# Urban point cloud classification with automated processing based on deep learning open-source solutions

Marcello La Guardia [1], Andrea Masiero [2,3], Valentina Bonora [1], Adriano Alessandrini [1]

[1] Department of Civil and Environmental Engineering (DICEA), University of Florence, Florence, Italy –
(marcello.laguardia, valentina.bonora, adriano.alessandrini)@unifi.it;
[2] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padua, Italy – andrea.masiero@unipd.it;
[3] Department of Land, Environment, Agriculture and Forestry (TESAF), University of Padua, Italy;

**Keywords:** Point Clouds, Deep Learning, LiDAR, Mobile Mapping, Mobile Laser Scanning.

**Abstract**

The recent development of technology in the field of point cloud acquisition in urban environment led experts to consider new needs in terms of big data management in Geomatics. MMS (Mobile Mapping Systems) allows to acquire in real time large number of points in seconds using mobile vehicles. On the other hand, the development and spread of DL (Deep Learning) CNN (Convolutional Neural Networks) applications for classification of 3D geospatial data allows to start first experimentations in urban environments. Considering this scenario, this work presents a framework for the segmentation and classification of urban point cloud datasets based on the integration of some machine and deep learning tools, namely RANSAC, Euclidean Cluster Extraction, PointNet++ and Support Vector Machines (SVM) algorithms, developed with open-source technologies. The case of study considered in this experimentation regards an urban point cloud dataset obtained by an MMS acquisition along the roads of Sesto Fiorentino (FI). The integration of different algorithms based on Python libraries allowed to obtain fast processing performance, optimizing the results and offering a low-cost and fast solution for experts involved in 3D geospatial data extraction from point cloud MMS acquisitions.

## 1. Introduction

Recent advances in computer science allowed researchers to focus the interest on three dimensional representations of the real environment, being such task of interest for a wide range of fields and applications, such as virtual reality, geomatics, robotics and autonomous driving (Blaszczak-Bak et al., 2024; Fernandes et al., 2022). In this scenario, point clouds play a key role in 3D environment representation, being the *raw 3D information* provided by many 3D acquisition sensors: nowadays, recently developed acquisition methods, such as mobile laser scanning (Toschi et al., 2015; Aminti et al., 2022; Aricò et al., 2023), with LiDAR sensors, and digital photogrammetry ensure fast acquisition of dense point clouds, characterized by rich geometric content information (Dominici et al., 2017; Masiero et al., 2018). In particular, MMSs (Mobile Mapping Systems) employ high-end remote/proximity sensing and localization sensors (GNSS and inertial measurement unit) that allow the real-time construction of high-detailed 3D point clouds (acquisition of millions of points per second) of the environment during vehicle movement (Wong et al., 2021).

Information extraction from point cloud data, garnered from state-of-the-art LiDAR sensors, stands as a crucial enabler for understanding the complexity of a device's spatial neighbourhood. The employment of point cloud information requires the development of effective strategies in order to extract such information from the huge amount of collected data. Important tasks within such information extraction methods are for instance segmentation, that is the ability of separating the points associated to different objects in a point cloud, and classification, used in order to identify which types of objects have been segmented. Such tasks, and in particular classification, are frequently implemented by using Deep Learning (DL) tools (Zhang et al. 2023).

To be more specific, a suitable interpretation of the surrounding environment usually requires the execution of the following activities: 3D data acquisition, 3D point cloud segmentation, object detection and classification (Li et al., 2022). Given the high acquisition rate, up to millions of points per second in recent devices, LiDAR datasets are usually voluminous, providing only an incomplete description of the surrounding objects. These factors make obtaining effective segmentation of such data into semantically meaningful categories a formidable challenge.

For this reason, the increasing demand for robust point cloud segmentation algorithms led to the employment of Artificial Intelligence (AI)-based methods, and, in particular, DL approaches based on 3D Convolutional Neural Networks (CNN) (Wang et al., 2019).

A CNN is a type of artificial neural network designed specifically for tasks related to pattern recognition and computer vision. During the last decade, CNNs have already largely been successfully applied to tasks involving 2D data, such as images, but they are also used in various other domains, such as natural language processing and speech recognition (Li et al., 2022).Extending their use to 3D data is not so obvious, and only some recently developed approaches allowed to obtain effective results (Qi et al., 2017a, Qi et al., 2017b, Quian et al., 2022).

