remote sensing

MDPI

Article

Semantic Mapping of Landscape Morphologies: Tuning ML/DL Classification Approaches for Airborne LiDAR Data

Marco Cappellazzo [1], Giacomo Patrucco [1,*], Giulia Sammartano [1,2], Marco Baldo [3] and Antonia Spanò [1,2]

[1] LabG4CH—Laboratory of Geomatics for Cultural Heritage, Department of Architecture and Design (DAD), Politecnico di Torino, Viale Mattioli 39, 10125 Turin, Italy; marco.cappellazzo@polito.it (M.C.); giulia.sammartano@polito.it (G.S.); antonia.spano@polito.it (A.S.)
[2] Polito FULL | The Future Urban Legacy Lab, OGR Tech., Corso Castelfidardo, 22, 10128 Turin, Italy
[3] Turin Support Unit, CNR IRPI—Research Institute for Geo-Hydrological Protection, Italian National Research Council, Strada delle Cacce 73, 10135 Turin, Italy; marco.baldo@irpi.cnr.it
* Correspondence: giacomo.patrucco@polito.it

Abstract: The interest in the enhancement of innovative solutions in the geospatial data classification domain from integrated aerial methods is rapidly growing. The transition from unstructured to structured information is essential to set up and arrange geodatabases and cognitive systems such as digital twins capable of monitoring territorial, urban, and general conditions of natural and/or anthropized space, predicting future developments, and considering risk prevention. This research is based on the study of classification methods and the consequent segmentation of low-altitude airborne LiDAR data in highly forested areas. In particular, the proposed approaches investigate integrating unsupervised classification methods and supervised Neural Network strategies, starting from unstructured point-based data formats. Furthermore, the research adopts Machine Learning classification methods for geo-morphological analyses derived from DTM datasets. This paper also discusses the results from a comparative perspective, suggesting possible generalization capabilities concerning the case study investigated.

Keywords: airborne LiDAR; deep learning; geometric filter; support vector machines; random forest

check for updates

Citation: Cappellazzo, M.; Patrucco, G.; Sammartano, G.; Baldo, M.; Spanò, A. Semantic Mapping of Landscape Morphologies: Tuning ML/DL Classification Approaches for Airborne LiDAR Data. *Remote Sens.* **2024**, *16*, 3572. https://doi.org/10.3390/rs16193572

Academic Editor: Kevin Tansey

Received: 5 August 2024
Revised: 10 September 2024
Accepted: 22 September 2024
Published: 25 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Landscape morphologies are now usually considered a concept linked to the physical constitution of a portion of land and the processes that create it [1]. In other words, they correspond to the landscape's material formation, including its shaping evolution, connected to the social and cultural structures with which it is associated. Although this research is oriented towards looking for anthropogenic forms in the landscape, it is possible to consider that the research tools, both concerning primary data, airborne LiDAR (Light Detection and Ranging), and, above all, those of semi-automatic semantic detection of the signs themselves (unsupervised and supervised macro-class classification), operate in the scenario of the definition and mapping of the microtopography of the land.

The concept of microtopography is also complex, as it considers terrain parameters such as slope, aspect, flow path, ruggedness index, wetness index, and curvature that characterize the surface. Still, in general, the microtopography studies in the literature converge on the preeminent consideration of the variation in elevation of the ground surface between 1 cm and 1 m [2]. Starting from this consideration, the studies focus on the use of both active (LiDAR) and passive (digital photogrammetry) remote sensing (RS) technologies for the basic definition of microtopographic characteristics starting from highly accurate elevation data over relatively large areas.

Therefore, DEMs are considered in the literature as the characteristic 3D reconstruction, derived from aerial acquisition by sensors equipped on vehicles (planes, helicopters, drones, etc.), and can be derived from LiDAR and photogrammetric methods. Their use

is traditionally related to the most varied application sectors, reducing time acquisition and ensuring high and accurate productivity, especially for urban 3D modeling or terrain morphology analysis.

Various studies can be cited among the environmental sectors most influenced by adopting automatic remote sensing techniques. These research works relate to hydrographic networks and morphological variations of watercourses concerning floods [3]; forest management applications for forestry censuses, canopy height, and biomass estimation [4,5]; coastal protection [6]; glacier monitoring [7]; and recognition and monitoring of avalanches and landslides [8,9]. Moreover, even Cultural Heritage (CH) domain disciplines related to archaeological studies have benefited from RS technology applications, specifically for airborne DEM data classification [10].

Concerning digital photogrammetric techniques, Unmanned Aerial Systems (UAS) have been recently recognized as a consolidated approach [11]. The theoretical and methodological developments in the direction of UAS photogrammetry have been determined by the need to continuously improve the density and accuracy of the results. Moreover, balancing the full exploitation of the radiometric data with the necessity to cover large surface extensions is equally essential. Integrating airborne far-range and close-range photogrammetry approaches can fully satisfy the needs of 3D documentation at the scale of archaeological landscapes highly characterized by human impact [12,13]. UAS photogrammetry also proved very efficient and effective in situations of different vegetation complexity and soil type, again in the environmental field, such as examples of peatland demonstrated [14]. Furthermore, the evolution of technologies has led to the consideration of integrating LiDAR sensors using UAS vectors. These compact solutions now optimize the on-flight stability issues and improve the estimation of the navigation path with real-time kinetic (RTK) GNSS. Moreover, the potential to estimate the ground surface is comparable, if not superior, to that obtainable using SfM-based UAS photogrammetry [15]. The recent diffusion of commercial UAS platforms equipped with compact LiDAR sensors provides cost-effective solutions for experimentation on landscape and heritage high-scale 3D mapping [16,17]. Despite the success of image/range-based UAS technology due to its low cost, the superiority of Airborne Laser Scanner (ALS) technology for detecting earthworks in forested areas has been well-established for a long time. This superiority is particularly evident when using a full waveform recording system. It is crucial to distinguish between filtering issues and the effective potential of the sensor for archaeological analysis applications, as highlighted in Ref. [18].

Today, automation in classification tasks acquires crucial importance both in the field of geospatial information aimed at urban contexts and 3D city models [19] and also in the vast application fields that search for artificial landforms using primary data consisting of both remotely sensed images and high-resolution DEMs [20]. Consistent and constant updating is provided to process 3D and 2.5D unstructured data. The semantic classification of these increasingly common data leads to a higher accuracy possibility in the digital twinning of real-world objects, whether they are anthropogenic artifacts or natural terrain features. Further detailed examination of classification approaches is addressed in Section 1.1.

In this framework, the presented research aims to support the documentation of artificial terrain shapes by automatically classifying high-scale aerial LiDAR survey data. Yet, this current study deals with the second step of previous research and, presently, focuses precisely on the classification phase. The first phase, according to the best practices of geomatic documentation and archaeological research, validated the LiDAR dataset acquired from a helicopter by comparing the results with those obtained on the ground using various and integrated methods such as terrestrial laser scanning and SfM-based UAV photogrammetry [21]. In addition, the previous work also anticipated the study of geomorphological analyses and the use of semi-automatic analysis of the DSM from ALS data, both of which are now extensively developed.

In this research, the proposed workflow focuses on studying the Spina Verde Park area in Como, Lombardia, Italy. The case study pertains to a landscape Cultural Heritage

context: the area is in a hilly, densely forested region (Figure 1) in the southern part of Como Municipality, surrounded by the city's urban edges.



Figure 1. The location of the case study. The area is located in the Lombardia region in the northern part of Italy. Specifically, the Spina Verde Park is embedded in the Como municipality. Aerial view of the north hill of the park, facing the southern edge of Lake Como (a) (source: authors).

Numerous geoarchaeological sites, linked to a network of proto-historic settlements already active during the first millennium BCE, can be found in this area. Moreover, the hiking paths of the park, vital for accessing archaeological sites and enhancing the park's accessibility, are tangible traces of the ancient communication routes between the various settlements. The manuscript is organized as follows: Section 1.1 deepens the related works on LiDAR applications and DEM processing based on 3D data or imagery classification approaches, particularly comparing supervised and unsupervised methods. Section 2, Materials and Methods, describes the airborne LiDAR dataset acquisition and preprocessing

and presents the classification approaches for Digital Terrain Model (DTM) generation: filtering, deep learning (DL) strategies, and Machine Learning (ML) classification exploiting multiband raster composition. Section 3, Results and Section 4 Discussion, methodically describes the classification accuracy of the MLCs and the DLMs, evaluating and discussing the comparison of predictions with ground truth labels. Conclusions and future perspectives follow in Section 5.

