The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-9-2025
30th CIPA Symposium “Heritage Conservation from Bits:
From Digital Documentation to Data-driven Heritage Conservation”, 25–29 August 2025, Seoul, Republic of Korea

Optimized Approach to Feature Selection for Semantic Segmentation Using Random Forest

Giuseppe Antuono [1], Valeria Cera[2], Daniela Ciarlo[3]

[1] Dept. of Civil, Building and Environmental Engineering, University of Naples Federico II, Naples, Italy - giuseppe.antuono@unina.it
[2] Dept. of Architecture, University of Naples Federico II, Naples, Italy - valeria.cera@unina.it
[3] Dept. of Civil, Building and Environmental Engineering, University of Naples Federico II, Naples, Italy - danielaciarlo.dc@gmail.com

Keywords: AI, Semantic Segmentation, Points Clouds Classification, Randon Forest, Features Importance, Machine Learning.

Abstract

The study proposes an optimized approach for feature selection in the semantic segmentation of point clouds within the architectural domain of cultural heritage, with a specific focus on historical monastic architecture. The goal is to enhance the automatic recognition and classification of architectural elements using the Random Forest algorithm, by reducing classifier dependency and increasing the model’s generalization capability. The developed method is based on a multiscale statistical selection of features, through p-value analysis and the optimization of influence radii, fully automating the process within a Python environment. The method was tested on a TLS point cloud dataset specifically built for Franciscan cloisters in the Campania region, segmented into ten architectural classes. The new approach builds upon the existing but implemented RF4PCC model, against which it was compared, showing significant improvements in the classification of minority classes, thanks to the adoption of the *class_weight=”balanced”* parameter and the expansion of the dataset. The analysis of *feature_importances_* revealed biases related to class imbalance, which were addressed through regularization strategies and complexity control of the decision trees. Experimental results show an increase in the macro F1-score and greater fairness in class classification. The proposed approach proves effective for applications in the cultural heritage field, offering an interpretable, efficient, and adaptable method for complex architectural contexts.

1. Introduction

The evolution of studies on point cloud segmentation using Artificial Intelligence (AI) algorithms has opened new opportunities in the field of cultural heritage (CH), enabling the interpretation and classification of architectural elements characterized by complex, often non-standardizable geometries (Cao et al., 2022; Zhao et al., 2023; Yang et al., 2023). One of the main challenges in developing AI-based approaches for semantic segmentation lies in the difficulty of collecting a sufficiently large dataset - especially in the context of Deep Learning (DL) - to train the algorithm and define a reliable predictive model (Terruggi et al., 2020). Considering these limitations, the development of Machine Learning (ML) approaches has led to promising results in recent years, even when working with small-scale datasets. In the CH domain, studies frequently reference the use of the Random Forest (RF) algorithm, where the computation and selection of features play a critical role in optimizing the classifier’s predictive accuracy (Pierdicca et al., 2020). This study develops an innovative methodological approach to overcome the limitations of current feature selection systems by introducing an external statistical validation framework that replaces traditional impurity-reduction-based methods. The primary objective is to fully optimize and automate the selection and computation of geometric features, significantly reducing classifier dependency and enhancing the model’s generalization capability across architectural contexts beyond the training data. This approach improves classification performance and computational efficiency, with particular focus on identifying minority architectural elements such as moldings, openings, and decorative details, which, despite their numerical underrepresentation, are crucial for the stylistic characterization of historic buildings. The analysis was conducted on point cloud datasets acquired through Terrestrial Laser Scanning, concentrating on the cloister architecture of Franciscan religious complexes. An advanced feature selection technique was developed based on multi-radius statistical analyses and p-value evaluations for each feature relative to target classes, employing triangular significance matrices to automatically determine the optimal radius for each feature and maximize class separability. The entire approach, implemented in Python, provides an interpretable, efficient, and adaptable method for complex architectural scenarios, with direct applications in digital documentation and cultural heritage analysis.

2. State of the Art

In recent years, the use of AI techniques - particularly ML and DL - has significantly enhanced automatic classification and semantic segmentation of point clouds for cultural heritage documentation (Xie et al., 2020; Gaber et al., 2023; Gîrbacia, 2024). Supervised ML techniques can perform well even with small datasets; however, the user’s role remains central in defining the geometric or radiometric features. The RF algorithm has shown good performance on limited datasets, provided that features are selected based on the characteristics of the target classes (Ni et al., 2017). Adding relevant features improves segmentation accuracy (Atik and Duran, 2022), but selecting the most appropriate ones and determining their influence radius remains crucial (Weinmann et al., 2013; Buldo et al., 2024), especially for model generalization (Grilli and Remondino, 2020). Targeted feature selection helps reduce model complexity and enhances predictive performance (Guyon and Elisseeff, 2003; Moyano et al., 2024). While wrapper methods often yield better results than filter-based methods, they carry a higher risk of overfitting, like embedded methods. Effective feature selection requires a balance between representativeness and generalizability (Harshit et al., 2022).

A critical issue is class imbalance, frequently encountered in CH datasets: some classes (e.g., walls) are overrepresented, while others (e.g., moldings) are underrepresented. This imbalance affects the assessment of feature importances in RF: features associated with majority classes tend to be overestimated, while those related to minority classes are undervalued, making accurate classification of the latter more difficult (Gu et al., 2022). RF calculates feature importance by summing the impurity reductions in the splits where each feature is used. However, imbalanced datasets distort this process: decision trees are more likely to generate splits that favor dominant classes, penalizing minority ones. Lin and Nguyen (2020) proposed oversampling and undersampling techniques to improve minority class prediction, though these methods often face practical limitations. Gu et al. (2022) also reported promising results using artificial balancing techniques, offering a solid methodological foundation for applications in CH. Nonetheless, parameter optimization of RF for architectural scenarios remains underexplored. This study introduces a new approach to feature selection that reduces classifier dependency and enhances generalization by adopting cost-sensitive learning strategies (Chen et al., 2004) and applying balancing techniques such as *class_weight = "balanced"*.

