

## UNSUPERVISED OBJECT-BASED CLUSTERING IN SUPPORT OF SUPERVISED POINT-BASED 3D POINT CLOUD CLASSIFICATION

E. Grilli<sup>1</sup>, F. Poux<sup>2</sup>, F. Remondino<sup>1</sup>

<sup>1</sup> 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy  
web: <http://3dom.fbk.eu> – Email: <grilli><remondino>@fbk.eu

<sup>2</sup> Geomatics Unit, University of Liège (ULiege), Liège, Belgium – Email: fpoux@uliege.be

Commission II, WGII/6

**KEY WORDS:** Point cloud, Segmentation, Classification, Machine learning, Clustering, Feature extraction.

### ABSTRACT:

The number of approaches available for semantic segmentation of point clouds has grown exponentially in recent years. The availability of numerous annotated datasets has resulted in the emergence of deep learning approaches with increasingly promising outcomes. Even if successful, the implementation of such algorithms requires operators with a high level of expertise, large quantities of annotated data and high-performance computers. On the contrary, the purpose of this study is to develop a fast, light and user-friendly classification approach valid from urban to indoor or heritage scenarios. To this aim, an unsupervised object-based clustering approach is used to assist and improve a feature-based classification approach based on a standard machine learning predictive model. Results achieved over four different large scenarios demonstrate the possibility to develop a reliable, accurate and flexible approach based on a limited number of features and very few annotated data.

### 1. INTRODUCTION

The current 3D research activities – which extends to various domains and applications – are dominated by classification endeavours. The semantic segmentation (or more commonly classification) task is a significant challenge for 3D unstructured datasets acquired with active or passive sensors (Xie et al., 2020). Recognising elements composing a scene is a crucial step, especially for Digital Twins (Stojanovic et al., 2018), Smart Cities (Nys et al., 2020) and Building Information Modelling (Bassier et al., 2017). In this context, there is a great demand for automated processes that can speed-up and improve the reliability of existing classification frameworks. However, we can safely say that there are no reliable and generalised methods for all the different scales and scenarios one can encounter. A classification method can hardly fulfil all domains since the semantic definitions attached to objects can differ depending on the domain. Most semantic segmentation methods aim to advance performances in one specific context, such as indoor structural recognition (ceiling, wall, floors, chair, etc.) (Dai et al., 2017; Stojanovic et al., 2019) or outdoor classification (street, building, car, vegetation, etc.) (Özdemir et al., 2019; Hu et al., 2020). This is also because modern approaches mainly rely on supervised methods based on neural networks (Guo et al., 2020), which necessitate annotated context-specific datasets such as the ones provided by Armeni et al. (2017) and Tan et al. (2020). These approaches are often fully supervised, rarely unsupervised and require a high level of expertise on top of high computing resource demands.

#### 1.1 Aim and structure of the paper

This paper aims at an easy-to-implement and user-friendly supervised method generalisable to several contexts and domains. To achieve this, we explore how unsupervised object-based features (Poux and Billen, 2019; Poux et al., 2020; Poux and Ponciano, 2020) can help a supervised point-based classification (Grilli et al., 2019; Grilli and Remondino, 2020). The goal is to combine two different classification approaches to maximise results' accuracy, minimise human efforts and deliver a 3D classification method that is case- and context-independent

while usable by non-experts. Experimental results are conducted on heterogeneous datasets (Section 1.2), including multi-scale urban areas (aerial LiDAR and photogrammetry), indoor buildings (RGB-D sensor) and architectural scenarios (terrestrial photogrammetry). The achieved results demonstrate the method's reliability and replicability. In the age of deep learning, the suggested method relies on a standard machine learning algorithm (Random Forests) to achieve fast and accurate point cloud classification, using reduced annotated samples and a minimal number of automatically computed features. Thus, primary purpose of this study is not the direct comparison with state-of-the-art methods but rather a study to evaluate how to improve 3D classification results by merging features and methods used in Poux et al. (2020) and Grilli et al. (2019). Following a summary of related works in Section 2, we define our methodology in Section 3. Section 4 presents the results over four heterogeneous scenarios to demonstrate the efficiency of the presented approach. Finally, in Section 5, we give some closing remarks as well as suggestions for future work.

#### 1.2 Considered scenarios

The presented method was evaluated on the following scenarios:

- Large scale urban point cloud (700 m × 700 m), derived with a hybrid aerial sensor over the city centre of *Bordeaux* (Toschi et al., 2021), including four classes (street, facades, roof, vegetation).
- Large scale urban point cloud (220 m × 200 m), belonging to the *Dublin* dataset (Zolanvari et al., 2019), acquired with a LiDAR sensor (without colour information) and featuring five classes (street, vegetation, façade, window, roof).
- Indoor point cloud (22 m × 28 m), available in the *S3DIS* dataset (Armeni et al., 2016), acquired with the Matterport sensor (Lehtola et al., 2017), including eight classes (roof, floor, facade, window, door, table, chair, and bookcase).
- Architectural point cloud (240 m length), derived with photogrammetry for the *Bologna*'s porticos and belonging to the ArCH benchmark (Matrone et al., 2020), composed of seven classes (arch, column, moulding, façade, vault, window, floor).

## 2. RELATED WORKS

The main methods for 3D classification reported in the most recent literature can be divided into two big macro-categories: machine and deep learning approaches. The feature engineering phase is one of the primary distinctions between standard machine learning methods and advanced deep learning methods. In the first case, the operator studies and selects the features, whereas in the second case, neural networks learn features after being fed large amounts of annotated data.

