of Artificial Intelligence for Cultural Heritage: Opportunities and Challenges. *IEEE Trans. Technol. Soc.* **2024**, *5*, 293–305. https://doi.org/10.1109/TTS.2024.3432407.
30. Pansoni, S.; Tiribelli, S.; Paolanti, M.; Di Stefano, F.; Frontoni, E.; Malinverni, E.S.; Giovanola, B. Artificial Intelligence and Cultural Heritage: Design and Assessment of an Ethical Framework. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 1149–1155. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-1149-2023.
31. Gîrbacia, F. An Analysis of Research Trends for Using Artificial Intelligence in Cultural Heritage. *Electronics* **2024**, *13*, 3738. https://doi.org/10.3390/electronics13183738.

32. Patrucco, G.; Setragno, F. Enhancing Automation of Heritage Processes: Generation of Artificial training Datasets from Photogrammetric 3D Models. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 1181–1187. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-1181-2023.
33. Brodu, N.; Lague, D. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 121–134. https://doi.org/10.1016/j.isprsjprs.2012.01.006.
34. Russo, M.; Grilli, E.; Remondino, F.; Teruggi, S.; Fassi, F. Machine Learning for Cultural Heritage Classification. *Represent. Challenges. Augment. Real. Artif. Intell. Cult. Herit. Innov. Des. Domain* **2021**, 209–214. https://doi.org/10.3280/oa-686.33.
35. Buldo, M.; Agustín-Hernández, L.; Verdoscia, C.; Tavolare, R. A Scan-to-BIM workflow proposal for Cultural Heritage. Automatic point cloud segmentation and parametric-adaptive modelling of vaulted systems. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 333–340. https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-333-2023.
36. Croce, V.; Caroti, G.; Piemonte, A.; Bevilacqua, M.G. From survey to semantic representation for Cultural Heritage: The 3D modelling of recurring architectural elements. *Acta IMEKO* **2021**, *10*, 98–108. https://doi.org/10.21014/ACTA_IMEKO.V10I1.842.
37. Fiorucci, M.; Khoroshiltseva, M.; Pontil, M.; Traviglia, A.; Del Bue, A.; James, S. Machine Learning for Cultural Heritage: A Survey. *Pattern Recognit. Lett.* **2020**, *133*, 102–108. https://doi.org/10.1016/j.patrec.2020.02.017.
38. Zhao, J.; Hua, X.; Yang, J.; Yin, L.; Liu, Z.; Wang, X. A Review of Point Cloud Segmentation of Architectural Cultural Heritage. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *X-1/W1-202*, 247–254. https://doi.org/10.5194/isprs-annals-x-1-w1-2023-247-2023.
39. Yang, S.; Hou, M.; Li, S. Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review. *Remote Sens.* **2023**, *15*, 548. https://doi.org/10.3390/rs15030548.
40. Kutlu, H.; Brucker, F.; Kallendrusch, B.; Santos, P.; Fellner, D.W. AI Based Image Segmentation of Cultural Heritage Objects used for Multi-View Stereo 3D Reconstructions. In Proceedings of the GCH 2023—Eurographics Workshop on Graphics and Cultural Heritage, Lecce, Italy, 4–6 September 2023; pp. 75–79.
41. Pellis, E.; Murtiyoso, A.; Masiero, A.; Tucci, G.; Betti, M.; Grussenmeyer, P. 2D to 3D Label Propagation for the Semantic Segmentation of Heritage Building Point Clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2022**, *43*, 861–867. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-861-2022.
42. Malinverni, E.S.; Pierdicca, R.; Sturari, M.; Colosi, F.; Orazi, R. Documentation and detection of colour changes of bas relieves using close range photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2017**, *42*, 203–210. https://doi.org/10.5194/isprs-Archives-XLII-5-W1-203-2017.
43. Vorobel, R.; Ivasenko, I.; Berehulyak, O.; Mandzii, T. Segmentation of rust defects on painted steel surfaces by intelligent image analysis. *Autom. Constr.* **2021**, *123*, 103515. https://doi.org/10.1016/j.autcon.2020.103515.
44. Abgaz, Y.; Souza, R.R.; Methuku, J.; Koch, G.; Dorn, A. A Methodology for Semantic Enrichment of Cultural Heritage Images Using Artificial Intelligence Technologies. *Technol. J Imaging* **2021**, *7*, 121. https://doi.org/10.3390/jimaging7080121.
45. Luo, X.; Wang, X.; Tan, X. Semantically Enriched Presentation for Cultural Heritage Image: A POI-Based Perspective. In Proceedings of the 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Champaign, IL, USA, 2–6 June 2019; pp. 410–411.