Finally, the proposed work introduces a methodological framework that integrates: - external statistical validation; - advanced parameter optimization; - class balancing techniques.

The goal is to develop a robust, interpretable, and generalizable system for semantic segmentation of point clouds within monastic architectural heritage.

3. Methodology and Materials

The proposed methodology combines multiscale statistical analyses with ML techniques to create an automated system for selecting geometric features in architectural 3D models. The innovative aspect lies in replacing traditional feature selection methods based on impurity reduction with an external statistical validation system that employs robust tests to evaluate differences between class distributions. This approach ensures an objective and reproducible identification of the optimal feature-radius combinations, eliminating interpretative subjectivity and enhancing the reliability of the predictive model's generalization. Specifically, the methodological workflow includes the following steps:

1. Comparison between classifications obtained using the predictive model from Random Forest for Point Cloud Classification (RF4PCC) (3DOM-FBK/RF4PCC 2024) and those derived from its implementation (RF4PCC - implemented), in terms of feature and radius selection using accuracy metrics and statistical significance criteria.
2. Analysis of feature importances and the impact of class frequency on the assigned weights, to assess the relative contribution of each feature in the decision-making process and understand how class size influences weight distribution.
3. Optimization of the new feature selection approach, aimed at reducing informational redundancy and improving feature selectivity, while ensuring high computational efficiency.
4. Implementation and experimental validation of the optimized method, including recalculation of the selected features, training of the new predictive model, and quantitative performance evaluation on test data, to assess the effectiveness of the proposed method and measure improvements over the RF4PCC - implemented approach.

The methodological process was applied and validated on a purpose-built dataset developed through collaboration between the Department of Civil, Building, and Environmental Engineering (DICEA) and the Department of Architecture (DiARC) at the University of Naples Federico II, in partnership with the Religious Provinces of the Monastic Order of Saint Francis in the Campania region.

The dataset focuses on the architectural typology of the cloister structure with a standardized geometric base (quadrangular or rectangular with arcades) yet marked by significant formal diversity. This combination of complexity and standardization makes the cloister an ideal case study for automatic morphological analysis using ML techniques.Specifically, the dataset consists of eight point clouds acquired using range-based survey techniques. These represent monastic cloisters from the following sites: the Convent of San Francesco in Montella (AV), the Convent of S.S. Pietà in Teggiano (SA), the Convent of San Francesco in Padula (SA), the Monastic Complex of San Lorenzo Maggiore in Naples (NA), the Convent of Sant'Andrea in Nocera Superiore (SA), the Convent of Sant'Antonio in Nocera Inferiore (SA), the Convent of San Francesco in Solofra (AV), and the Convent of San Francesco in Benevento (BN). The discrete models in the form of point clouds were captured using phase-based Terrestrial Laser Scanners (TLS), configured to record a scan grid of 7 mm at 10 m, combined with color information. For each point cloud, manual annotation was performed on a significant portion of the dataset, segmenting and classifying groups of points into 10 architectural classes, corresponding to: "wall", "floor", "column", "molding", "vault", "arch", "stair", "window/door", "roof", and "other".

3.1 Implementation of RF4PCC model and Comparison of Results

The first step of the adopted methodology involved applying the pre-trained RF4PCC predictive model to the dataset of monastic cloisters. This model had been previously developed and trained by the 3DOM research unit of the Bruno Kessler Foundation (FBK). The aim of this phase was to assess the model's performance in a specific architectural context and to identify any areas for improvement to optimize classification results. However, the application of the pre-trained model revealed significant limitations in segmenting and recognizing the typical macro-elements of cloisters. This highlights the need for a novel methodological approach tailored to the specific characteristics of historical architecture, especially monastic architecture.

The RF4PCC model adopts a feature selection method based on Random Forest (RF) impurity reduction. This selection is applied iteratively to an initial feature set. Based on a multi-scale analysis of the training set, the authors progressively selected the most relevant features using RF - such as Planarity, Omnivariance, Surface Variation, and Verticality - computed at specific predefined radii, while also integrating the z coordinate. Although this method is computationally efficient, it relies solely on the internal discriminative power of the RF classifier, without considering the statistical significance of class distributions in the specific domain. In fact, applying this approach to the Franciscan dataset revealed systematic misclassifications between architecturally distinct classes, such as "column" and "wall", as well as "molding" and "window/door". These results underscore the limitations of a selection process driven exclusively by impurity reduction, without any external statistical validation tailored to the cloister (CH) domain.

As a result, the methodology proposed here, RF4PCC – implemented, introduces several modifications compared to the original RF4PCC approach, replacing impurity reduction with an external statistical validation framework, structured around four key methodological innovations:

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-41-2025 | © Author(s) 2025. CC BY 4.0 License.