According to the application's purpose of this research, this section is first focused on the point cloud feature selection, then on the existing approaches based on a combination of clustering and point-based classification.

**Feature selection.** Establishing the features to be used in the model is a critical step in the supervised classification analysis. Most of the similar studies depend on geometric features to classify the points of a considered point cloud based on their local neighbourhood. The neighbourhood to be used can be defined using either an established radius, which can be spherical (Lee and Schenk, 2002) or cylindrical (Filin and Pfeifer, 2005) or a K number of nearest neighbours (KNN) (Linsen and Prautzsch, 2001). The sampling rate resulting from data acquisition and the items of interest influence the choice of an acceptable value (radius or K). For this reason, all these neighbourhood types have been and are still broadly explored in literature within single or multi-scale approaches (Weinmann et al., 2015). On the other, multi-scale approaches have proved to be the most efficient, whether used for spherical/cylindrical neighbourhoods (Brodu and Lague 2012; Niemeyer et al. 2014; Thomas et al., 2018), or KNN (Hackel et al. 2016). Regarding the geometric features themselves, height based (Hackel et al., 2016; Özdemir et al., 2019), as well as angular-based features (Muñoz et al. 2008) are commonly used for point cloud classification purposes. Besides, the eigenvalue-based features proposed by Pauly et al. (2003) have proved to help describe the local geometry of the point. These features are shape descriptors derived from a combination of the eigenvalues ( $\lambda_1 > \lambda_2 > \lambda_3$ ) extracted from the covariance matrix (Blomley et al., 2014). The features values illustrate the main linear (1D), planar (2D), or volumetric (3D) structure of the point cloud in the neighbourhood. Blomley et al. (2014) offer detailed insights into the traditional eigen-based covariance features in differing scale scenarios. Deeper analyses have then been proposed in Weinmann et al. (2014, 2015, 2017a).

**Classification methods.** The semantic segmentation methods-denoted classification in this article- can vastly differ depending on the feature set provided as an input to the machine learning classifier. In the literature, we usually distinguish point-based classifiers (that reason from a per-point feature set) from segment-based classifier (per-segment labelling). The latter usually relies on a segmentation step, where the point cloud is partitioned into subsets of points called 'segments'. In addition to neighbourhood definitions found at the point-level to achieve point-based classification, such as shown in Bremer et al. (2013), other characteristics can describe each segment to guide the process. The result is a set of internally homogeneous segments, i.e. groups of points representing the basic units for classification. In many cases, segmentation procedures aim to produce relatively small segments (over-segmentation), representing only object parts (sub-objects) rather than the final objects of interest directly. In Chehata et al. (2009) first, a supervoxel-based segmentation is used to segment point cloud data, then different machine learning algorithms are tested to label the point cloud. Luo et al. (2018) proposed a supervoxel-based classification; their method used Conditional Random Field matching to

classify supervoxels. Sun et al. (2018) used a Random Forest classifier to classify point cloud based on supervoxels. Some authors rely on a region growing algorithm for segmentation of point cloud followed by an object-based classifier such as SVM (Yang et al., 2017) or a Bagged Tree Classifier (Bassier et al., 2020). use segment-based shape analysis relying on semantic rules.

This article investigates the merging of both clustering and point-based classification to develop a user-friendly approach. A small number of similar attempts have been proposed in the literature, such as presented in Weinmann et al. (2017b) where segment-based shape analysis relies on semantic rules. This approach motivates the gains of a "higher level" understanding of the scene translated into features that can help achieve better inference. Other works which relies on "segment features" in point-based classification frameworks can be found in Guinard et al. (2017), Landrieu et al. (2017) and Landrieu and Simonovsky (2018).

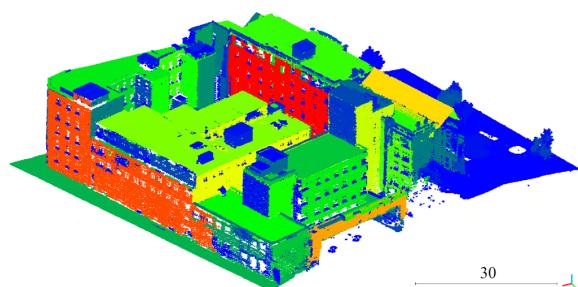
## 3. METHODOLOGY

This section first describes the different steps of our framework (Section 3.1) and then explains the features used in our combined classification approach (Section 3.2).

### 3.1 Framework

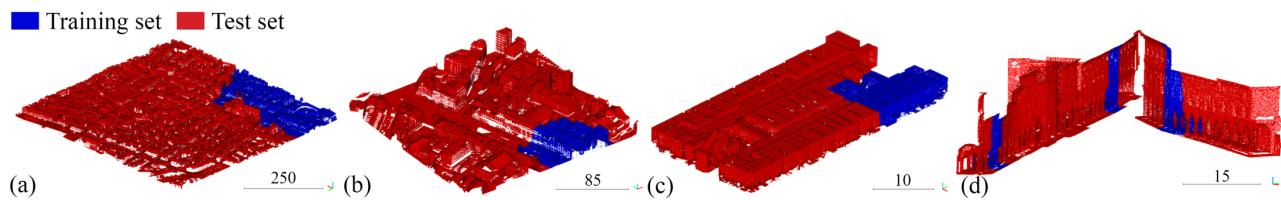
The combined approach presented in this study follows these main steps:

- Apply unsupervised clustering, following the approach presented in Poux et al. (2020) to segment the datasets.
- Extract of a small set of geometric and covariance-based features which are effective for heterogeneous scenarios: the feature selection and their computation within heuristic neighbours is automatically performed to by-pass the otherwise laborious feature design process (Section 3.2).
- Manually annotate a reduced portion of the point cloud, facilitated and supported by the clustering results (Figure 1): although the datasets used in this paper contains fully annotated point clouds, the data were divided into training (30%) and test (70%) sets (Figure 2). Our idea is that, when training is not available, the training's size should be as limited as possible and the annotation step rapid and user friendly.