The application of DL algorithms finalized to segmentation, detection and classification led specialists to focus the interest on several problems and challenges that regard the data acquisition and the DL-based models. Geospatial data typically provided by mobile laser scanning systems are typically characterized by a diversified point density (that depends on the distances between LiDAR sensor and objects), noise, and the incomplete nature of mobile acquisitions: all such factors contribute to making MMS point cloud understanding quite challenging (Kumar et al., 2019). In addition to such criticalities, processing big data acquired by MMS requires high computation and time-consuming effort, in particular for DL operations (Liu et al., 2019).

The classification of urban road inventory represents one of the main challenging topics where supervised and unsupervised DL algorithms are tested (Ma et al., 2024). Nowadays several sensor acquisition technologies are widely used to extract point clouds data from urban environment analysis, and several datasets are shared on the web for object classification purposes (Bayrak et al., 2024).

The application of DL technologies for segmentation and classification of elements inside the urban environment from MMS acquisitions represents a research topic that involves the realm of urban mobility for Smart Cities. In fact, this study is also included in such field, being a scientific contribution within the Italian Sustainable Mobility Centre (MOST). MOST is a center of research developed with the Next Generation EU recovery initiative, aimed at the establishment of the Italian leadership for inclusive and sustainable mobility solutions (Alberti et al., 2023).

In particular, this work focuses the interest on the application of DL-based models that employ CNN, working on urban point cloud dataset generated from a mobile mapping acquisition. The procedure followed in this study considered as testing dataset a mobile mapping campaign along the streets of the city centre in Sesto Fiorentino (Florence, Italy).

Direct segmentation and classification based on CNN resulted to be a quite time-consuming activity, hence this work experimented an integrated approach for the fast segmentation and classification of MMS urban point clouds, employing open-source technology.

The dataset acquired in Sesto Fiorentino was used for testing the proposed integrated approach, which is summarized by the following workflow:
- point cloud segmentation, using Python scripts implementing ground detection and Euclidian distance-based clustering
- classification of the segmented elements with PointNet++
- refinement of PointNet++ results using Support Vector Machine (SVM) on certain critical cases.

The overall framework represents a low cost and fast solution for urban point cloud processing, obtaining acceptable results in terms of quality.

The next paragraph, materials and methods, presents the proposed procedure in detail, then, the following section, discussion and results, analyses the goals achieved in this experimentation, and, finally, some conclusions are drawn in order to resume the research objectives and results, along with highlighting on some possible future integrations on this subject.

## 2. Materials and methods

This work aims at testing a workflow in order to properly segment and classify objects in a 3D point cloud, acquired by a MMS survey of an urban environment (Figure 1), using open software tools.

The instrumentation used for the acquisition on the streets of the city centre of Sesto Fiorentino (FI) was a ChcNav AU20 LiDAR installed on a car mount kit, including a Ladybug5+ high-resolution panoramic camera, a GNSS receiver and an Inertial Measurement Unit (IMU).

The extension of the acquired point cloud was sufficient to contain different urban environment combinations (low and high-density buildings, parks, etc) with different geometric features and relationship between the elements.

First, the original point cloud was partitioned into smaller point clouds, obtained by subsampling and cropping the original one, setting bounding boxes of 30×30×50 m (Figure 2) and spatially uniformly subsampling at 0.025 m.