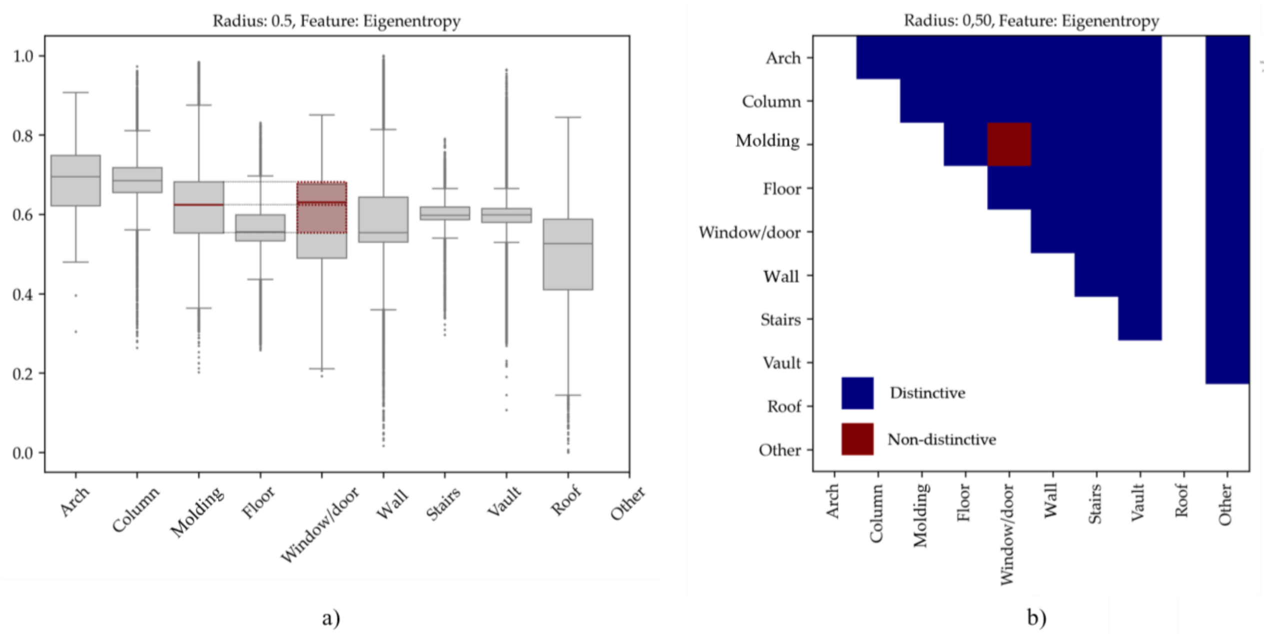


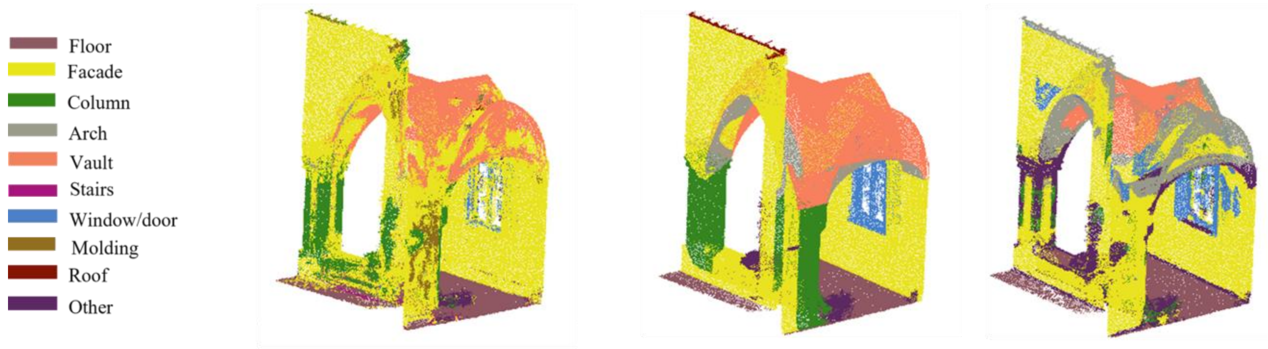
Figure 1. Selection of the optimal feature–radius combinations was carried out through an integrated analysis based on two tools: boxplots, used to visually compare class distributions and identify the most effective radii for each feature (a); triangular p-value matrices, used to assess the statistical significance of differences between class distributions (b).

1. Expansion of the geometric feature set: the new approach extends the initial set (Planarity, Omnivariance, Surface Variation, Verticality, RGB, b*, R+G+B/3) by integrating Linearity, Sphericity, Anisotropy, and Eigenentropy, specifically selected to capture morphological characteristics typical of CH;
2. Multiscale radius optimization: unlike RF4PCC, which uses predefined fixed radii, the new method implements systematic optimization (range 0.25–1.00 m) identifying the optimal radius for each feature through comparative analysis of class distributions using boxplots;
3. Independent statistical validation: Impurity reduction in Random Forest is replaced by triangular p-value matrices used to evaluate the statistical significance of differences between class distributions, offering external and objective validation of discriminative power;
4. Automated feature computation: All features are computed via a fully automated Python script, eliminating reliance on external software and ensuring full reproducibility and control over the feature extraction process.

The boxplot analysis examines the relationships between classes through the position of the median and the shape of the boxes: overlapping boxes and coinciding medians indicate similar distributions. The p-value matrices identify indistinguishable classes and features that are ineffective for discrimination. By combining boxplots and p-value matrices, it was possible to identify the optimal radii and geometric features for class differentiation (Figure 1). The new model was then trained using a dataset consisting of four monastic cloister point clouds, three for training and one for validation. The resulting predictive model was tested on two different point clouds, one partially and one completely unknown to the system. The comparison between the original and the implemented RF4PCC shows significant improvements: accuracy increased from 0.565 to 0.615, and weighted F1-score from 0.569 to 0.587, accounting for the unequal distribution of classes, while the macro F1-score decreased from 0.434 to 0.333, indicating lower average performance across all classes. So, the new model performs better in classifying major categories ("floor”, "wall" and "vault") but struggles with morphologically similar classes ("arch"/"vault”, "moldings"/"wall") (Figure 2). The "other" class still aggregates a high number of misclassifications. Notable improvements are observed for the "column" class (from 2490 to 7398 correctly classified points) and the "vault" class (from 15549 to 22113). Challenges persist for the "roof" class (due to poor representation in the dataset) and for classes with weakly distinctive geometric features, highlighting the need for dataset enrichment.