1.1. Related Works

The literature has highlighted that employing ALS data to structure semantic information is a powerful technology, especially within densely forested areas. In such a context, the efficient enhancement of landscape morphology investigation is strictly determined by tuning the methodological approach. Specifically, good domain-related knowledge and proper classification of the considered features are crucial aspects of this task [22]. The discriminating factors related to the classification approach for deriving DTMs are mainly related to the nature of preliminary unstructured points and the complexity of the context and their microtopographic features [23]. The definition of the semantic content of unstructured data is fundamental for interpreting and implementing surface object detection from digital elevation models (DEMs). The American Society of Photogrammetry and Remote Sensing (ASPRS) defines LAS specifications, describing the characteristics of the LiDAR sensor, whose recording returns data from the pulse embedded in the point cloud *.las format, as scalar fields, which can be useful for classification purposes [24].

In the framework of DSM classification, filtering terrain, and off-ground objects, the literature has largely identified and tested a wide range of LiDAR-based methods and applications [25], providing reference methods and integrated strategies for DTM generation based on filtering approaches [26]. More recently, specifically for LiDAR-based DSMs, using data richness derived from the full waveform LiDAR has been demonstrated as an improved and promising approach [27]. Defining and extracting land-surface parameters are the primary tasks for any classification workflow implementation [28]. The richness of the semantic content of unstructured data plays an important role in point classification for reliable labeling of 3D points at different scales [29]. Integrating semantic information from image-based data, as in Ref. [30], contributes to workflow efficiency.

As introduced, the heritage and landscape domain has greatly benefited over the last decades from using remote sensing approaches to investigate and interpret anthropogenic evidence on terrain morphology. In this context, LiDAR data potential has been largely explored from both terrestrial and aerial points of view [31–33]. In archaeological mapping, airborne DSM data have been used also for surface filtering and terrain extraction for 3D mapping and anthropic settlement analysis [34]. Further, surface reconstruction from DSM and DTM classification is fundamental for extracting anthropogenic traces [35]. In Ref. [36], multiple visualizations of the DSM and multispectral data have been used to analyze and cluster the distribution of ancient Maya urban settlements in Mexico and Guatemala. Ref. [17] proposes a pipeline for 3D semantic segmentation of UAS-based LiDAR data for microrelief classification and extraction of ancient structure traces.

In recent epochs, much attention has been devoted to the study of Artificial Intelligence's (AI) contribution to optimizing classification pipelines and efficient learning tasks. Numerous Machine Learning (ML) and deep learning (DL) methods have been developed and tested to increase automation and ensure classification workflow accuracy. However, the countless variety of LiDAR or image-based methodological approaches developed in the literature in the past decades demonstrates that integrated approaches based on multi-source data prove to be a successful approach [27].

The following paragraphs specify the two main categories of ALS data classification that the authors considered: image-based classification approaches using raster data and a 3D unstructured point cloud.

1.1.1. ML Classification Applications Based on Imagery Data

Although the automatic learning approaches have been explored in research fields involving raster-based data and 3D data processing, imagery classification on optical data and grid DSM framework have been more comprehensively explored, primarily based on remote sensing applications. The necessity to enhance the automation of image classification tasks led to the development of various methodologies, each based on the specific criteria assumed by different algorithm applications.

Considering the different categories of image classification methods deeply examined by Ref. [37], the high efficiency of non-parametric, per-pixel, and supervised strategies for remotely sensed imagery and DSM has been established. Research of the literature demonstrates that AI techniques based on Machine Learning classifiers (MLCs) are particularly suitable for geospatial imagery [38]. However, some recent studies have investigated ML classification techniques adopting an integration strategy on DSM using optical sensor data and ALS 3D data [17,39,40].

Among the variety of available ML algorithms, such as random forest (RF) [41], Artificial Neural Networks (ANNs) [42–44], and support vector machine (SVM) [45], it is possible to underline two extensively used approaches that will be implemented in the workflow of this present research (Section 2.3).

ML classification approaches have been efficiently investigated for treating SAR data using SVM methods [46]. To classify satellite data, Ref. [46] presented the Sentinel-2 Time-Series, which exploited the Google Earth Engine computational platform. Moreover, [47] effectively used SVM for hills micro-landform classification based on grid DTM.

A specific category of raster processing and interpretation, related to the present research, is the use of geomorphological layer (GML) methods, composite raster data, and visualization techniques (VTs) [48]. It is identified for DSM analysis and classification tasks and employed as a training dataset for Machine Learning classification algorithms. Ref. [49] presented an application of ML and multiple layer combination for a multicollinearity analysis for the classification of potential areas for groundwater, with an accuracy of almost 80%. Furthermore, as preliminarily investigated in Ref. [21], ML and VTs have already been experimented with for supervised automated classification of micro-topography in Como Park.

Several specific disciplinary sectors have widely explored the potential connected to DL applications related to image classification. Recent studies have also investigated DL research for multi-class semantic classification, as in Ref. [50], where starting from 1 billion masks in 11 million images, a classification model characterized by a significant generalization capability was developed, enabling the possibility to segment a massive variety of images with heterogeneous features. Concerning environmental studies, Ref. [51] developed a solution to detect crop pest species, implementing semantic enrichment and comparing the results obtained using RGB and NIR UAS imagery. Also, heritage and archaeological domains have widely benefited from the advancement of technology in the automatic classification framework. For example, in Ref. [52], starting from a manually annotated dataset, a Convolutional Neural Network (CNN) was trained to enhance the photogrammetric pipeline for the 3D reconstruction of heritage objects. Moreover, Ref. [53] exploited an integration of CNN UNet-3 and Watershed algorithms while defining a methodological pipeline to generate a digital vectorized representation of decorative apparatus.

1.1.2. Integrated Classification Approaches of 3D Unstructured Data

The features and performance of 3D point-based classification methods differ from those of raster-based approaches for imagery classification, mainly due to the unstructured nature of the primary data [54,55]. In this regard, point cloud semantic segmentation and classification is a crucial task today, characterized by continuous technological and methodologic advancements [56]. Several experiments have been carried out to enhance the heritage documentation field, where the semantic contribution is crucial for evaluating automation approach performances both in ML and DL directions [57,58].

In this framework, DL networks and unsupervised algorithms can implement learning tasks directly in the 3D point cloud data. Among the available solutions, the geometric voxel analysis defines one of the most diffuse approaches [19]. A broad spectrum of unsupervised filtering algorithms related to geometric relationship features can be mentioned. The competition among them is mainly based on the accuracy of the results related to the context-related capabilities, the nature of the starting data, and the objectives of the application. In this context, a Simple Morphology Filter (SMRF) aims to segment the ground class points, operating a division of the point cloud into grids along an XY plane [59]. Moreover, in Ref. [60], a method based on the cloth simulation technique is used to extract the ground points in this framework. This algorithm has been implemented in Open-Source software platforms, such as CloudCompare v2.12.0.

A combination of geometric features can be used to recognize, extract, and label contour detector candidates from 3D point clouds based on geometric features (for example, for 3D modeling purposes [60]). Today, this issue is crucial in overcoming the need for Digital Twin Cities (DTCs) [61]. In this regard, several experiments on the ISPRS H3D benchmark dataset [62] have been conducted to enhance the segmentation of urban 3D unstructured data and ensure the semantic content's correctness in imbalanced class frequency datasets. For example, in Ref. [63], the integration of generalized-class point-transformer segmentation and a refined object-based approach fully address the imbalanced class distribution problem. In Ref. [64], a RF model was trained using geometric, radiometric, and echo support features derived from point clouds and mesh data.

Also, concerning supervised methods, several recent approaches explore DL techniques applied to 3D point cloud data, showing effective results [65–69]. Following the necessity of DTCs to constantly update 3D point cloud urban data, in Ref. [70], a semantic segmentation urban model was trained using the RandLA-Net architecture.

However, in this framework, a significant bottleneck is represented by the generation of reference data for the training and validation issues of a DL model. This task is demanding and time-consuming, requiring significant effort. Furthermore, the training process requires Python libraries that need powerful and updated graphic hardware (CUDA enabled). Memory consumption is thus a relevant issue while training a DL model [71]. In Ref. [72], a methodological example for training and validation data generation is addressed. Also, a reference dataset preparation strategy has been proposed in the literature for DL 3D point classification model evaluation, balancing the training and validation dataset in 80–20% [73].

This research also addresses the experimentation of some of the most innovative semantic segmentation solutions of point clouds in the heritage domain context based on supervised deep learning approaches.

2. Materials and Methods

The integrated methodology proposed in this present contribution (Figure 2) involves applying semi-automatic segmentation and classification approaches on aerial LiDAR datasets based on Machine Learning technologies through a combination of methods.

The aim is to adequately document the landscape morphology and to detect and classify ground-related features connected with anthropogenic landforms. The proposed integrated methodology thus aims to leverage heterogeneous 3D data to map these non-documented landscape morphological features, bridging the gap in the existing geographic datasets and updating and enriching the domain-related spatial database. In fact, within the complex framework of Cultural Heritage documentation, this work aimed at integrating comprehensive, current methodologies for the semantic classification of 3D point clouds [58] and raster data [38]. Therefore, the whole research is evaluated and validated, considering each approach from a holistic perspective.