**Figure 1.** Clustering output for the training set of the Dublin dataset.

- Train a standard machine learning algorithm: based on previous performance analysis (Grilli et al., 2019; Grilli and Remondino, 2020; Matrone et al., 2020), we selected the widely known Random Forest classifier (Breinman, 2001) among different classifiers available in the literature (SVM, Decision tree, etc.).



**Figure 2.** Training and test sets for the four case studies: Bordeaux (a), Dublin (b), S3DIS (c), Bologna (d).

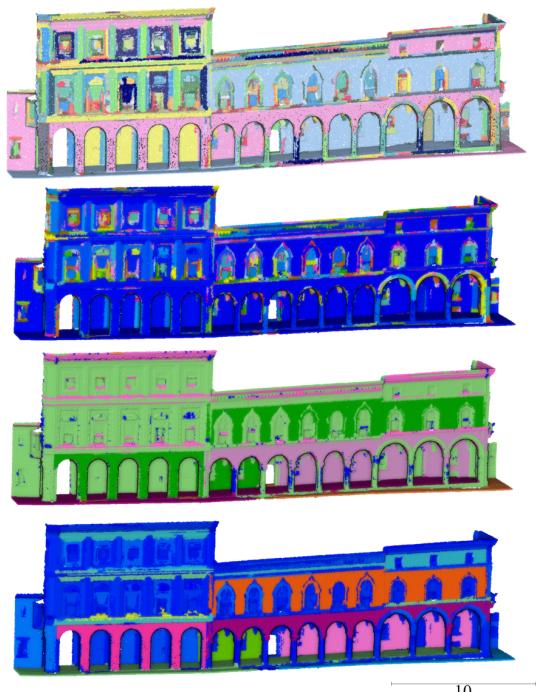
- Assess the achieved point-wise classification outcomes through quality metrics extracted for the entire test set: among the several metrics existing in the literature (Goutte and Gaussier, 2005), the Overall Accuracy (OA) is used to evaluate the classifier's ability to predict labels based on all observations. In addition, it is considered the F1-score, as it's a good measure of how well the classifier performs, being an average of Precision and Recall.

### 3.2 Feature engineering

We aim to design a reduced number of meaningful features that can be used and adapted in different scenarios for point cloud classification. These features can then be fed into standard classifiers to train machine learning predictive models.

Three main categories of features are combined: a) radiometric (RGB values), b) clustering and c) geometric features.

**Clustering features** (Figure 3). Clustering methods have two major advantages: (i) they don't use prior knowledge on discriminating variables and (ii) they find answers directly in the data. This allows exploring fed variables and highlight unsuspected (or suspected) relationships. The clustering features are computed following an unsupervised scheme, where the point cloud is partitioned into subsets of neighbouring points called segments (Poux et al., 2020; Poux and Ponciano, 2020; Bassier et al., 2020). We aim at a set of internally homogeneous segments that will host the cluster features at four different aggregation levels.



**Figure 3.** Four levels of clustering features: from top to bottom level 1, 2, 3 and 4.

The procedure aims to yield relatively small segments, representing only object parts (sub-objects) rather than the final objects of interest directly, which, sorted, constitute the Cluster feature Level 1. Then adjacent segments with similar properties are merged to spatially contiguous objects (Cluster feature level 2) by considering the results of the first clustering pass instead of the initial point cloud as input. Such a step-wise procedure, based on an initial over-segmentation, permits reducing the risk of combining multiple real-world objects in one segment (under-segmentation). The principle of the approach is to limit any domain knowledge and parameter tweaking to provide a fully unsupervised clustering featuring. Finally, level 3 is based on a k-means clustering using the Hartigan's rule (Chiang and Mirkin, 2010), whereas level 4 is a graph-based centrality-measure of the clusters weighted over the first three levels.

The initial parameters involved in the definition of these multi-level clustering features are automatically extracted through an automatic heuristic determination of three RANSAC-inspired clustering parameters:

- a distance threshold for neighbourhood definition ( $\varepsilon$ );
- the threshold for the minimum number of points needed to form a valid planar region ( $\tau$ );
- the decisive criterion for adding points to a region ( $\alpha$ ).