3.2 Study of *features_importances_* and the Influence of an increased number of Classes on Assigned Weights

To better understand the importance of each feature in the training process and to evaluate the reliability of the model's predictions, an in-depth analysis of the feature importance weights was conducted using the *feature_importances_* indicator from the Random Forest (RF) algorithm (Figure 3).
This analysis was necessary to identify potential biases - i.e., systematic errors or distortions that may undermine the model’s performance - and to assess whether an imbalanced class distribution in the dataset affects the estimation of feature importance. Indeed, it was observed that when a particular class is overrepresented compared to others, the model tends to systematically favor features that better discriminate the dominant class, thereby reducing the weight assigned to features that could be essential for the correct recognition of minority classes (Figure 4).This issue is particularly critical in the context of cultural heritage (CH), where elements such as moldings, capitals, and decorative details are underrepresented compared to structural elements like walls and floors. The imbalance leads to overfitting on the dominant class (i.e., excessive specialization on majority patterns by memorizing specific features rather than learning generalizable relationships), underfitting on minority classes, and a reduced ability to generalize to new data containing examples of the underrepresented classes.



a) b) c)

Figure 2. Comparison of the classification of an unknown portion of the point cloud, obtained through: the implemented RF4PCC model (a) and the RF4PCC model (c), compared to the ground truth (b) derived from the manual segmentation of the point cloud module of the monastic cloister of S. Francesco (BN).

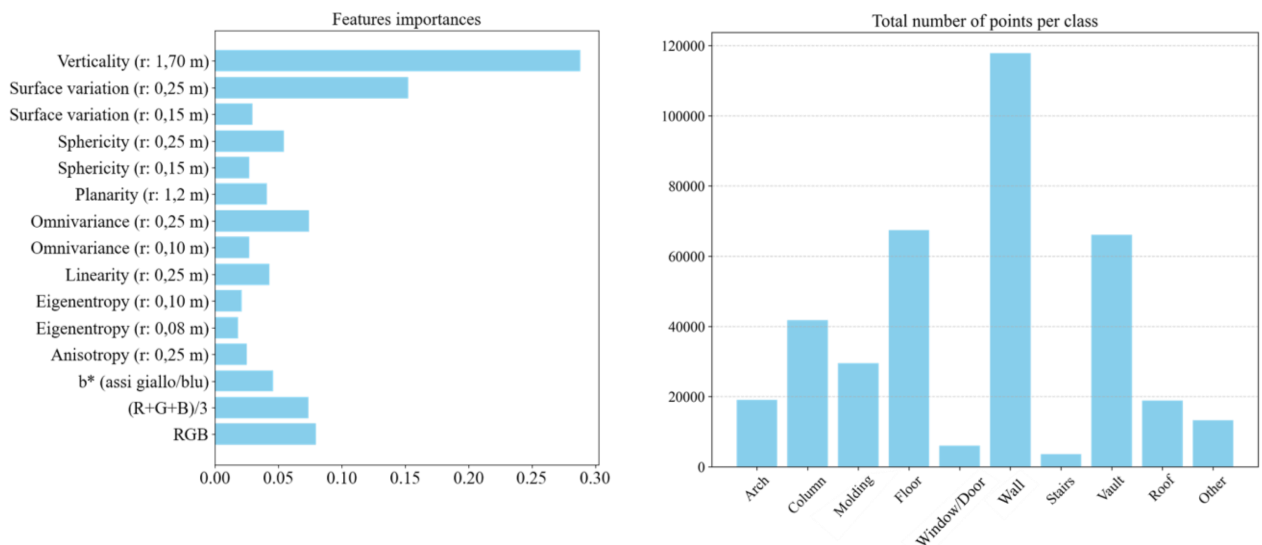
So, to address class imbalance, a systematic approach was adopted based on the optimization of the *class_weight='balanced'* parameter in Random Forest, which automatically applies inverse weighting: underrepresented classes are assigned higher weights, while overrepresented classes receive lower weights. This artificial balancing improves both problem understanding and generalization capacity, allowing each class to contribute equally to the learning process regardless of its frequency in the dataset, and reducing the tendency to favor dominant classes. However, artificial balancing can paradoxically lead to overfitting, which is particularly problematic when working with small datasets. To leverage the benefits of class balancing while avoiding overfitting, two complementary strategies were implemented:

- Dataset expansion: Four new point clouds were added to increase category representativeness and enable the learning of more robust patterns, reducing the memorization of specific features and mitigating the distorting effects of artificial balancing.
- Tree complexity control: The parameters *max_depth*=10 (an optimal compromise to capture complex relationships without excessive specialization) and *min_samples_leaf*=5 (preventing terminal nodes for single examples and leaves tailored to individual samples from minority classes) were optimized. These values were selected empirically, considering that excessive *max_depth* leads to overfitting in small or imbalanced datasets, while overly high *min_samples_leaf* limits the model's ability to learn relevant details, thus impairing predictions.

This approach results in a robust, interpretable, and reliable model capable of accurate predictions regardless of class distribution, while avoiding the overfitting risks associated with artificial balancing techniques.

3.3 Optimization of the New Feature Selection Approach

Following the analysis of feature importance and the implementation of strategies to mitigate overfitting, a systematic and objective method was developed to further accelerate and optimize the feature selection process.



a) b)

Figure 3. Relationship between feature importance (a) and class point counts (b). The most important feature is Verticality (radius 1.70 m), and the class with the most points is "wall".