Then, segmentation was carried out employing integrated freely available segmentation tools, based on geometry-based algorithms using Python libraries. In particular Open3D opensource Python libraries were loaded into an Anaconda developing environment, where the Python script was executed. The script considered the segmentation as a two-step process: first, planar surfaces were detected using a RANSAC (Random Sample Consensus) algorithm, and, in particular, ground was removed from the point cloud, then, the remaining points were segmented in different objects by means of Euclidean Cluster Extraction (Figure 3).
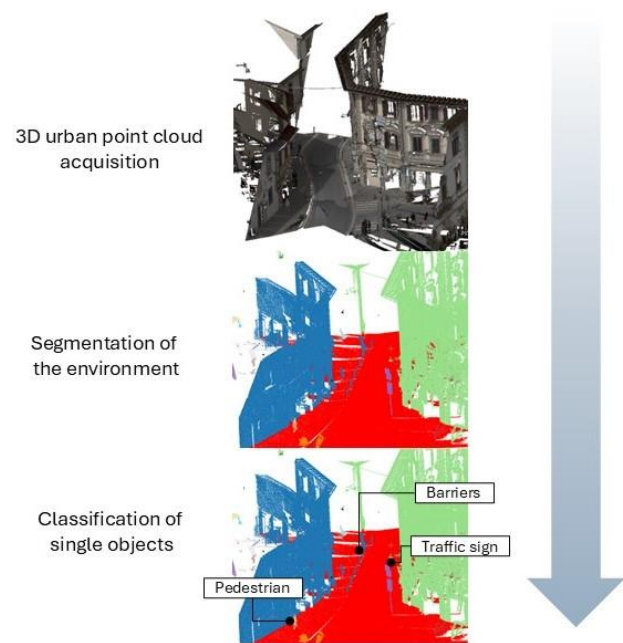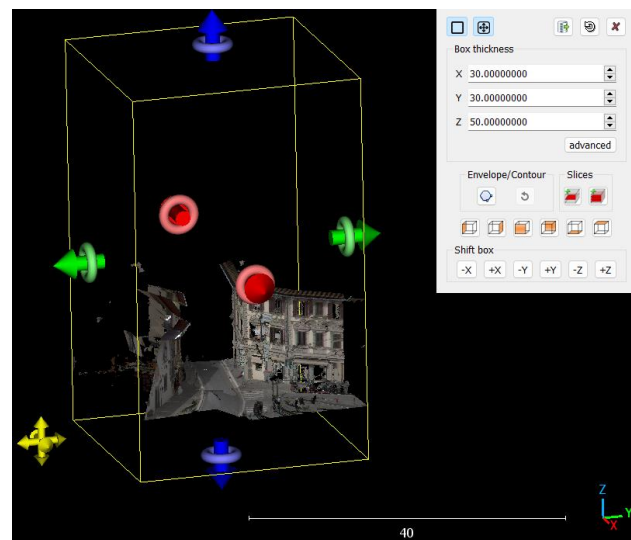


Figure 1. The workflow



Figure 2. Example of bounding box extraction of the dataset.

RANSAC represents a robust statistical method, widely used in computer vision, machine learning and data fitting applications, that allows to extract parameters of a mathematical model on the basis of observed data that contain outliers (Schabel et al., 2007; Yang et al., 2022). It allows to identify the larger subset of inliers, that consist of data points that fit the model within a fixed tolerance.

The Euclidean Cluster Extraction, instead, allowed the partitioning of the remaining data points employing the information that regards their Euclidean distribution in space (Liu et al., 2021).
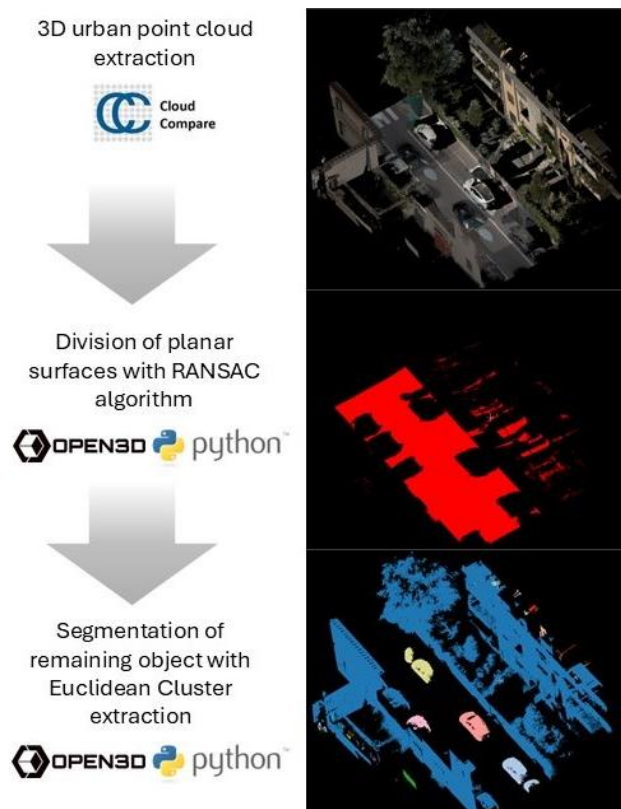


Figure 3. The segmentation process.