**Geometric features** (Figure 4). Covariance features (Blomey et al., 2014), or eigen-based features, are commonly used in segmentation and classification procedures due to their ability to provide in-depth information on the geometrical layout of the reconstructed scene. The most common covariance features include (Table 1):

$$\text{Linearity} \quad L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1} \quad (1)$$

$$\text{Planarity} \quad P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1} \quad (2)$$

$$\text{Sphericity} \quad S_\lambda = \frac{\lambda_3}{\lambda_1} \quad (3)$$

$$\text{Omnivariance} \quad O_\lambda = \sqrt[3]{\prod_{j=1}^3 \lambda_j} \quad (4)$$

$$\text{Anisotropy} \quad A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1} \quad (5)$$

$$\text{Eigenentropy} \quad E_\lambda = -\sum_{j=1}^3 \lambda_j \ln(\lambda_j) \quad (6)$$

$$\text{Sum of eigenvalues} \quad \Sigma_\lambda = \sum_{j=1}^3 \lambda_j \quad (7)$$

$$\text{Surface variation} \quad C_\lambda = \frac{\lambda_3}{\Sigma \lambda} \quad (8)$$

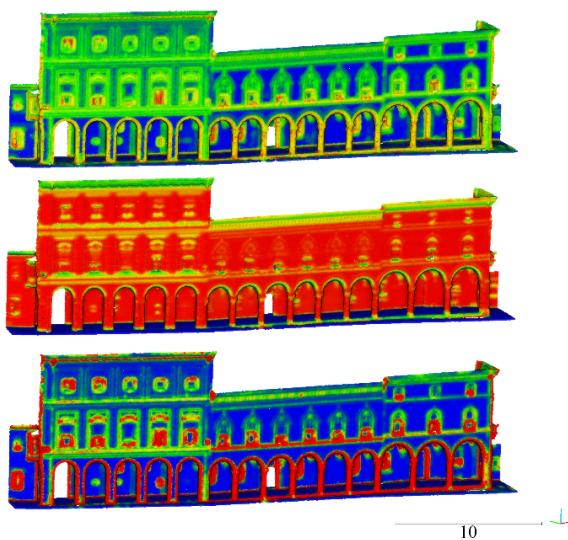
**Table 1.** Covariance features.

Grilli et al. (2019) presented a study about the behaviour of the covariance features measured within spherical neighbourhoods at increasing radius sizes in order to pick a smaller number of features that could be useful for heritage case study classification. In the same way, in this paper, we aim to identify a reduced number of geometric features that can be used in any possible environment, notwithstanding a fast computation.

The number of covariance features used in the presented approach is reduced to the only two Omnivariance and Surface Variation, chosen because of their ability to distinguish macro-elements and entities (Teruggi et al., 2020). In order to demonstrate the effectiveness of the reduced selection, the classification experiments were also carried out using the entire set of features (Section 4).

In addition, a height-based feature (Distance from Ground) and a normal-based one (Verticality) are considered. In particular, we have noticed that the feature Verticality is typically needed, independently from the scenario, to differentiate precisely horizontal and vertical artefacts. Directly related, the use of a height-based feature like the Distance from ground ( $\Delta z$  component) becomes essential to distinguish the different horizontal elements (i.e., street and roof).

It has to be underlined that the selected features are extracted within a spherical radius  $\epsilon$  (offered by the Clustering features) in a multi-scale approach. Experimentally we observed that a maximum number of  $4\epsilon$  was optimal for all scenarios.



**Figure 4.** Examples of geometry-based features: from top to bottom Omnivariance, Verticality and Surface Variation.

#### 4. EXPERIMENTS AND RESULTS

For each case study, the Random Forest classifier was trained with different combinations of features to have an internal comparison between the proposed method and the ones we combined. In particular:

- A. *Clustering*: clustering features (treated at a point level) are given in input to the classifier. Further experiments will include a segment-level classification. Total number of features: 4.
- B. *Covariance features in a multi-scale approach*: all the different covariance features (Table 1) are extracted at increasing radii, together with Verticality and Distance from ground. Total number of features: 73 (1 height-based, 9 geometric x 8 scales).
- C. *Proposed approach*: combination of clustering features and few selected geometric features calculated at four scales ( $\epsilon$ ,  $2\epsilon$ ,  $3\epsilon$ ,  $4\epsilon$ ). Total number of features: 17 (1 height-based, 3 geometric x 4 scales, 4 clustering).

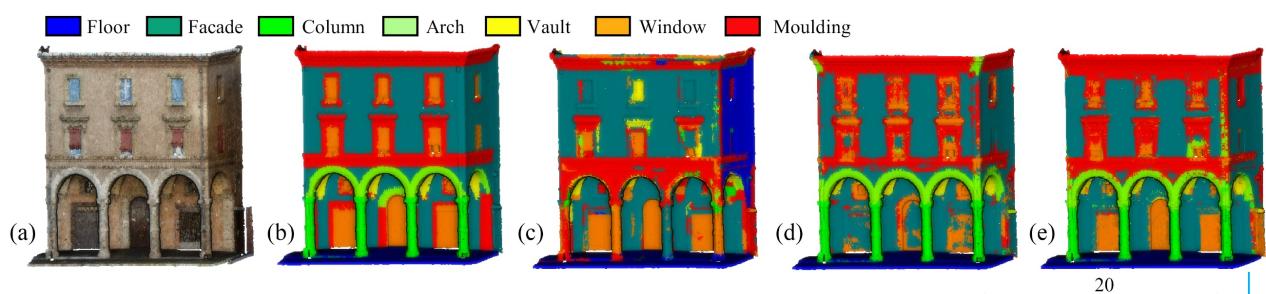
Besides, all datasets were treated with and without their radiometric attributes to further test the approach's reliability in different conditions.