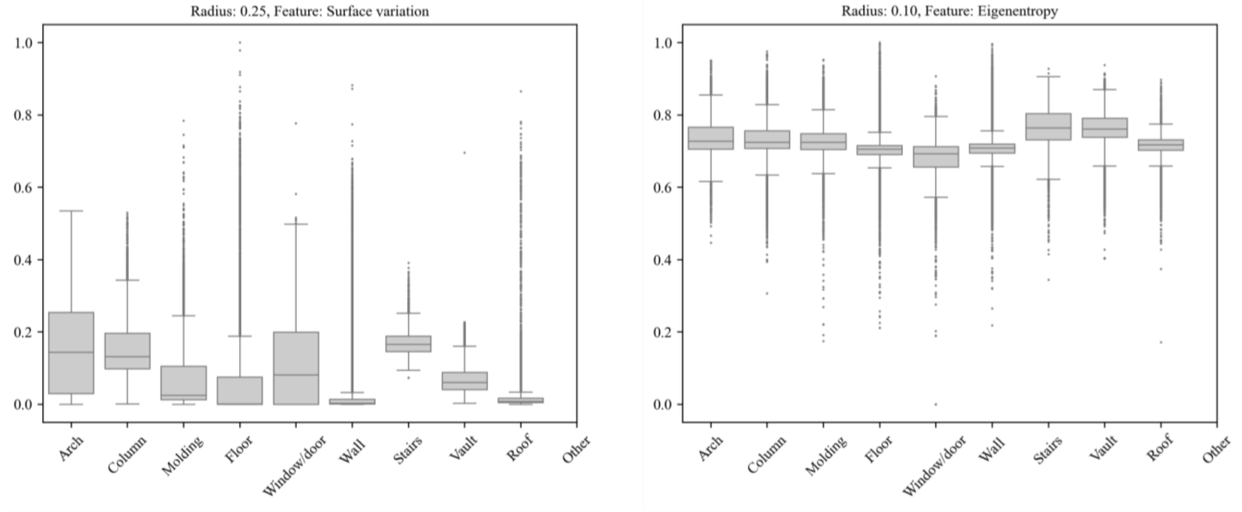


Figure 4. The boxplot analysis of the feature Surface variation, calculated with a 0.25 m radius (a), and the feature Eigenentropy (b), calculated with a 0.1 m radius - with feature importance values of 0.17 and 0.02 respectively - highlights a relationship between the importance assigned to a feature and the number of points in the class it helps to distinguish.

The main goal was to overcome the limitations of the previous approach by reducing the risk of interpretive errors and computational inefficiencies.

As previously mentioned, the feature selection process adopted in the earlier RF4PCC–implemented model required a multi-step approach involving the simultaneous comparison of p-value matrices (statistical significance of feature-radius combinations) and boxplots (class distribution analysis across the 0.25–1.00 m range). So, to eliminate subjectivity in evaluating overlaps, a single triangular matrix was developed, providing an immediate overview and significantly reducing analysis time (Figure 5). Information previously dispersed across boxplots is now directly embedded in each cell through the automatic computation of the classes best discriminated by each feature-radius combination.

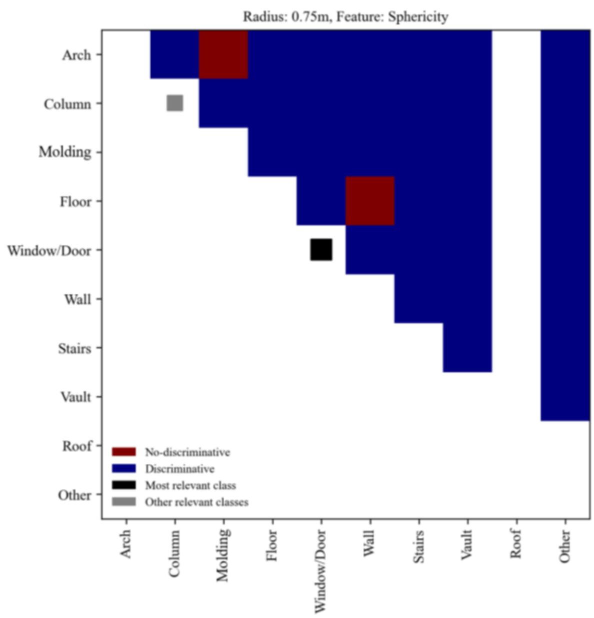


Figure 5. Triangular summary matrix of the feature-radius combinations adopted in the optimized model, where each cell automatically indicates the best-discriminated classes, integrating information obtained from boxplots and p-value matrices.

The analysis employs an automated summary table that simultaneously evaluates all triangular matrices from the training dataset, automatically identifying the optimal feature-radius combinations (Figure 6). The result is a system that directly delivers the optimal feature and radius selections for each class, removing the need for manual interpretation, reducing the risk of decision-making errors, eliminating inter-operator variability, and significantly shortening analysis time.

3.4 Implementation and Experimental Validation of the Optimized Method

Once the methodological framework for optimizing the RF4PCC algorithm was defined, an automated feature selection process was implemented on the monastic cloisters dataset. The automated summary table identified the optimal feature-radius combinations through the simultaneous analysis of all point clouds in the training set, ensuring statistical significance in discriminating between architectural classes specific to the monastic environment. This approach significantly reduced the dimensionality of the feature space, yielding a highly selective and computationally efficient subset, while minimizing the interpretive subjectivity inherent in the previous process.