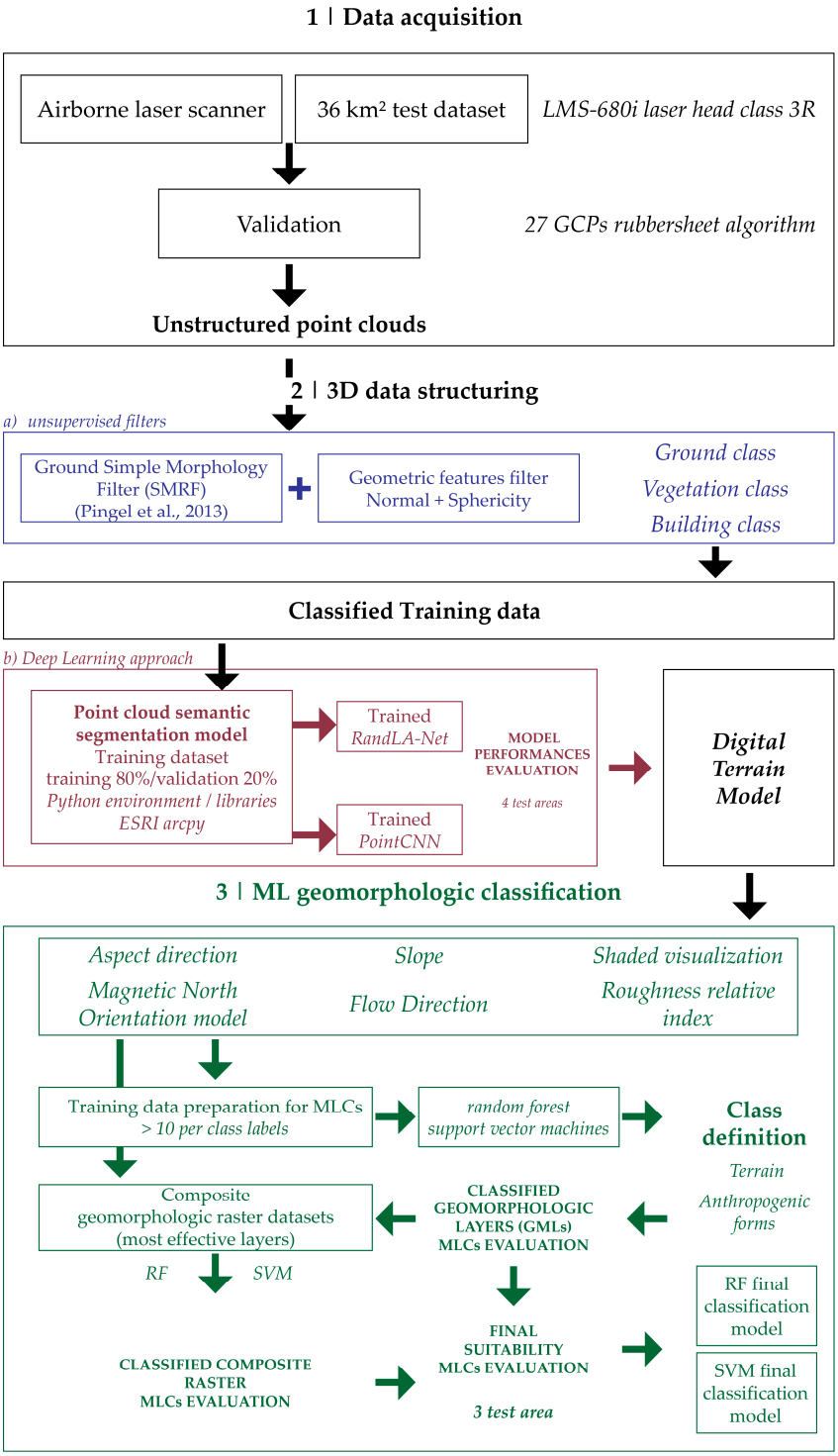


Figure 2. General framework schema. The integrated methodology provides a pipeline to document heritage ground features starting from a tailored high-scale airborne LiDAR survey (1). The second step provides a generalized class point cloud segmentation by integrating unsupervised and supervised DL approaches (2). Subsequently, the derived DTM and the geomorphological layer have been used for the application of ML classification methodology to map the impact of anthropogenic shapes on the ground.

After the first phase related to the data acquisition, the second part of the methodology consists of integrating unsupervised geometric filters with supervised Neural Network (NN) models for a generalized macro-class automatic segmentation of 3D point cloud data (Section 2.2). The macro-classes are intended to be compatible with the ASPRS specifications, even if, in this case, an exception has been addressed. The vegetation class encompasses various entities other than buildings and terrain. The third step provides an overview of the application of ML classification techniques using unconventional non-optic imagery data, such as DTM and geomorphological analysis (Section 2.3). The results of both approaches will be discussed in Section 4.

2.1. Primary Dataset (Airborne LiDAR)

The research area, located on the southern side of the Como Lake, features a vast dimension (almost 36 km^2), with a dense and irregular forest coverage of the Spina Verde Park, and evergreen plants in certain areas resulting in varying densities of point clouds. The area elevation is between 164 m and 597 m, and it is featured by steep gradient slopes with 0° min, 66° max, and 13° mean (Figure 3). Given these characteristics, the application of a low-altitude integrated LiDAR/photogrammetric flight survey was meticulously tailored to facilitate the efficient acquisition of comprehensive and uniform data and to achieve an optimal point density for the micro topography investigation. To mitigate the effects of dense vegetation on hilly terrain, the flight was performed before the spring season to minimize the vegetation growth. Moreover, as a terrestrial complement, an extensive survey campaign made by ground-based surveys has been planned to monitor individual archaeological sites. These ground-based data ensured a multi-scale approach and enriched the comparative analysis between aerial and terrestrial datasets, as already extensively discussed in Ref. [21]. An integrated UAS photogrammetric and terrestrial LiDAR campaign was carried out to document the park landscape and its archaeological sites. In the specific framework of this research, the campaign primarily delivered GNSS measurements that were used to validate the ALS dataset.

The flight was performed using a Eurocopter/Airbus AS350 equipped with multi-sensor hardware composed of the Litemapper 6800 LiDAR System with a RiEGL LMS-Q680i full waveform laser head operating with a maximum LPRR (Laser Pulse Repetition rate) of 400 kHz at a 60° scan angle, with nominal laser beam divergence < 0.5 mrad. The aerial flight positioning and attitude computation were performed using an IGI AEROControl GNSS/IMU composed of a 256 kHz IMU and a full GNSS Septentrion. Finally, a 150 MP PhaseOne iXM-RS150F medium format aerial camera (50 mm focal length) for stereo pair images was equipped to the platform.

Since the external aerial LiDAR POD is engineered to be removed each time, a dedicated calibration flight is mandatory to obtain a reliable dataset. For this reason, after takeoff, a special pattern composed of eight cross-flight strips at different AGLs over the helipad base was executed to calculate a boresight calibration. The helipad was fully resolved from a topographic/geodetic point of view to furnish Ground Control Points, Check Points, and Tie Points useful for estimating boresight misalignments (different for each flight) and solving camera parameters.

The flight was planned to provide a 70–50% overlap in front and lateral directions, considering an average AGL (Above Ground Level) of 1600 ft (approx. 490 m). The flight pattern comprised eleven flight strips, nine with NNE-SSO and two orthogonally to them, used for on-site boresight calibration purposes. Two cross-strips were added over the survey area to check and improve boresight (roll, pitch, heading, and mirror scale) parameters. The area was overflown to obtain high spatial resolution images with a GSD between 3 and 4.8 cm/pixel. Finally, a dense point cloud was generated with an approximate density of 40–50 pts/m^2 over border areas and more than 100 pts/m^2 over the top hilly area (with smaller AGL). The average density of data is thus estimated at 75 pts/m^2.