46. Grilli, E.; Dininno, D.; Petrucci, G.; Remondino, F. From 2D to 3D Supervised Segmentation and Classification for Cultural Heritage Applications. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2018**, *42*, 399–406. https://doi.org/10.5194/isprs-archives-XLII-2-399-2018.
47. Fatiguso, F.; Buldo, M. Complesso della SS. Trinità di Venosa (PZ). In *La Diagnostica per Gli Edifici Storici: Metodi non Distruttivi e Tecnologie Innovative per la Valutazione e il Controllo. Collana Architettura Sostenibile/Culture Costruttive per il Recupero Sostenibile*; De Fino, M., Fatiguso, F., Eds.; EdicomEdizioni: Monfalcone (GO), Italy, 2020; pp. 169–180, ISBN 978-88-96386-92-7.
48. Quintilla-Castán, M.; Agustín-Hernández, L. 3D survey and virtual reconstruction of heritage. The case study of the City Council and Lonja of Alcañiz. *Vitruvio* **2021**, *6*, 12–25. https://doi.org/10.4995/vitruvio-ijats.2021.16567.
49. 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.* **2020**, *12*, 2598. https://doi.org/10.3390/RS12162598.
50. 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-Information* **2020**, *9*, 535. https://doi.org/10.3390/ijgi9090535.
51. Hu, X.; Yuan, Y. Deep-Learning-Based Classification for DTM Extraction from ALS Point Cloud. *Remote Sens.* **2016**, *8*, 730. https://doi.org/10.3390/rs8090730.
52. Malinverni, E.S.; Pierdicca, R.; Paolanti, M.; Martini, M.; Morbidoni, C.; Matrone, F.; Lingua, A. Deep Learning for Semantic Segmentation of 3D Point Cloud. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2019**, *42*, 735–742. https://doi.org/10.5194/isprs-archives-XLII-2-W15-735-2019.
53. Battini, C.; Ferretti, U.; De Angelis, G.; Pierdicca, R.; Paolanti, M.; Quattrini, R. Automatic generation of synthetic heritage point clouds: Analysis and segmentation based on shape grammar for historical vaults. *J. Cult. Herit.* **2024**, *66*, 37–47. https://doi.org/10.1016/j.culher.2023.10.003.
54. Pellis, E.; Masiero, A.; Grussenmeyer, P.; Betti, M.; Tucci, G. Synthetic data generation and testing for the semantic segmentation of heritage buildings. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2023**, *48*, 1189–1196. https://doi.org/10.5194/isprs-Archives-XLVIII-M-2-2023-1189-2023.

55. Siountri, K.; Anagnostopoulos, C.N. The Classification of Cultural Heritage Buildings in Athens Using Deep Learning Techniques. *Heritage* **2023**, *6*, 3673–3705. https://doi.org/10.3390/heritage6040195.
56. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.
57. Weinmann, M.; Jutzi, B.; Mallet, C. Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *2*, 313–318. https://doi.org/10.5194/isprsannals-II-5-W2-313-2013.
58. Rodríguez-Gonzálvez, P.; Rodríguez-Martín, M. Weld Bead Detection Based on 3D Geometric Features and Machine Learning Approaches. *IEEE Access* **2019**, *7*, 14714–14727. https://doi.org/10.1109/ACCESS.2019.2891367.
59. Chehata, N.; Guo, L.; Mallet, C. Airborne LiDAR feature selection for urban classification using random forests. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2009**, *38*, 207–212.
60. Dahl, V.A. Structure Tensor. Computation, Visualization, and Application. In *Quantitative Analysis of Tomographic Volumes*; 2019; pp. 1–14. Available online: https://people.compute.dtu.dk/vand/notes/ST_intro.pdf (accessed on 24 October 2023)
61. Harshit, H.; Kushwaha, S.K.P.; Jain, K. Geometric Features Interpretation of Photogrammetric Point Cloud from Unmanned Aerial Vehicle. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *10*, 83–88. https://doi.org/10.5194/isprs-annals-X-4-W2-2022-83-2022.
62. Getty, J.P. Art & Architecture Thesaurus® Online. Available online: https://www.getty.edu/research/tools/vocabularies/aat/ (accessed on 24 October 2023).
63. Agarwal, A.; Kenney, A.M.; Tan, Y.S.; Tang, T.M.; Yu, B. MDI+: A Flexible Random Forest-Based Feature Importance Framework. *arXiv* **2023**, arXiv:2307.01932. https://doi.org/10.48550/arXiv.2307.01932.
64. Scornet, E. Trees, forests, and impurity-based variable importance. *arXiv* **2021**, arXiv:2001.04295. https://doi.org/10.48550/arXiv.2001.04295.