Tables 2 shows results coming from the above-mentioned feature combinations in the four considered scenarios (Section 1.2). It can be seen that the proposed approach, which combines clustering features and a few selected geometric features, leads to an improvement in results. It is also noticeable that the combined approach performed better when radiometric features were included. Besides, we can see that the better improvements were achieved for the urban scenarios. In fact, quite similar accuracy values were reached between the standard multi-scale (Approach B) and the combined approach for the indoor and architectural datasets. However, from a qualitative point of view, classification results look much "cleaner" when cluster and geometric features are combined (Figure 5 and 6). In addition, it has to be considered that only 17 features were used for the proposed approach, against the 73 of the standard one.

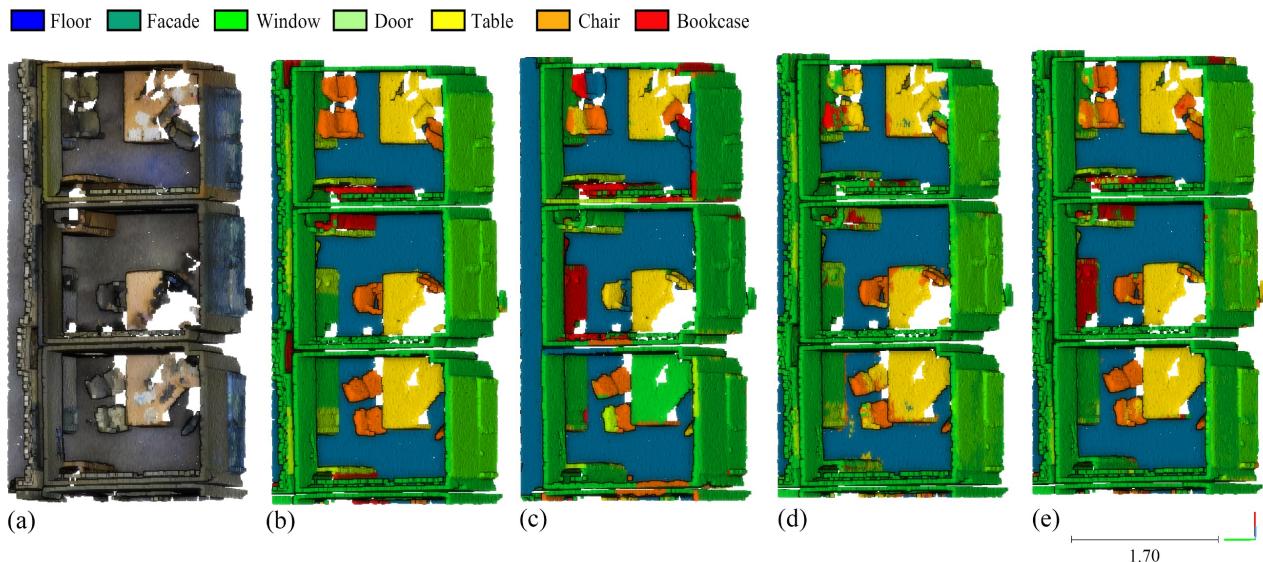
For more details about the quality of the results, please check Figures 7-10, comparing hand-annotated and predicted point clouds. Finally, in Tables 3-6, all the per-class F1 scores are reported.

	URBAN - Bordeaux		URBAN - Dublin		INDOOR - S3DIS		ARCHIT. - Bologna	
CLASSIF. APPROACH	OA %	F1 %	OA %	F1 %	OA %	F1 %	OA %	F1 %
No RGB	A	81.43	81.07	84.91	84.45	69.54	67.54	63.04
	B	88.31	88.14	90.82	90.54	83.25	81.00	79.05
	C (Proposed)	<b>92.52</b>	<b>92.47</b>	<b>93.13</b>	<b>92.89</b>	<b>84.96</b>	<b>84.11</b>	<b>79.09</b>
RGB	A	82.78	82.46	-	-	73.25	72.71	60.39
	B	88.46	88.28	-	-	87.55	86.63	80.31
	C (Proposed)	<b>92.56</b>	<b>92.51</b>	-	-	<b>88.59</b>	<b>88.07</b>	<b>80.38</b>

**Table 2.** Classification metrics achieved in the four scenarios using different features.



**Figure 5.** A portion of the Bologna dataset coloured (a) and hand-annotated (b), with its classification results achieved using clustering features (c), geometry-based features in a standard approach (d), combining clustering and geometry-based features (e).



**Figure 6.** A portion of the S3DIS dataset coloured (a) and hand-annotated (b), with its classification results achieved using clustering features (c), geometric features in a multiscale approach (d), combining clustering and geometry-based features (e).

F1 scores				
	Ground	Façade	Roof	Vegetation
<b>A</b>	87.17	71.04	87.88	71.38
<b>B</b>	85.3	84.21	91.97	73.4
<b>C</b>	94.79	85.01	94.89	84.85

**Table 3.** Per-class F1 scores for the Bordeaux dataset.

F1 scores					
	Ground	Veg	Façade	Window	Roof
<b>A</b>	94.42	44.07	62.59	10.02	86.51
<b>B</b>	97.13	76.31	86.63	43.6	94.11
<b>C</b>	96.3	93.58	87.66	40.76	94.33