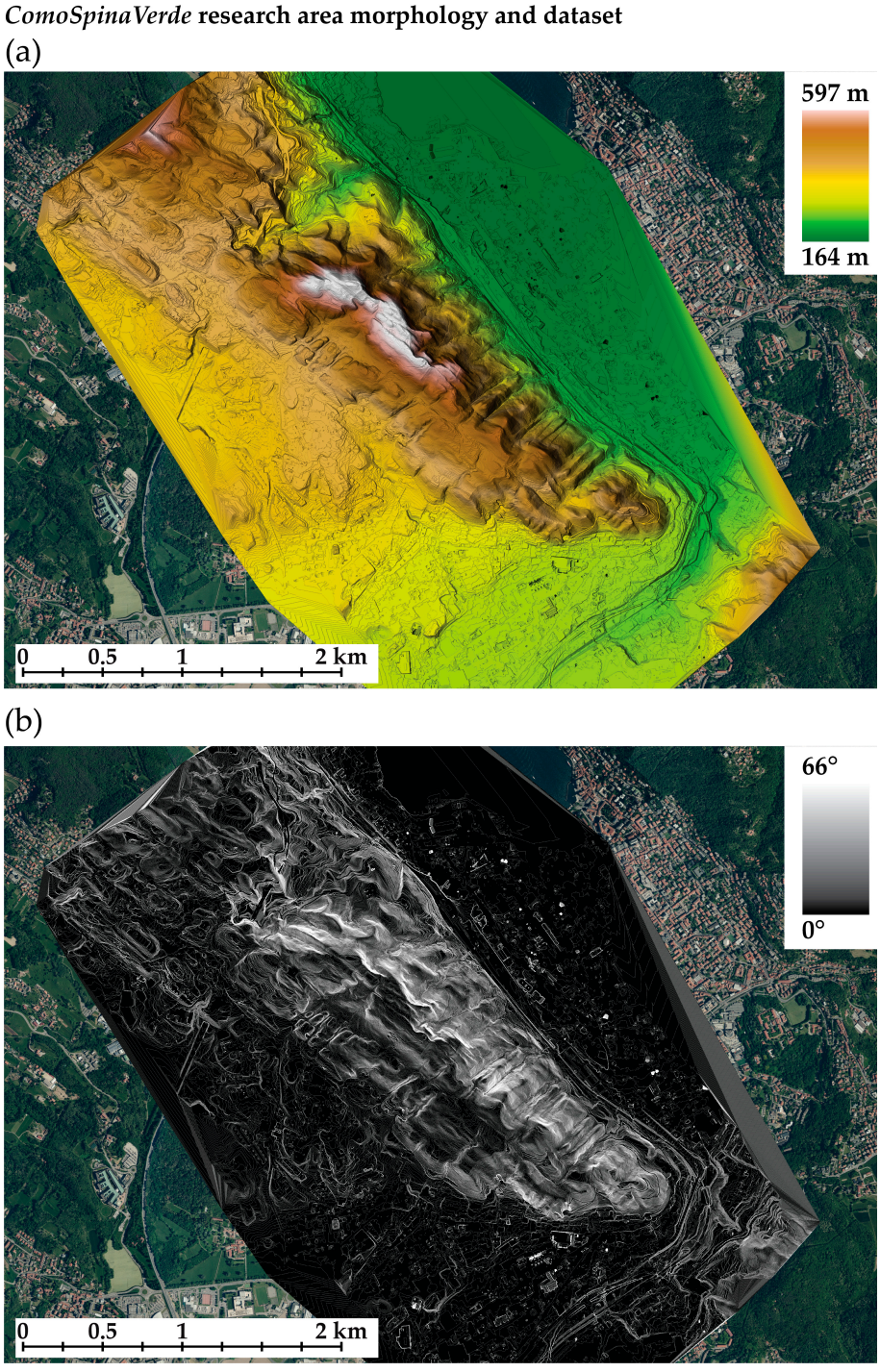


Figure 3. The Spina Verde Site's map displays the area of the airborne acquisition and the site's morphology: DTM (**a**) and slope direction analysis (**b**). The DTM was calculated from the filtered point cloud that had a density of 75 points per square meter (not filtered). Ground class had an average point spacing of 20 cm.

The trajectory on aerial LiDAR application in this case is solved in post-processing using Precise Point Positioning (PPP), i.e., when no GNSS ground station of a geodetic reference frame is present. The presence of a GNSS CORS (Continuously Operating Reference Station) over the survey area permitted to solve aerial N, E, and H separation (256 Khz IMU with two epochs per second) with an RMSE (Root Mean Square Error) of +/− 0.030 (ETRF2000—ellipsoidal). After that, to reduce the ellipsoid height to a

local orthometric datum, an official geoid model provided by the IGM (Italian Military Geographic Institute) based on the ITALGEO2005 geoid model was applied (declared accuracy is +/− 0.035 m at 1σ or +/− 0.100 m at 3σ).

Validation was finally verified using 27 Ground Control Points (GCPs) taken with GNSS real-time positioning techniques all around the survey areas, used to check the plano-altimetric consistency of the point cloud. This was first presented in Ref. [21]. In particular, a rubbersheet algorithm based on GCPs *Z*-axis residuals was used to smooth the altimetric fluctuation of the point cloud (Table 1). The GCPs were used as east–north constraints for the absolute orientation of the photogrammetric block. Consequently, a model keypoint ground surface was generated and used for image projection and the final orthomosaic generation.

Table 1. Rubbersheet algorithm and GCPs *Z*-axis residual weighted average were used as parameters for the algorithm (the control points are thus not absolutely fixed but have a variable variance to guarantee best fitting over the local plane).

Minimum Absolute ΔZ [m]	Maximum Absolute ΔZ [m]	RMSE [m]	Std. Deviation [m]
0.001	0.12	0.052	0.052

2.2. Supervised and Unsupervised Approaches Macro-Class Classification Approaches for DTM Generation

Two different approaches for the classification of the ALS point cloud have been followed to generate effective and accurate DTM for the analysis of the landscape topology. The intention was to establish a methodological low-time consumption pipeline to semantically classify an ALS LiDAR point cloud into a highly generalized three-class model: ground, building, and vegetation [24]. However, it is crucial to specify that the vegetation class was not intended to contain only the points referring to the vegetation but also all the other points remaining from the ground and building classes (e.g., powerlines). The first attempt used a Simple Morphology Filter (SMRF) filter for ground extraction and a tailored voxel-based analysis to filter the points pertaining to building and vegetation classes (see Section 3.1.1). The second involved DL approaches to train a transferable classification model for airborne LiDAR point clouds (Section 3.1.2). For the development of both methods, it was crucial to have access to adequately structured LiDAR data and process the dataset in such a way as to save computational power. Especially concerning DL approaches, as widely demonstrated by the state of the art [67], using original-resolution LiDAR point clouds with global coordinates and non-normalized scalar fields can impact processing times and consume excessive memory.

2.2.1. Unsupervised Approach: SMRF and Geometric Features Filter Integration

Regarding the geometrical filter application, the objective was to evaluate the effectiveness of combining different unsupervised filters: the consolidated SMRF filter [59] for the ground extraction step (Figure 4) and a tailored geometric filter to distinguish between buildings and vegetation.

The SMRF filter divides the point cloud into grids along an XY plane. The lowest elevation value is computed within each grid element to create a minimum elevation surface map. Through an iterative segmentation of the surface map and the elevation and slopes' calculation from an estimated DTM, it is possible to distinguish between ground and non-ground pixels referring to the optimal threshold for elevation difference and slope tolerance. The segmentation mask obtained from this process is then applied to the original point cloud, removing non-ground points. The integrated unsupervised filter uses the eigen-based geometric features of normal (λ_3) and curvatures (1 m radius neighborhood) to distinguish between points pertaining to vegetation or building classes. In this function, the variance in normal and curvature values was used to differentiate between buildings

and vegetation. Buildings typically exhibit greater planarity conditions than vegetation, resulting in less curvature variation and relative normal differences among neighboring points in a certain cluster. In contrast, vegetation points are scattered, leading to higher curvature variations than buildings. The result of the filtering process returns a classified point cloud. The filtering process could then be optimized for every case study using eigen features to disambiguate with more accurate results. The final structure of the point cloud is x, y, z, I, and C:

$$I = Intensity$$
$$C = Classification\ ground\ truth\ label$$

In this case, a reduced field structure has been conceived to evaluate a transferable model for ALS point clouds. One of the aims was to evaluate a strategy using the least number of parameters possible, avoiding storing radiometric and return pulse values.