The new predictive model was trained using the optimized feature dataset, retaining the previously validated RF configuration with regularization parameters (*max_depth*=10, *min_samples_leaf*=5, *class_weight='balanced'*) to ensure robustness against overfitting and fair treatment of all classes. As an additional anti-overfitting strategy, the training dataset was expanded from four to eight monastic cloisters, significantly increasing the morphological and stylistic variety of the training set and offering a more comprehensive representation of monastic architectural variability, thereby reducing the risk of memorizing specific patterns. So, the optimized model was validated following a rigorous experimental protocol using two types of test data: portions of point clouds partially seen during training and point clouds entirely unseen by the model. This dual approach enabled the assessment of both generalization capabilities on unexplored sections of architectures included in the training set, and predictive robustness on completely novel case studies, offering a comprehensive evaluation of performance in real-world operational scenarios.

Class	Top Features	Best Radii
Arch	Verticality; Sphericity	1.0m (Ve); 0.75m (Ve); 1.0m (Sp)
Column	Eigenentropy; Surface variation; Omnivariance	0.75m (Ei); 0.25m (Su); 0.75m (Om)
Molding	Verticality	0.1m (Ve); 0.5m (Ve)
Floor	Verticality; Omnivariance	0.5m (Ve); 0.75m (Ve); 0.1m (Ve); 0.25m (Ve); 1.0m (Ve); 0.5m (Om)
Window/Door	Linearity; Verticality	1.0m (Li); 0.5m (Ve); 1.0m (Ve)
Wall	Verticality; Omnivariance	0.25m (Ve); 0.75m (Ve); 0.5m (Ve); 0.1m (Ve); 0.5m (Om)
Stairs	Verticality	1.0m (Ve); 0.75m (Ve)
Vault	Verticality	1.0m (Ve); 0.75m (Ve); 0.25m (Ve)
Roof	Sphericity	0.75m (Sp)
Other	Anisotropy; Sphericity; Surface variation	0.1m (An); 1.0m (Sp); 0.75m (Su); 0.75m (Sp)

Figure 6. Automated summary table of the optimal feature-radius combinations for the different point clouds in the training dataset. The system simultaneously aggregates and analyzes the results of the triangular matrices, automatically identifying the most effective choices for class discrimination, while reducing analysis time, subjectivity, and operator variability.

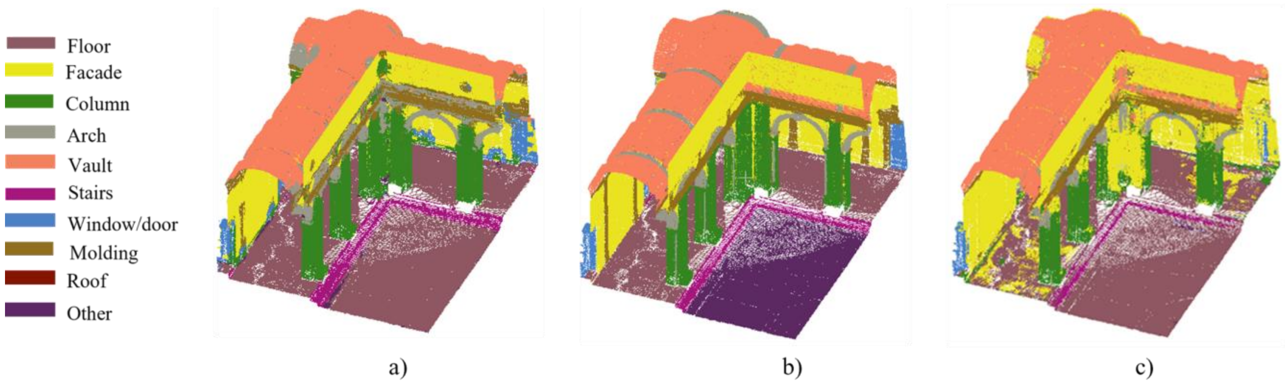


Figure 7. Comparison between the classification of the portion of the point cloud partially known to the classifier, obtained through: the optimized model (a) and the implemented RF4PCC model (c), versus the ground truth (b) derived from the manual segmentation of the point cloud module of the monastic cloister of San Francesco in Padula (SA).

3.4.1 Performance Comparison: Implemented RF4PCC vs. Optimized Model on Known Dataset

The comparative analysis of performance metrics reveals a targeted pattern of improvements with the optimized method, aligned with the objectives of class balancing and enhanced identification of minority architectural elements (Figure 7).
Overall accuracy shows a slight decrease from 66.9% to 65.3% (-2.4%), offset by a significant increase in the macro F1-score from 0.499 to 0.512 (+2.6%) and in the weighted F1-score from 0.623 to 0.632 (+1.4%), indicating more balanced performance across the different architectural classes.
The most notable results emerge in the analysis of minority classes, which have historically been underrepresented in architectural datasets. The “molding” class shows an F1-score improvement from 0.487 to 0.549 (+12.7%), driven by a substantial increase in recall from 34.4% to 46.3% (+34.6%), indicating a significantly improved ability to identify decorative elements within the point cloud. Although precision drops from 83.2% to 67.4%, this trade-off is strategically beneficial in the context of monastic heritage analysis, where capturing all significant decorative features - even at the cost of requiring subsequent manual validation - is preferred.
The “column” class reflects a strategic shift in classification approach, with a notable increase in recall from 69.4% to 90.8% (+30.8%), showing that the optimized model identifies over 90% of columns present. The reduction in F1-score from 0.733 to 0.621 (-15.3%) is primarily due to a decline in precision (from 77.8% to 47.2%), reflecting a more "inclusive" strategy that favors comprehensive identification of vertical structural elements, crucial for the morphological analysis of monastic cloisters.