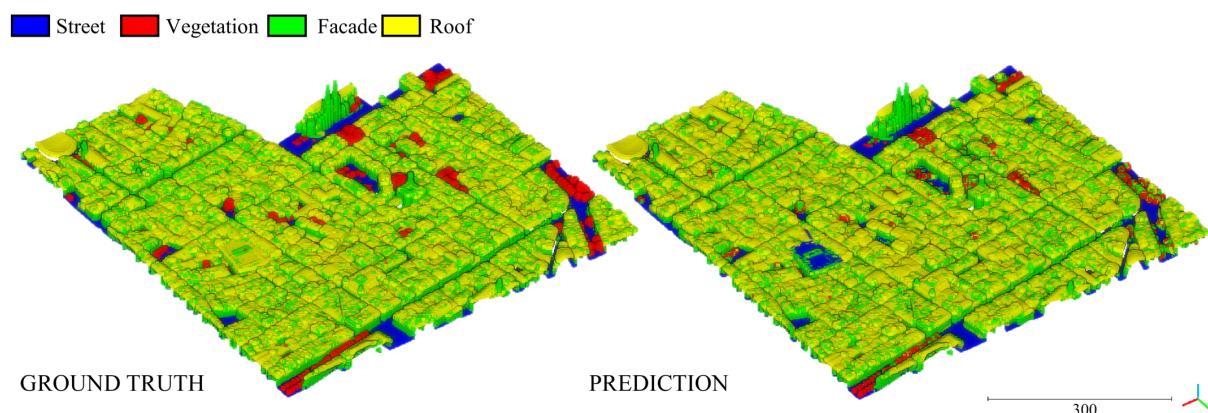
**Table 4.** Per-class F1 scores for the Dublin dataset.

F1 scores								
	Roof	Ground	Wall	Wind	Door	Table	Chair	Book
<b>A</b>	81.15	87.09	78.98	49.25	38.48	51.71	34.51	33.29
<b>B</b>	98.3	95.83	90.3	77.74	69.38	70.28	59.55	33
<b>C</b>	98.08	97.44	90.25	73.5	66.86	73.87	72.88	49.06

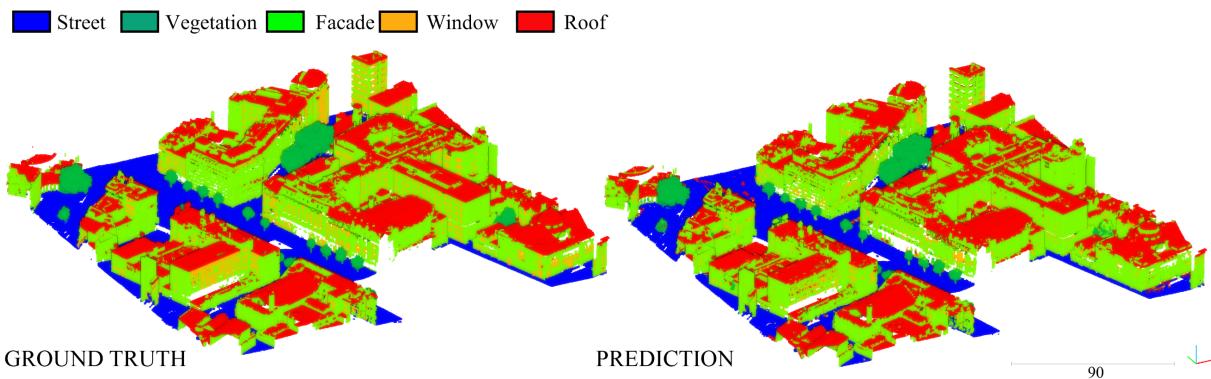
**Table 5.** Per-class F1 scores for the S3DIS dataset

F1 scores							
	Ground	Façade	Col	Arch	Vault	Wind	Mould
<b>A</b>	88.49	56.96	26.86	28.59	78.75	50.13	48.82
<b>B</b>	98.16	79.26	81.14	66.16	90.42	71.32	67.25
<b>C</b>	99.96	79.65	85.75	65.52	90.79	70.56	69.09

**Table 6.** Per-class F1 scores for the Bologna dataset.



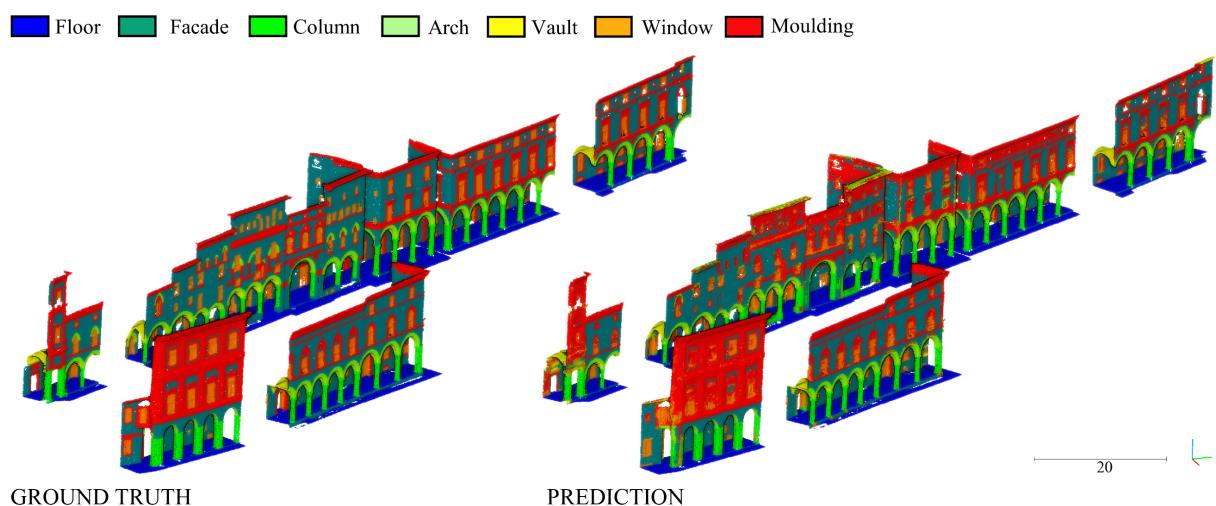
**Figure 7.** Visual comparison between the hand-annotated Bordeaux dataset (left) and the prediction achieved using the proposed method (right).