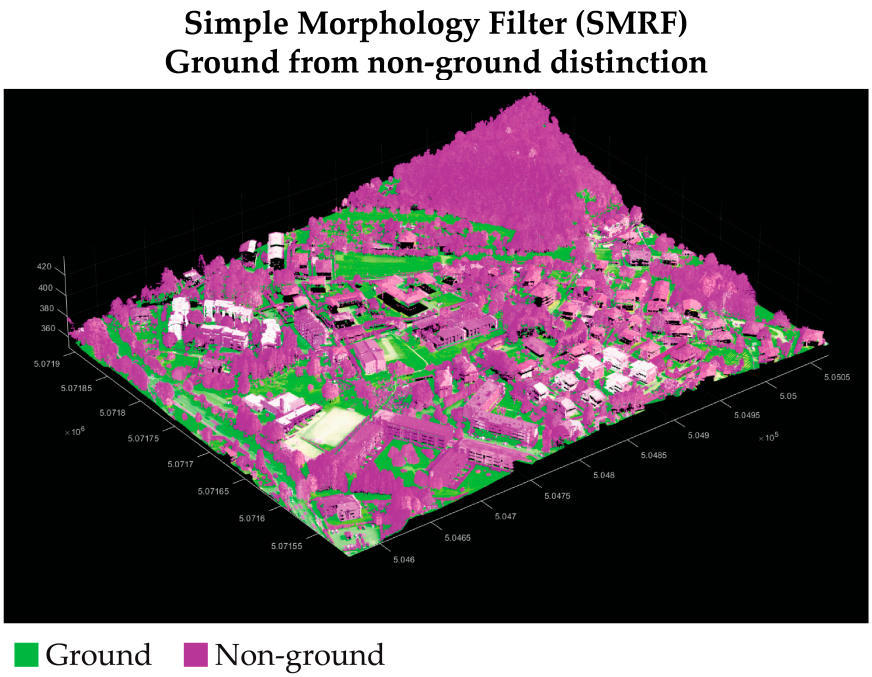


Figure 4. First step: ground from non-ground filtering with SMRF.

The purpose of the steps previously described was to validate an approach to the use of DL methods evaluating an alternative to the use of unsupervised filters, as well as to assess the use of the filters themselves as a replacement or aid for manual segmentation in data preparation to train ANN classification models (DLMs) of ALS point clouds. In this specific case study, the authors evaluated the method's reliability using the graphical user interface (GUI) solution in ArcGIS Pro 2.9 instead of a common code-based solution. Today, software houses like ESRI are deepening the research of integrating Machine Learning tools inside software to allow spatial data scientists to easily include GeoAI processing tasks in the same environment. Regarding the processing, two different workstations were used to train the models (Table 2), both with CUDA-enabled NVIDIA RTX graphic cards.

Table 2. Workstation comparison.

Workstation	GPU	VRAM (GB)	Accessible RAM (GB)
1	NVIDIA RTX A4000	16	32
2	NVIDIA RTX A5000	24	32

2.2.2. Supervised Approach: Training and Validation Datasets

The first step in the supervised classification workflow is to train deep learning models (DLMs) aiming at converting unstructured 3D datasets into a semantically organized dataset. This dataset contains significant semantic information about the class of each point, which is used as primary data input for the ANN.

Two separate neural architectures were used in the training process: RandLA-Net [67] and PointCNN [66]. A CE (Cross-Entropy) loss function was used for both architectures that pertain to point-based approaches. As previously mentioned, RandLA-Net and Point CNN significantly differ for two reasons. Firstly, PointCNN is a point convolution network using only coordinates as the input, while RandLA-Net is a point-wise multilayer perceptron architecture. Moreover, PointCNN is particularly suitable for analyzing small-scale object clouds. At the same time, RandLA-Net is designed to efficiently process urban or territorial-scale data, incorporating a random sampling phase for each network node. For this reason, the dataset was employed, adopting different downsampling strategies and normalization techniques for the point cloud values.

The PointCNN models were trained using sub-sampled data (2 pt/m^2) with a minimum distance sub-sample strategy, while the RandLA-NET models were trained using the original resolution data (75 pt/m^2). The main reason for adopting this sub-sample strategy was to reduce the inferential training time and memory consumption. Additionally, the intention was to evaluate the methodology's scalability using a strong sub-sample strategy. Moreover, the classification of "older" datasets with lower density >5 pts/m^2 represents a crucial issue in the environmental and heritage landcover mapping framework.

One of the aims was also to evaluate the impact of geographical coordinates in the training processing. Despite previous research experiences, this aspect has been considered extremely important to avoid memory exceeding failure [57]. The training data were thus generated using shifted classified point clouds resulting from the integrated unsupervised filter, where some checks and few corrections related to the building class were applied. In this case, the aim was to limit the manual operator intervention, working on a reduced dataset that could be used as a training dataset (Table 3).

Table 3. Data preparation for DL training.

Dataset	Training/Validation	Tiles Number (100 m × 100 m)	Total Area [km^2]	Training/Validation Percentage	T-V on Total Dataset Percentage
ComoSpinaVerde	Total		36.00		
T-V-a *	Train	20	0.20	80%	0.7%
	Validation	5	0.05	20%	
	Total	25	0.25		

* Testing with workstation 2 [Table 2].

Since the park area is characterized by dense vegetation, the T-V-a dataset was chosen to adequately represent the class distribution of the area. In fact, it is possible to observe from Figure 5 that the most populated class is the one pertaining to high vegetation. The class imbalance issue observed in this dataset and its generalized class scheme is a well-known problem. This issue also arises when working with denser datasets [62], where there is a need to provide a less generalized class schema [74].

Two final predictive DLMs were thus trained for T-V-a and will be discussed in Section 4.1:

- Model 1, trained original resolution data (architecture: RandLA-Net);
- Model 2, trained using sub-sampled data (architecture: PointCNN).

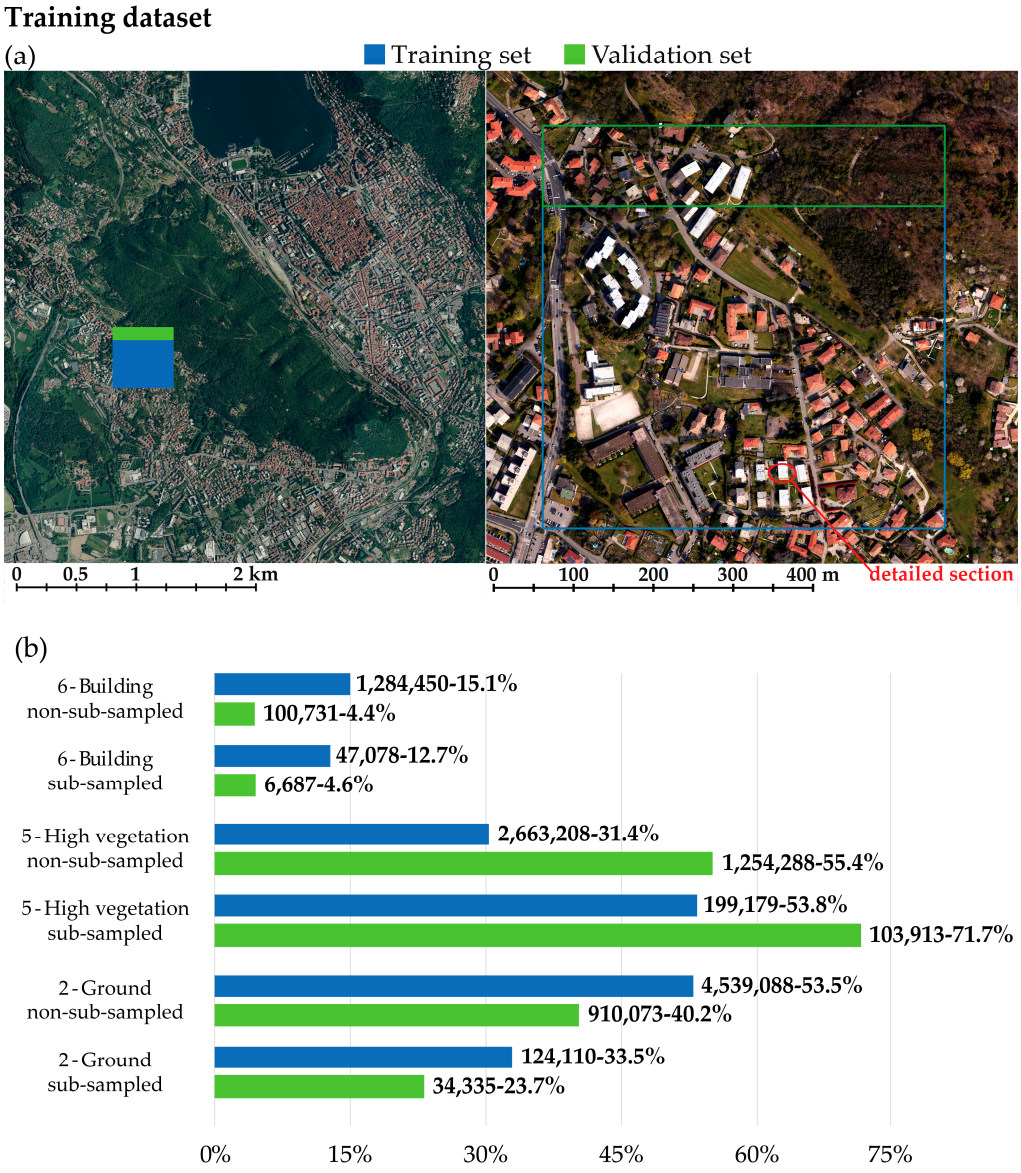


Figure 5. T-V-a location map (**a**) and dataset class frequency graph (**b**). As can be observed, the park area is characterized by dense vegetation. For this reason, the reference datasets were chosen to adequately represent the class distribution of the area. In this sense, the most populated class is the one pertaining to high vegetation.