In this way, most of the dominant planar surfaces in the scene were removed (ground and building facades), whereas the remaining objects were quickly separated. The segmentation was tested on different urban environment cases present in the dataset of Sesto Fiorentino (Figures 4-5).
The classification of the urban objects extracted from segmentation processing was based on PointNet++ CNN Deep Learning algorithm using Pytorch libraries (Yan et al., 2021).
The original library of PointNet++ (Yan et al., 2021) was designed for the classification of generic objects provided by Modelnet database. We adapted the code to be suitable for the classification of urban environment' objects, allowing the possibility to edit the training and the testing operations with our dataset. Both training and testing operations on classification were fast, because the algorithm worked on subsampled and segmented point clouds.
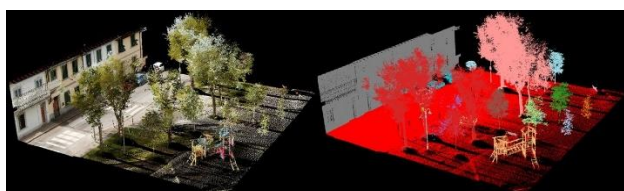


Figure 4. Example of point cloud segmentation in a park environment
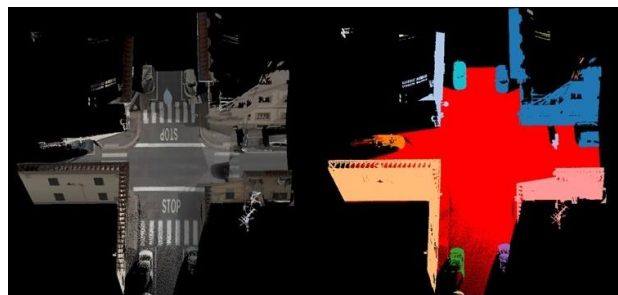


Figure 5. Example of point cloud segmentation in a crossroads with buildings facing it.

The aim of the experimentation was to use a general dataset of urban objects generated from MMS point cloud acquisition on the road, and then to test the algorithm on the objects segmented from the dataset of Sesto Fiorentino. Instead, in order to increase the training dataset size, training objects were provided both from the Sesto Fiorentino acquisition and another dataset acquired using MMS in similar conditions (De Deuge et al., 2013): this allowed to build a training dataset not strictly referred to the urban environment of Sesto Fiorentino. Training and test datasets were clearly separated, i.e. training objects extracted from the Sesto Fiorentino dataset were clearly not used in the test dataset. The categories considered for the classification were: barrier, car, motorcycle/bicycle, pedestrian, pillar, light pole, traffic sign, tree (see also Table 1).

| Category | Number | format | example |
|---|---|---|---|
| barrier | 14 | .txt (x,y,z, Nx,Ny,Nz) | |
| car | 47 | .txt (x,y,z, Nx,Ny,Nz) | |
| motorcycle/ bicycle | 24 | .txt (x,y,z, Nx,Ny,Nz) | |
| pedestrian | 22 | .txt (x,y,z, Nx,Ny,Nz) | |
| pillar | 22 | .txt (x,y,z, Nx,Ny,Nz) | |

| | | | |
|---|---|---|---|
| light pole | 19 | .txt (x,y,z, Nx,Ny,Nz) |  |
| traffic sign | 25 | .txt (x,y,z, Nx,Ny,Nz) | |
| tree | 22 | .txt (x,y,z, Nx,Ny,Nz) | |
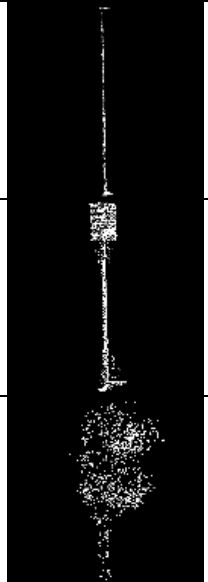
Table 1. Classes considered for classification.

Then, the trained network was tested on a dataset composed by 13 objects for each class. The confusion matrix in Figure 6 shows the results obtained after PointNet++ classification.

| confusion matrix | | Real label | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | barr. | car | moto/byc. | ped. | pillar | pole | tr. sign | tree |
| **Model pred.** | barr. | 6 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | car | 3 | 13 | 3 | 2 | 0 | 3 | 0 | 3 |
| | moto/byc. | 1 | 0 | 6 | 1 | 0 | 0 | 1 | 0 |
| | ped. | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| | pillar | 1 | 0 | 0 | 1 | 8 | 0 | 1 | 0 |
| | pole | 2 | 0 | 1 | 2 | 1 | 5 | 2 | 0 |
| | tr. sign | 0 | 0 | 1 | 2 | 4 | 5 | 9 | 4 |
| | tree | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 |

Figure 6. Confusion matrix after PointNet++ classification.