**Figure 8.** Visual comparison between the hand-annotated Dublin dataset (left) and the prediction achieved using the proposed method (right).



**Figure 9.** Visual comparison between the hand-annotated S3DIS dataset (left) and the prediction achieved using the proposed method (right).



**Figure 10.** Visual comparison between the hand-annotated Bologna dataset (left) and the prediction achieved with the proposed method (right).

## 5. CONCLUSIONS

The paper presented a combined approach, based on clustering and covariance features, for point cloud classification based on a traditional machine learning predictor. Four heterogeneous datasets were considered, featuring different type of classes and scenarios. Experiments proved that unsupervised object-based features help supervised point-based classification. Therefore, the combined method offers reduced labelling efforts, speeds up classification processing, improves accuracy, requires low computational power and is generalisable to various scenarios, making it suitable for daily work in various fields. As future work, we plan to compare the presented approach with other state-of-the-art methods for benchmarking purposes, including deep learning methods.

## REFERENCES

- Armeni, I., Sax, S., Zamir, A.R. and Savarese, S., 2017. Joint 2d-3d-semantic data for indoor scene understanding. *arXiv preprint arXiv:1702.01105*.
- Armeni, I., Sener, O., Zamir, A.R., Jiang, H., Brilakis, I., Fischer, M. and Savarese, S., 2016. 3d semantic parsing of large-scale indoor spaces. *CVPR*, pp. 1534-1543.
- Bassier, M., Vergauwen, M. and Poux, F., 2020. Point Cloud vs. Mesh Features for Building Interior Classification. *Remote Sensing*, 12(14), p.2224. doi:10.3390/rs12142224
- Bassier, M., Vergauwen, M. and Van Genechten, B., 2017. Automated classification of heritage buildings for as-built BIM using machine learning techniques. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(2W2), pp.25-30.
- Blomley, R., Weinmann, M., Leitloff, J. and Jutzi, B., 2014. Shape distribution features for point cloud analysis-a geometric histogram approach on multiple scales. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3), p.9.
- Breiman, L., 2001. Random forests. *Machine learning*, 45(1), pp.5-32.
- Bremer, M., Wichmann, V., Rutzinger, M., 2013. Eigenvalue and graph-based object extraction from mobile laser scanning point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. II-5-W2, pp. 55-60.
- Brodu, N. and Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 68, pp.121-134.
- Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., 2015. ShapeNet: An Information-Rich 3D Model Repository. *arXiv*, arXiv:1512.03012.
- Chehata, N., Guo, L. and Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 38(3), pp. 207-212.
- Chiang, M.M.T. and Mirkin, B., 2010. Intelligent choice of the number of clusters in k-means clustering: an experimental study with different cluster spreads. *Journal of Classification*, 27(1), pp. 3-40. doi:10.1007/s00357-010-9049-5
- Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T. and Nießner, M., 2017. Scannet: Richly-annotated 3d reconstructions of indoor scenes. *CVPR*, pp. 5828-5839.
- Filin, S. and Pfeifer, N., 2005. Neighborhood systems for airborne laser data. *Photogrammetric Engineering & Remote Sensing*, 71(6), pp.743-755.
- Goutte, C. and Gaussier, E., 2005. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European conference on information retrieval*, pp. 345-359, Springer, Berlin, Heidelberg.
- Grilli, E. and Remondino, F., 2020. Machine learning generalisation across different 3D architectural heritage. *ISPRS International Journal of Geo-Information*, 9(6), p. 379.
- Grilli, E., Farella, E.M., Torresani, A. and Remondino, F., 2019. Geometric feature analysis for the classification of cultural heritage point clouds. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. XLII-2/W15, 541-548.
- Guinard, S., Landrieu, L., Vallet, B., 2017. Pré-segmentation pour la classification faiblement supervisée de scènes urbaines à partir de nuages de points 3D LIDAR.
- Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M., 2020. Deep learning for 3D point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* doi:10.1109/tpami.2020.3005434
- Hackel, T., Wegner, J.D. and Schindler, K., 2016. Fast semantic segmentation of 3D point clouds with strongly varying density. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, 3, pp.177-184.
- Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N. and Markham, A., 2020. Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11108-11117.
- Landrieu, L., Raguet, H., Vallet, B., Mallet, C., Weinmann, M., 2017. A structured regularisation framework for spatially smoothing semantic labelings of 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132, pp.102-118.
- Landrieu, L. and Simonovsky, M., 2018. Large-scale point cloud semantic segmentation with superpoint graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4558-4567. doi:10.1109/CVPR.2018.00479
- Lee, I. and Schenk, T., 2002. Perceptual organisation of 3D surface points. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/A), pp.193-198.
- Lehtola, V.V., Kaartinen, H., Nüchter, A., Kaijaluoto, R., Kukko, A., Litkey, P., Honkavaara, E., Rosnell, T., Vaaja, M.T., Virtanen, J.P. and Kurkela, M., 2017. Comparison of