2.3. MLCs for Grid-Based DTM Classification

The following part of this present research develops a methodological approach based on 2.5D raster data from a previous work. It is worth understanding that this previous research [21] was based on assessing the feasibility of using dense terrain models for feature extraction and, consequently, semantic segmentation aimed at automatically recognizing anthropogenic shapes (namely, trails). In this regard, as mentioned in Section 1.1, employing ALS data represents a powerful technology to enhance the study and the analyses of the sites' landscape morphology within densely forested areas [22]. Moreover, even if in recent years it was stated that an image-based approach is particularly powerful in the framework of classification and semantic segmentation tasks, working in a densely forested area does not allow—in this present research—for the use of a photogrammetric approach for the generation of a dense surface model, in favor of ALS data [22]. Thus, starting from the semantic segmentation of the ALS point clouds presented in Section 2.2, an accurate DTM with a 20 cm Ground Sampling Distance (GSD) was generated (Section 2.3.1), and geomorphological layers

(GLMs) were exploited as primary data for ML classification approaches to improve the level of automation and quality in detecting the human-made terrain features (Section 2.3.2).

2.3.1. DTM and Geomorphologic Analyses Preparation

Generally, various features related to the ground surface can be detected, such as trails, human earthworks, and similar elements present, even if not well represented, in the raw DTM. Considering the research aims, the interpretation of visual features related to the ground surface, such as anthropogenic elements, represents a relevant challenge. Therefore, the issues connected to difficulties in identifying and interpreting these visual features became evident when studying the DTM. In fact, it was not easy to unambiguously detect these features in ALS point cloud data and the subsequent raster interpolation of DTM. It is indeed clear from Figure 6 that both the orthoimage and raw DTM were not helpful in reaching the research aims. Concerning the RGB orthoimage, the vegetation covers almost the entire underlying terrain, even though aerial photogrammetric acquisitions have been performed during the leafless period of the year.

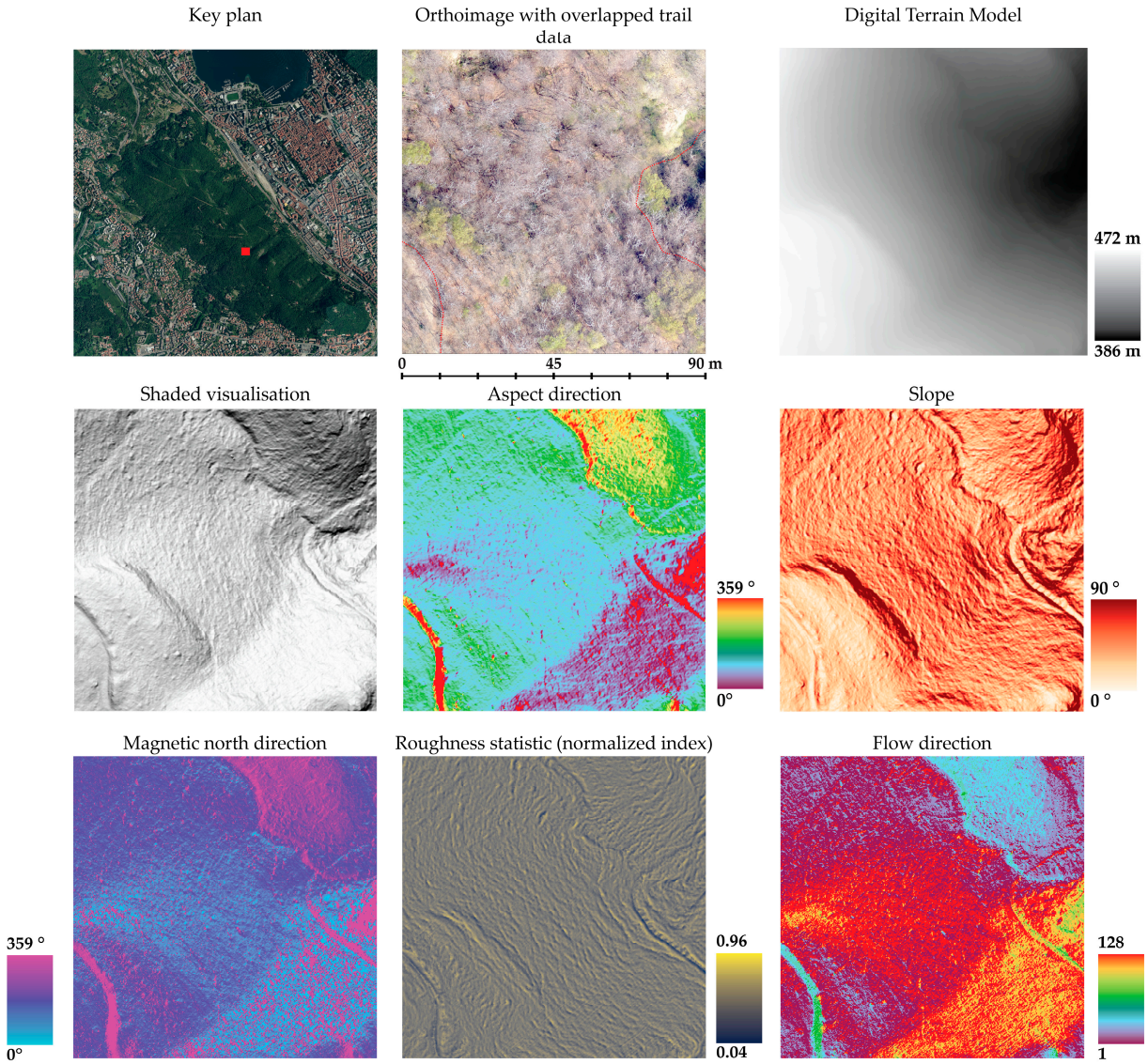


Figure 6. A comparison of the orthoimage, DTM, and subsequent geomorphological analyses was performed on a sample site, showing the park's trails. Each raster analysis was visually inspected, considering a sample area, in order to understand its suitability for the rapid identification of the anthropogenic shapes of the terrain. The red square in the key plan represents the extension of the analyzed area.

To address the issues related to ground feature identification, an investigation was carried out to generate geomorphological raster datasets that could enhance the interpretation of traces in the terrain morphology.

Different geomorphologic analyses have been conducted to determine the most significant geometrical and numerical signatures, with the aim of detecting the significant ground shapes (Figure 6). The shaded visualization, as well as the slope, aspect, and flow direction analyses, were generated from the DTM surface.

Moreover, from a Focal Statistic (FS) analysis on a five-pixel neighborhood cluster of the DTM, a Relative Roughness index r_i [75] raster was estimated using the raster calculator.

From a spatial analysis of the 3D point cloud, the direction value of the magnetic north was calculated for each point. This function calculates the orientation angle for each point using the supplementary attributes of magnetic north vectors acquired during the flight [76]. The resulting scalar field is then interpolated to create a shaded false-color visualization. From this visualization, human artifacts such as trails become more visible. This method is considered an effective alternative to simple DTM Hillshade, which often fails to highlight man-made paths, especially under vegetation.

From the visual inspection of the geomorphological analyses (Figure 6), it is evident how specific ground features such as trails, depressions, and rock formations are detectable. On the contrary, the elevation model and the orthoimage have not proven suitable for this type of task since the searched features are not unambiguously evident. Therefore, geomorphological layers (GMLs) were used as training datasets for applying Machine Learning (ML) classification algorithms.

After evaluating the classification accuracy (Section 3.2) using the individual geomorphological analyses, several composite raster datasets were generated by combining the GMLs (Figure 7).