Figure 8. Confusion matrix of the optimized model compared to the classification of the point cloud partially known to the classifier. Significant increases are observed in correctly classified points for the classes “column” (43694), “molding” (46856), “floor” (87389), and “stairs” (8201). The “vault” class shows a slight decrease (117532), while confusions between adjacent classes are reduced, confirming the model’s effectiveness in improving the discrimination of architectural elements in the point cloud.

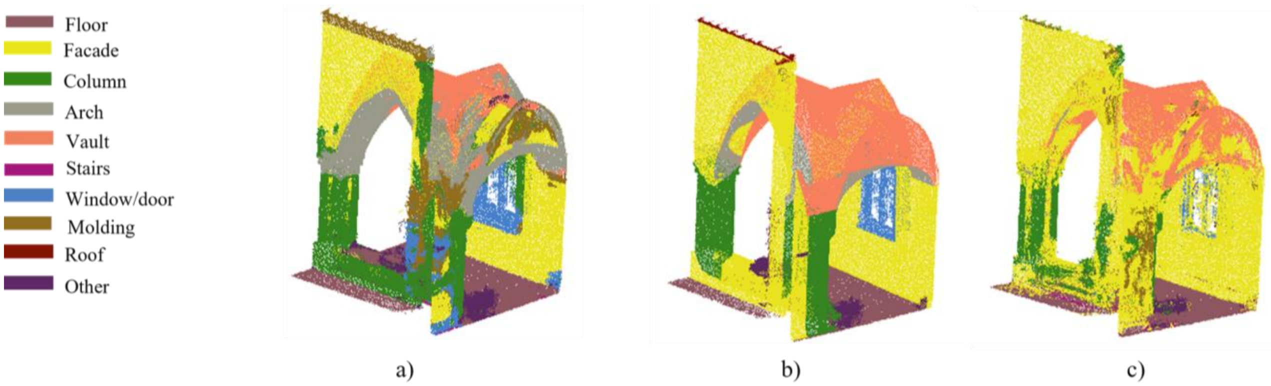


Figure 9. Comparison between the classification of the portion of the point cloud unknown to the classifier, obtained through: the optimized model (a) and the implemented RF4PCC model (c), against the ground truth (b) derived from the manual segmentation of the point cloud module of the monastic cloister of S. Francesco in Benevento (BN).

The comparative analysis of the confusion matrices (Figure 8) highlights significant improvements with the optimized method when classifying the point cloud partially known to the model, confirming the trends observed in the performance metrics.
The most significant improvement is seen in the classification of the “column” class, which increased from 33370 to 43694 correctly classified points. This is accompanied by a notable reduction in confusion with the “molding” and “wall” classes, indicating a stronger discriminative capability of the optimized model in distinguishing vertical structural elements, consistent with the recall increase (from 69.4% to 90.8%).
Similarly, the “molding” class shows a substantial improvement from 34846 to 46856 correctly classified points, reflecting better discrimination of decorative elements, which have historically posed challenges for automated classification systems in historic architecture. This result aligns with the +12.7% F1-score improvement for the “molding” class and confirms the effectiveness of the optimized method in identifying underrepresented decorative elements.

3.4.2 Performance Comparison: Implemented RF4PCC vs. Optimized Model on Unknown Dataset

The analysis of results on the entirely unseen point cloud provides a critical assessment of the generalization capabilities of the two classification approaches (Figure 9).
Contrary to the results on the partially known point cloud, the RF4PCC method achieves higher accuracy (61.5% vs. 57.7%), indicating greater stability on novel data. However, the optimized method shows a significant increase in the macro F1-score (from 0.333 to 0.414, +24.3%), demonstrating a superior ability to handle minority classes even under completely new dataset conditions.
The comparison of the confusion matrices (Figure 10) confirms the strategic differences between the two approaches.
The “column” class exhibits the most significant improvement, with an F1-score increase from 0.444 to 0.578 (+30.2%), resulting from a substantial rise in recall from 36.7% to 71.0%, which translates to an increase in correctly classified points from 7398 to 14335. Particularly noteworthy is the improvement in the classification of “window/door” elements, where the optimized method achieves an F1-score of 0.499 compared to 0.227 in the previous method, driven by a dramatic recall increase from 13.1% to 52.4%.
The comparative analysis on unseen data confirms the effectiveness of the optimization framework in reducing bias toward dominant classes and significantly enhancing the identification of minority architectural crucial elements for comprehensive documentation of historical heritage.

Figure 10. Confusion matrix of the optimized model on point clouds unknown to the classifier, showing significant increases in correctly classified points for “column” (14335) and “window/door” (4478).

The optimized method successfully meets its objectives of improving the recognition of minority elements while maintaining competitive performance on dominant classes, thus providing a more balanced approach tailored to the specificities of the monastic heritage analyzed.

4. Conclusions and future developments

The proposed implementation method proved effective in improving the segmentation and classification of point clouds within the context of architectural cultural heritage, with specific reference to Franciscan monastic cloisters. The results confirm that the systematic optimization of feature selection, combined with external statistical validation, can overcome the limitations of approaches based solely on impurity reduction in the Random Forest classifier. Feature selection based on variable radii improved model accuracy by reducing classification ambiguities, while the use of p-value matrices made the process more objective and automated, enhancing generalization capability. The creation of a dedicated cloister dataset enabled rigorous and context-specific validation, demonstrating the model’s ability to adapt to complex morphologies. However, generalization remains limited to the cloister typology, highlighting the need for testing on different structural types. Balancing precision and recall for minority classes requires further development, especially in cases where minimizing false positives is critical, and the method still depends on the availability of high-quality annotated datasets. Future developments aim to extend the application to architectural subcategories, increasing the analysis granularity to include the recognition of specific details such as moldings, capitals, and decorative elements, which will require dataset diversification and integration with hybrid learning systems to balance interpretability and performance. Overall, the framework represents a significant contribution to the automatic classification of point clouds for cultural heritage, demonstrating that rigorous, statistically grounded approaches can surpass the limitations of traditional methods and marking a substantial advancement in the digital documentation of heritage assets.