In order to improve the classification results, a final classification step has been introduced, based on the use of SVM classifier. SVM allowed to explicitly introduce in the classification step the use of certain specific geometric features of the considered objects, i.e. the object sizes, extracted for instance from the corresponding bounding box.

The introduction of SVM algorithm in the workflow improved the results obtained before, using only PointNet++ classification, as shown in figure 7. The details of the results and the comments will be object of the next paragraph.

| confusion matrix | | Real label | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | barr. | car | moto/byc. | ped. | pillar | pole | tr. sign | tree |
| **Model pred.** | barr. | 6 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | car | 1 | 13 | 2 | 2 | 0 | 3 | 0 | 1 |
| | moto/byc. | 3 | 0 | 8 | 1 | 0 | 0 | 1 | 2 |
| | ped. | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| | pillar | 1 | 0 | 0 | 0 | 8 | 0 | 1 | 0 |
| | pole | 2 | 0 | 1 | 2 | 2 | 7 | 2 | 0 |
| | tr. sign | 0 | 0 | 0 | 2 | 3 | 3 | 9 | 4 |
| | tree | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 |

Figure 7. Confusion matrix after PointNet++ classification and SVM application.

## 3. Discussion and results

The proposed processing procedure integrates different open-source algorithms, hence resulting to be low cost, nevertheless it allowed to realize fast segmentation and classification steps, with good results on the considered tests.

Among the criticalities highlighted in such experimental tests, the segmentation process behaviour was not constant, but varied depending on the features of the segmented urban environment. For instance, considering the separation of the planar surfaces using RANSAC algorithm, the sensitivity should be properly designed to cover the sidewalks, and it changed on the basis on their height and the slope of the road. Instead, considering the segmentation of the remaining environment in single objects using Euclidean Cluster Extraction, sometimes either divided single objects in more elements (over-segmentation) or grouped more objects (under-segmentation) as seen in figure 8.
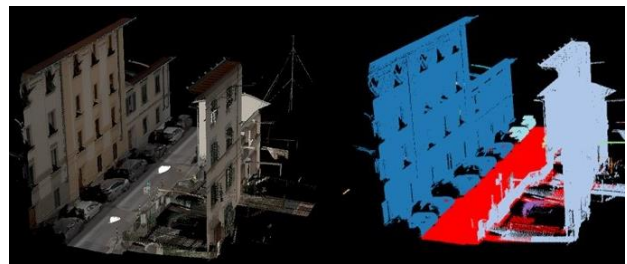


Figure 8. Example of under segmentation of urban environment.

A careful analysis of the results shown in Figure 6 reveals that PointNet++ classification worked quite well for objects characterized by well-defined geometries as cars. Instead, it highlights also some weaknesses, in particular categories of objects with similar shapes were sometimes confused: for instance, poles and pillars were sometimes confused with traffic signs (Figure 9).
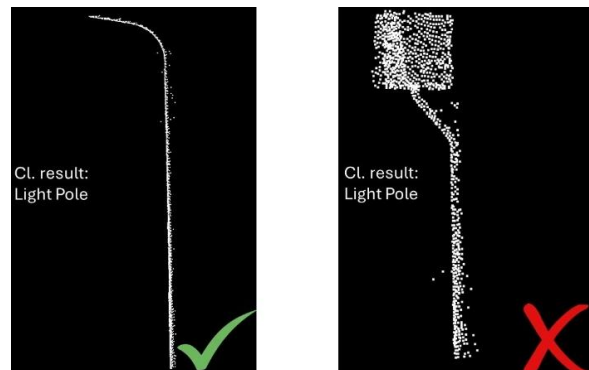


Figure 9. Example of two test classifications. The first, on the left, is the correct classification of a light pole, the second, on the right, is the wrong classification of a traffic sign confused for a light pole by the DL algorithm.

The introduction of an additional SVM-based classification step allowed to improve PointNet++ results (Figure 10), as shown in Figure 6 and 7. This additional step limited classification errors between classes of objects of similar shape.