- the selected state-of-the-art 3D indoor scanning and point cloud generation methods. *Remote sensing*, 9(8), p.796.
- Linsen, L. and Prautzsch, H., 2001. Local versus global triangulations. Proc. of *Eurographics*, pp. 257-263.
- Luo, H., Wang, C., Wen, C., Chen, Z., Zai, D., Yu, Y., Li, J., 2018. Semantic Labeling of Mobile LiDAR Point Clouds via Active Learning and Higher Order MRF. *IEEE Trans. Geosci. Remote Sens.* Vol. 56, pp. 3631-3644.
- Matrone, F., Lingua, A., Pierdicca, R., Malinverni, E.S., Paolanti, M., Grilli, E., Remondino, F., Murtiyoso, A. and Landes, T., 2020. A benchmark for large-scale heritage point cloud semantic segmentation. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. XLIII-B2, 1419-1426.
- Munoz D, Vandapel N, Hebert M, 2008. Directional associative Markov network for 3-d point cloud classification. In: Proceedings of the International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'08). IEEE, 63–70.
- Niemeyer, J., Rottensteiner, F. and Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. *ISPRS journal of photogrammetry and remote sensing*, 87, pp.152-165.
- Nys, G.A., Poux, F. and Billen, R., 2020. CityJSON Building Generation from Airborne LiDAR 3D Point Clouds. *ISPRS International Journal of Geo-Information*, 9(9), p.521.
- Özdemir, E., Remondino, F. and Golkar, A., 2019. Aerial point cloud classification with deep learning and machine learning algorithms. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. XLII-4/W18, pp. 843-849.
- Pauly, M., Keiser, R. and Gross, M., 2003. Multi-scale feature extraction on point-sampled surfaces. In *Computer graphics forum*, Vol. 22, No. 3, pp. 281-289.
- Poux, F. and Billen, R., 2019. Voxel-based 3D point cloud semantic segmentation: unsupervised geometric and relationship featuring vs deep learning methods. *ISPRS International Journal of Geo-Information*, 8(5), p.213.
- Poux, F. and Ponciano, J.J., 2020. Self-Learning Ontology For Instance Segmentation Of 3d Indoor Point Cloud. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43(B2), pp. 309-316.
- Poux, F., Mattes, C. and Kobbelt, L., 2020. Unsupervised segmentation of indoor 3D point cloud: application to object-based classification. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 44(W1-2020), pp. 111-118.
- Stojanovic, V., Trapp, M., Richter, R., Hagedorn, B. and Döllner, J., 2018. Towards the generation of digital twins for facility management based on 3D point clouds. In *Proceeding of the 34th Annual ARCOM Conference*, Vol. 2018, pp. 270-279.
- Stojanovic, V., Trapp, M., Döllner, J. and Richter, R., 2019. Classification of indoor point clouds using multiviews. In *The 24th International Conference on 3D Web Technology*, pp. 1-9.
- Sun, Z., Xu, Y., Hoegner, L. and Stilla, U., 2018. Classification of MLS point clouds in urban scenes using detrended geometric features from supervoxel-based local contexts. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, IV-2, 271-278.
- Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K. and Li, J., 2020. Toronto-3D: A large-scale mobile lidar dataset for semantic segmentation of urban roadways. *CVPR*, pp. 202-203.
- Thomas, H., Goulette, F., Deschaud, J.E., Marcotegui, B. and LeGall, Y., 2018. Semantic classification of 3D point clouds with multi-scale spherical neighborhoods. In *2018 International conference on 3D vision (3DV)*, pp. 390-398. IEEE.
- Teruggi, S., Grilli, E., Russo, M., Fassi, F. and Remondino, F., 2020. A Hierarchical Machine Learning Approach for Multi-Level and Multi-Resolution 3D Point Cloud Classification. *Remote Sensing*, 12(16), p.2598.guo
- Toschi, I., Farella, E.M., Welponer, M. and Remondino, F., 2021. Quality-based registration refinement of airborne LiDAR and photogrammetric point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 172, pp.160-170.
- Weinmann, M., Jutzi, B. and Mallet, C., 2014. Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3), p.181.
- Weinmann, M., Jutzi, B., Hinz, S. and Mallet, C., 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, pp.286-304.
- Weinmann, M., Jutzi, B., Mallet, C., and Weinmann, M., 2017a. Geometric features and their relevance for 3d point cloud classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-1/W1, pp. 157-164.
- Weinmann, M., Hinz, S. and Weinmann, M., 2017b. A hybrid semantic point cloud classification-segmentation framework based on geometric features and semantic rules. *PFG-Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(3), pp.183-194.
- Xie, Y., Tian, J. and Zhu, X.X., 2020. Linking points with labels in 3D: A review of point cloud semantic segmentation. *IEEE Geoscience and Remote Sensing Magazine*, 8(4), pp.38-59.
- Yang, B., Dong, Z., Liu, Y., Liang, F., Wang, Y., 2017. Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126, pp. 180-194.
- Zolanvari, S.M., Ruano, S., Rana, A., Cummins, A., da Silva, R.E., Rahbar, M. and Smolic, A., 2019. DublinCity: Annotated LiDAR point cloud and its applications. *arXiv preprint arXiv:1909.03613*.