Acknowledgements

This article is the result of the joint research efforts of the authors. The specific written contributions are attributed as follows: G.A. wrote paragraphs 3, 3.2, 3.4.1 and 4; V.C. wrote paragraphs 1, 3.3 and 3.4.2; D.C. wrote paragraphs 2, 3.1 and 3.4. The project is part of an agreement between the University of Naples Federico II, and the Religious Provinces of the Franciscan Order in Campania. The scientific coordinators of the project are G. Antuono and V. Cera.

References

Atik, M. E., Duran, Z., 2022. Selection of Relevant Geometric Features Using Filter-Based Algorithms for Point Cloud Semantic Segmentation. *Electronics*, 11(20), 3310. doi.org/10.3390/electronics11203310.

Buldo, M., Agustín-Hernández, L., Verdoscia, C., 2024. Semantic Enrichment of Architectural Heritage Point Clouds Using Artificial Intelligence: The Palacio de Sástago in Zaragoza, Spain. *Heritage*, 7(12), 6938-6965. doi.org/10.3390/heritage7120321.

Cao, Y., Teruggi, S., Fassi, F., Scaioni, M., 2022. A Comprehensive Understanding of Machine Learning and Deep Learning Methods for 3D Architectural Cultural Heritage Point Cloud Semantic Segmentation. In: Borgogno-Mondino, E., Zamperlin, P. (eds). *Geomatics for Green and Digital Transition*, 1651, 329-341. Springer, Cham. doi.org/10.1007/978-3-031-17439-1_24.

Chen, C., Liaw, A., Breiman, L., 2004. *Using random forest to learn imbalanced data*. University of California, Berkeley, Department of Statistics, Technical Report. https://statistics.berkeley.edu/sites/default/files/techreports/666.pdf.

Gaber, J.A., Youssef, S.M., Fathalla, K.M., 2023. The Role of Artificial Intelligence and Machine Learning in Preserving Cultural Heritage and Art Works via Virtual Restoration. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-1/W1-2023, 185-190. doi.org/10.5194/isprs-annals-X-1-W1-2023-185-2023.

Gîrbacia, F., 2024. An Analysis of Research Trends for Using Artificial Intelligence in Cultural Heritage. *Electronics*, 13(18), 3738. doi.org/10.3390/electronics13183738.

Grilli, E., Remondino, F., 2020. Machine Learning Generalisation across Different 3D Architectural Heritage. *ISPRS International Journal of Geo-Information*, 9(6), 379. doi.org/ 10.3390/IJGI9060379.

Gu, Q., Tian, J., Li, X., Jiang, S., 2022. A novel random forest integrated model for imbalanced data classification problem. *Knowledge-Based Systems*, 250(C), 109050. doi.org/10.1016/j.knosys.2022.109050.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.

Harshit, H., Kushwaha, S.K.P. Jain, K., 2022. Geometric Features Interpretation of Photo-grammetric Point Cloud from Unmanned Aerial Vehicle. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W2-2022, 83-88. doi.org/10.5194/isprs-annals-X-4-W2-2022-83-2022.

Lin, H.-I.,Nguyen, M. C., 2020. Boosting Minority Class Prediction on Imbalanced Point Cloud Data. *Applied Sciences*, 10(3), 973. doi.org/10.3390/app10030973.

Moyano, J., Musicco, A., Nieto-Julian, J.E., Domínguez-Morales, J.P., 2024. Geometric characterization and segmentation of historic buildings using classification algorithms and convolutional networks in HBIM. *Automation in Construction*, 167, 105728. doi.org/10.1016/j.autcon.2024.105728.

Ni, H., Lin, X., Zhang, J., 2017. Classification of ALS point cloud with improved point cloud segmentation and random forests. *Remote Sensing*, 9(3), 288. doi.org/10.3390/rs9030288.

Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E.S., Frontoni, E., Lingua, A.M., 2020. Point cloud semantic segmentation using a deep learning framework for cultural heritage. *Remote Sensing*, 12(6), 1005. doi.org/10.3390/rs12061005.

Terruggi, S., Grilli, E., Russo, M., Fassi, F., Remondino, F., 2020. A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification. *Remote Sensing*, 12(16), 2598. doi.org/10.3390/ RS12162598.

Weinmann, M., Jutzi, B., Mallet, C., 2013. Feature Relevance Assessment for the Semantic Interpretation of 3D Point Cloud Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5-W2, 313-318. doi.org/10.5194/isprsannals-II-5-W2-313-2013.

Xie, Y., Tian, J., Zhu, X.X., 2020. Linking Points with Labels in 3D: A Re-view of Point Cloud Semantic Segmentation. *IEEE Geoscience and Remote Sensing Magazine*, 8(4), 38-59. doi: 10.1109/MGRS.2019.2937630.

Yang, S., Miaole, H., Songnian, L., 2023. Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review. *Remote Sensing*, 15(3), 548. doi.org/10.3390/rs15030548.

Zhao, J., Hua, X., Yang, J., Yin, L., Liu, Z., Wang, X., 2023. A review of point cloud segmentation of architectural Cultural Heritage. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-1/W1-2023, 247-254. doi.org/10.5194/isprs-annals-X-1-W1-2023-247-2023.

3DOM-FBK/RF4PCC. Available online: https://github.com/3DOM-FBK/RF4PCC (accessed on 02 September 2024).