Another weakness is represented by the scarce number of objects involved in training and testing dataset. In fact, this work shows only some preliminary results, obtained by using only a limited number of objects, due to the time-consuming activity of extraction and quality check of urban point cloud objects.
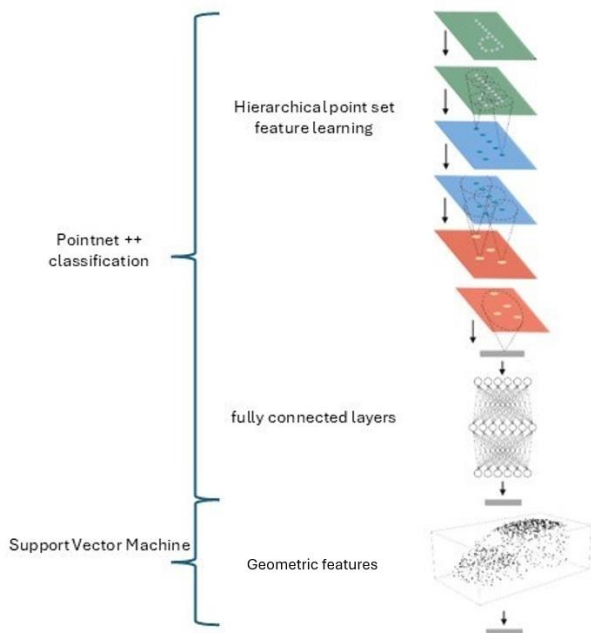
Figure 10. The process of classification.

## 4. Conclusions

The research presented in this paper shows a low-cost framework for the segmentation and classification of point clouds acquired from MMS acquisition on urban environments. The employed technologies based on Python language allowed to build an overall framework for the segmentation and classification of urban point cloud dataset, employing open-source libraries. The obtained results, after the integration of Pointnet++ CNN and SVM, highlight acceptable results but underlines possibilities of future improvements of the work.

Our future works foresees the use of much larger training and testing datasets, in order to obtain more reliable results in terms of both training and validation.

Then, the use of alternative deep and machine learning methods will be considered in order to further improve the classification step.

Independently on such potential improvements, the use of open-source software libraries will be kept as a strength of the proposed approach: indeed, despite requiring some programming skills, it allows a full customization of the proposed solution.

Furthermore, the proposed integrated approach for point cloud segmentation and classification appears as a reasonable solution to reduce the computational burden, implementing a fast and low-cost framework for urban environment segmentation and classification avoiding the use of expensive commercial software solutions and preferring the employment of free Python libraries.

## References

Alberti, F., Alessandrini, A., Bubboloni, D., Catalano, C., Fanfani, M., Loda, M., Marino, A., Masiero, A., Meocci, M., Nesi, P., Paliotto, A., 2023: Mobile mapping to support an integrated transport-territory modelling approach. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-1/W1-2023, 1–7. DOI: 10.5194/isprs-archives-XLVIII-1-W1-2023-1-2023.

Aminti, P., Bonora, V., Mugnai, F., Tucci, G., 2022: Arno Riverbed Survey in Florence 1935 - 2019: From the Integrated Survey to the Geomatic Monitoring. *Geomatics and Geospatial Technologies. ASITA 2021. Communications in Computer and Information Science*, vol 1507. Springer, Cham. DOI: 10.1007/978-3-030-94426-1_6

Aricò, M., La Guardia, M., Lo Brutto, M., Rappa, E. M., Vinci, C., 2023: Mobile Mapping for Cultural Heritage: the Survey of the Complex of St. John of The Hermits in Palermo (Italy). *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XLVIII-1/W1-2023, 25–32. DOI: 10.5194/isprs-archives-XLVIII-1-W1-2023-25-2023

Bayrak, O. C., Ma, Z., Farella, E. M., Remondino, F., Uzar, M., 2024: ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLVIII-2-2024, 25–32. DOI: 10.5194/isprs-archives-XLVIII-2-2024-25-2024

Blaszczak-Bak, W., Masiero, A., Bąk, P., Kuderko, K., 2024: Integrating Data from Terrestrial Laser Scanning and Unmanned Aerial Vehicle with LiDAR for BIM Developing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-1-2024, 25–30. DOI: 10.5194/isprs-archives-XLVIII-1-2024-25-2024

De Deuge, M., Quadros, A., Hung, C., Douillard, B, 2013: Unsupervised Feature Learning for Classification of Outdoor 3D Scans. *Australasian Conference on Robotics and Automation (ACRA)*, 2013.