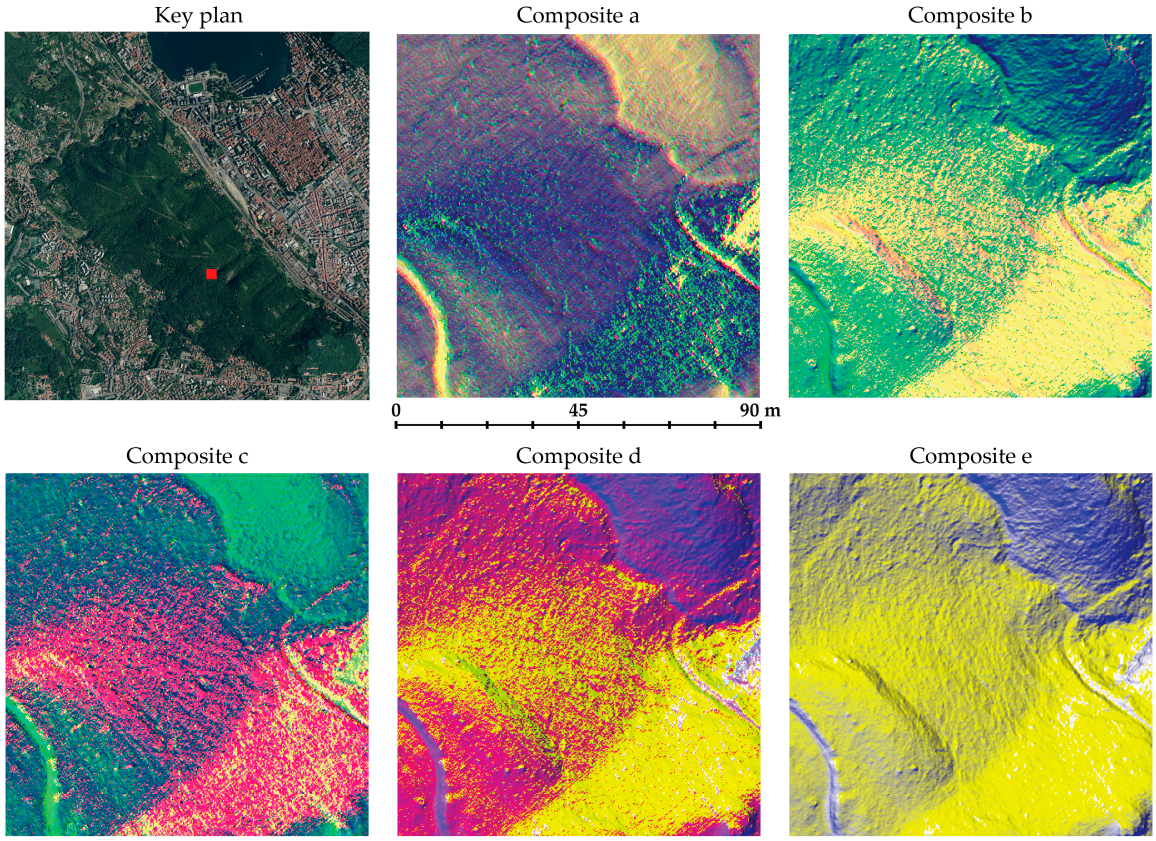


Figure 7. Comparison of the chosen composite geomorphological raster. In Table 4, it is possible to observe the RGB band disambiguation for each raster. The red square in the key plan represents the extension of the analyzed area.

In this case, the aim was to assess whether generating these composite rasters could enhance the classification effectiveness. For this reason, five composite raster data sets were generated. The combination of GMLs in each composite raster is reported in Table 4.

Table 4. RGB bands disambiguation in the composite geomorphological raster (Figure 7).

Composite Raster	Red Band	Green Band	Blue Band
Composite a	Aspect	Magnetic north direction	Roughness index
Composite b	Flow direction	Shaded visualization	Roughness index
Composite c	Flow direction	Magnetic north direction	Roughness index
Composite d	Shaded visualization	Flow direction	Aspect
Composite e	Shaded visualization	Shaded visualization	Aspect

2.3.2. Geomorphological Layer (GLM) Data Preparation Strategies for ML Classifiers

Since it has been established that Machine Learning classifiers (MLCs) are particularly suitable for geospatial imagery [38], the integration of Machine Learning techniques within the methodological workflow was tested. However, in contrast to what has been developed so far, some recent studies [17,39,40] have investigated ML classification techniques adopting an optical sensor data integration strategy with digital elevation model data. In this step of the pipeline, the goal was to achieve a high level of automation for the macro-scale land-cover mapping task, which was the ultimate objective of the proposed methodology, thus adopting the use of MLCs. These classification algorithms belong to the supervised learning category, where a predictive model is generated from an input set of training samples. Additionally, it is widely acknowledged that MLCs can handle large datasets effectively, producing more accurate results for complex data than parametric classifiers [38]. Parametric classifiers, such as maximum likelihood, require the assumption of normal data distribution and are slightly influenced by training data. Moreover, the landscape's morphology complexity led to noisy classification results [37]. Thus, this study tested two different nonparametric, per-pixel, and supervised Machine Learning classifiers to reach the desired aim.

- **SVM** is a Machine Learning method based on supervised statistical learning theory exploiting the theory of small samples [77]. For classification learning tasks, SVM evolved as a nonlinear probabilistic algorithm focusing on the distribution and distinction of training samples (support vectors) among two classes based on optimal hyperplane [45]. SVM thus efficiently suits remote sensing applications for high-dimension pattern recognition and binary classification tasks [77].
- **RF** combines multiple Decision Tree classifiers (DTs) [41]. Each DT uses a random sample of the training data to reduce the variance, finally assigning a unique class. The final prediction is then assigned by considering all the DT results through the majority voting method.

As demonstrated, the training samples have a crucial role in the effectiveness of the MLCs, even if there is not a generalized model for the training data preparation. However, it can be assumed that the numerosity and size of training samples must be planned accurately, depending on the classification algorithm, the number of class variables, the quality of the primary data, and the variety within the whole spatial extension [38]. The area and the support vectors that have been used during the training data preparation are displayed in Figure 8. For this task, a predefined dual-class classification scheme was employed to categorize anthropogenic landforms (namely trails) and the ground. The first objective was to minimize human involvement in the labeling process while maximizing label coverage across all regions of the training area.

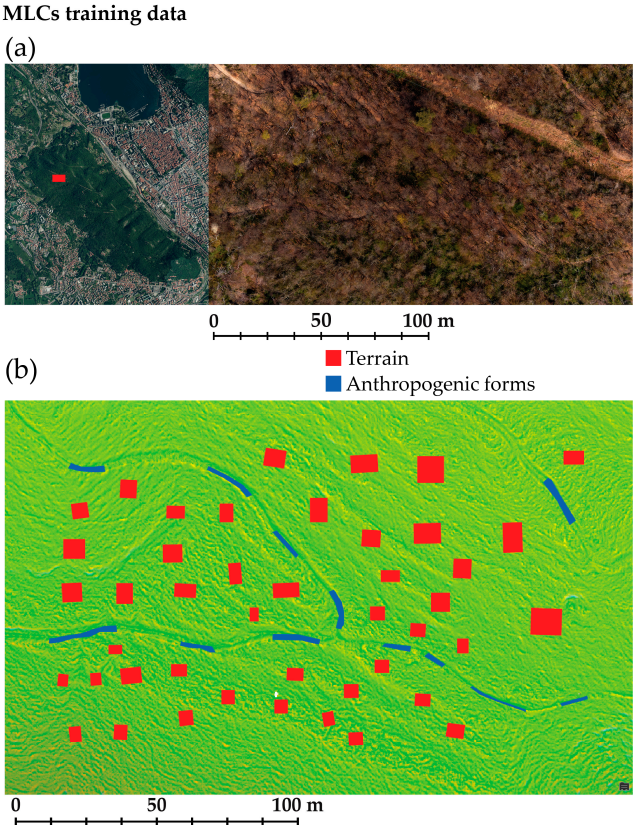


Figure 8. Map of the training area location (**a**) and training data label generation (**b**).

The main aim was to optimize the labeling process by using the minimum number of samples up to the recommended minimum [38]. In this sense, the number of labels is >10 per class, and the training samples' class frequency is coherent with the distribution of the features within the raster extension (Table 5). The results of this approach will be discussed in Section 3.2.

Table 5. Training data labeling strategy comparison.

Class	% m^2/Total	m^2 Average
Anthropogenic forms	26.27%	21.89
Terrain	73.73%	32.11

3. Results

Following the classification processes outlined in Section 2, the performance of the predictive models and algorithms was evaluated to assess their effectiveness. The classification accuracy of the DLMs, unsupervised filter (Section 2.2), and the MLCs (Section 2.3) was evaluated by comparing the predictions with the ground truth labels. Starting from the data reported in the confusion matrices (true positives, true negatives, false positives, false negatives), the following metrics [78] have been estimated to evaluate the achieved results: Accuracy (2), Precision (3), Recall (4), F1 score (5). The metrics have been calculated as follows:

$$\text{Accuracy} = \frac{True\ positives(TP) + True\ negatives(TN)}{True\ positives(TP) + True\ negatives(TN) + False\ negatives(FN) + False\ positives(FP)} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

3.1. Point Cloud Macro-Class Classification Evaluation: Unsupervised and Supervised Approaches

This section reports the results relative to the methods developed in Section 2.2. The performances of the DLMs (trained with the T-V-a dataset) and unsupervised filters were evaluated using the confusion matrix derived from the validation dataset. Finally, the DL models were tested on four different datasets, and the results are available in Appendix A.