Dominici, D., Alicandro, M., Massimi, V., 2016: UAV photogrammetry in the post-earthquake scenario: case studies in L'Aquila. *Geomatics, Natural Hazards and Risk*, 8(1), 87–103. DOI: 10.1080/19475705.2016.1176605

Fernandes, D., Silva, A., Névoa, R., Simões, C., Gonzalez, D., Guevara, M., Novais, P., Monteiro, J., Melo-Pinto, P., 2021: Point-cloud based 3D object detection and classification methods for self-driving applications: A survey and taxonomy *Information Fusion*, Volume 68, 2021, Pages 161-191, ISSN 1566-2535. DOI: 10.1016/j.inffus.2020.11.002.

Kumar, B., Pandey, G., Lohani, B., Misra, S. C., 2019: A multi-faceted CNN architecture for automatic classification of mobile LiDAR data and an algorithm to reproduce point cloud samples for enhanced training, *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, pp. 80-89, ISSN 0924-2716. DOI: 10.1016/j.isprsjprs.2018.11.006

Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J., 2022: A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects, *IEEE Transactions on Neural Networks and Learning Systems*, 33(12), pp. 6999-7019. DOI: 10.1109/TNNLS.2021.3084827.

Liu, H., Song, R., Zhang, X., Liu, H., 2021: Point cloud segmentation based on Euclidean clustering and multi-plane extraction in rugged field. *Meas. Sci. Technol*., 32 095106. DOI: 10.1088/1361-6501/abead3

Liu, Y., Fan, B., Xiang, S., Pan, C., 2019: Relation-Shape Convolutional Neural Network for Point Cloud Analysis, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8895-8904.

Ma, Z., Bayrak, O. C., Remondino, F., 2024: Automatic Point Cloud Classification of Under-Represented Pole-Like Objects Based On Hierarchical Directed Graph, *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, Athens, Greece, 2024, pp. 8307-8311. DOI: 10.1109/IGARSS53475.2024.10641299.

Masiero, A., Guarnieri, A., Fissore, F., Piragnolo, M., Pirotti F., Vettore, A., 2018: TLS and photogrammetry for 3D modelling of a low relief: case study of ancient archive, Palazzo Bo, Padua. *Metrology for Archaeology and Cultural Heritage (MetroArchaeo)*, pp. 431-436. DOI: 10.1109/MetroArchaeo43810.2018.9089777

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2017a. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 652-660.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017b. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30.

Qian, G., Li, Y., Peng, H., Mai, J., Hammoud, H., Elhoseiny, M., Ghanem, B., 2022: Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *Advances in Neural Information Processing Systems*, 35, 23192-23204.

Schnabel, R., Wahl, R., Klein, R., 2007: Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum*, 26, 214-226. DOI: 10.1111/j.1467-8659.2007.01016.x

Toschi, I., Rodríguez-Gonzálvez, P., Remondino, F., Minto, S., Orlandini, S., Fuller, A., 2015: Accuracy evaluation of a mobile mapping system with advanced statistical methods. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-5/W4, 245–253. DOI: 10.5194/isprsarchives-XL-5-W4-245-2015.

Wang, Y., Sun, Y., Liu, Z., E. S., Sanjay, Bronstein, M. M., Solomon, J. M., 2019: Dynamic Graph CNN for Learning on Point Clouds. *ACM Trans. Graph*. 38, 5, Article 146, 12 pages. DOI: 10.1145/3326362

Wong, K., Gu, Y., Kamijo, S., 2021: Mapping for Autonomous Driving: Opportunities and Challenges, *IEEE Intelligent Transportation Systems Magazine*, 13(1), pp. 91-106. DOI: 10.1109/MITS.2020.3014152.

Yan, X., Gao, J., Li, J., Zhang, R., Li, Z., Huang, R., Cui, S., 2021: Sparse Single Sweep LiDAR Point Cloud Segmentation via Learning Contextual Shape Priors from Scene Completion. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(4), 3101-3109. DOI: 10.1609/aaai.v35i4.16419

Yang, L., Li, Y., Li, X., Meng, Z., Luo, H., 2022: Efficient plane extraction using normal estimation and RANSAC from 3D point cloud. *Computer Standards & Interfaces*, Volume 82, 2022, 103608, ISSN 0920-5489. DOI: 10.1016/j.csi.2021.103608

Zhang, H., Wang, C., Tian, S., Lu, B., Zhang, L., Ning, X., Bai, X., 2023: Deep learning-based 3D point cloud classification: A systematic survey and outlook. *Displays*, Volume 79, 102456, ISSN 0141-9382. DOI: 10.1016/j.displa.2023.102456