3.1.1. Unsupervised Geometric Filter Results

Concerning the application of the integrated filtering approach based on the SMRF filter (ground) and geometric feature-based filter (off-ground) for deriving the three classes (Section 2.2.1), the following results can be analyzed.

Considering the unsupervised approach (Figure 9), positive outcomes have been experienced for terrain and vegetation classes. In contrast, the building class is the one characterized by the worst results in terms of metrics (Table 6). The high Accuracy metric achieved (≈99%) is due to the heavy numerical imbalance of points belonging to this class. For this reason, the very high number of true negatives (10,776,575 out of 10,990,089, ca. 98% of total points) significantly affects the Accuracy of this class. However, considering also the Precision, Recall, and F1 score, it can be observed that the building class is the one characterized by the lowest performance. This highlights the difficulties of the employed unsupervised algorithm in unambiguously recognizing the considered class. Overall, the results can be considered adequate for the accurate generation of a digital model of the terrain. Specifically, a sensitive evaluation of the results for individual classes demonstrates how the SMRF filter [59] manages to apply good generalization in ground points. However, cases of under-prediction are observed in some areas (Precision ≈ 76%).

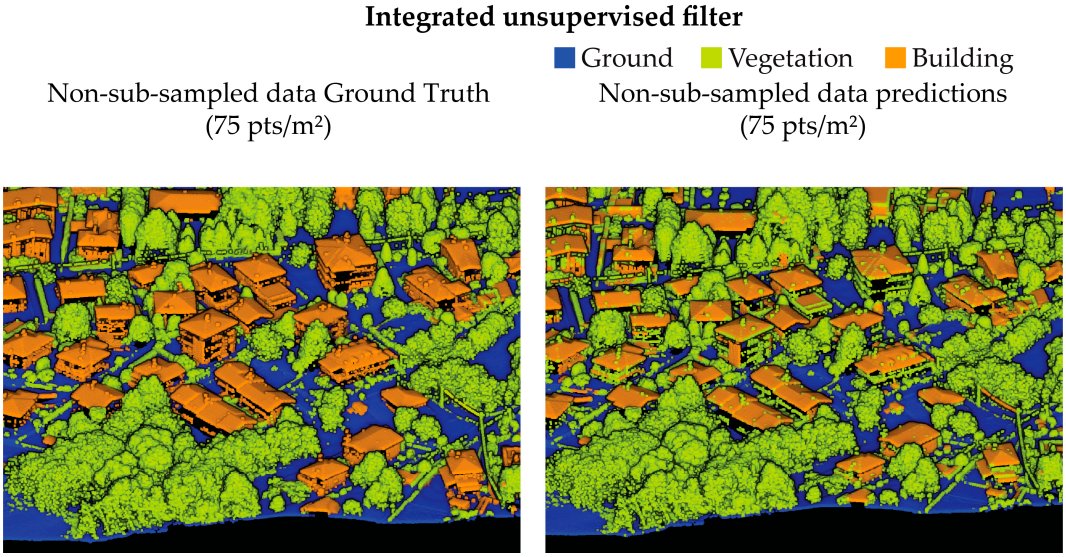


Figure 9. Unsupervised filter results: Ground Truth comparison with prediction results on the T-V-a validation dataset.

Integrating a geometric feature filter algorithm allows for disambiguation between building classes using only the eigenvalue of normal (λ_3) and geometric values of sphericity, which was then evaluated. In this case, the significant imbalance between the two classes does not lead to easily interpretable results, especially regarding points related to buildings.

However, it is evident from Table 6 that the vegetation class does not demonstrate an optimal sensitivity (Recall ≈ 86%) if compared to the one achieved for the ground class (Recall ≈ 99%). As shown in Figure 10, a high percentage of over-prediction can be observed, particularly localized on building roof areas with higher noise and characterized by architectural objects deviating from the average surface (e.g., chimneys). Similar behavior occurs in the case of the vertical elements related to facades, which are generally less dense and noisier than the horizontal ones. Due to this topological difference, the points pertaining to these surfaces are less assimilable as belonging to a plane and are thus associated with the vegetation class.

Table 6. Accuracy metric assessment of the unsupervised geometric filters carried out on the T-V-a validation dataset.

Class	Accuracy	Precision	Recall	F1 Score	Elapsed Time [hh:mm:ss]
Macro Average	0.93	0.79	0.86	0.82	00:14:33
2—Ground	0.90	0.76	0.99	0.86	
5—High Vegetation	0.90	0.99	0.86	0.92	
6—Building	0.99	0.62	0.73	0.67	

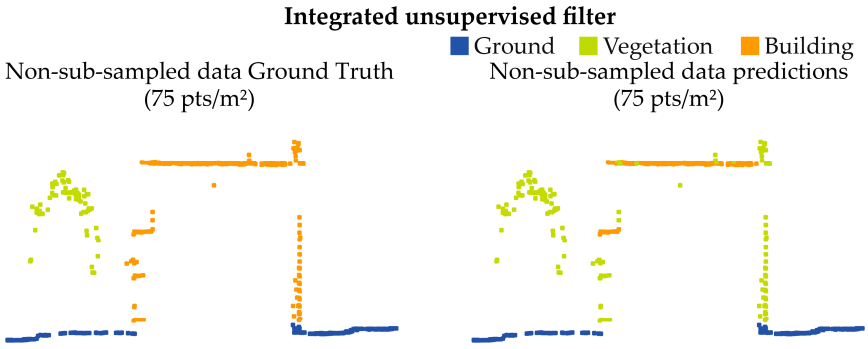


Figure 10. Example of filter results ambiguities. Comparison of Ground Truth with prediction results of the unsupervised filter adopting a common situation using a reduced cloud sample.

However, considering that the main objective was to segment ALS data in order to evaluate the effectiveness of the unsupervised integrated filter, this objective is not precluded or influenced by the results obtained. As observable by the evaluation metrics related to the ground class and reported in Table 6, this step of the proposed methodology represents a valuable solution for generating accurate DTMs.

The classification results of the unsupervised filter have been thus evaluated with a visual inspection of a human operator and, where necessary, refined in order to be used as the primary dataset to train a DL classification model.

3.1.2. Deep Learning Models Results

As stated in Section 2.2.2, two DLMs were trained from the finally selected T-V-a training dataset (Figure 5): Model 1 and Model 2. While Model 1 was trained using the original resolution dataset, Model 2 was derived instead from a sub-sampled dataset.

The DLM 1, trained with the RandLA-Net architecture, stopped on the 25th epoch after approximately 7 h, resulting in an overall accuracy of 92% (Table 7) observed in the classified validation dataset (Figure 11).

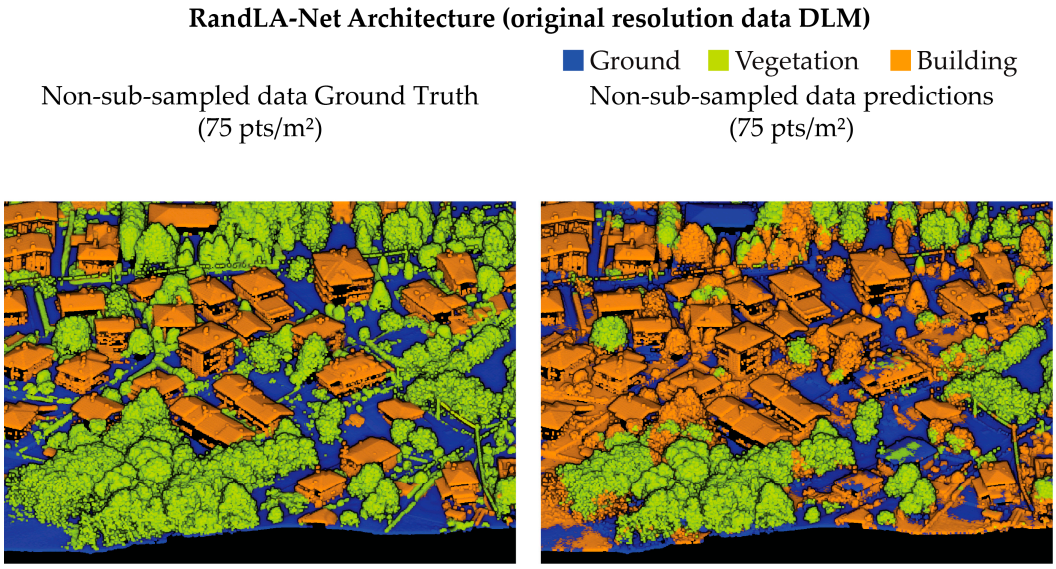


Figure 11. Ground Truth comparison with prediction results on validation dataset of the trained DLM 1 with the RandLA-Net architecture.

Table 7. Accuracy metric assessment of the trained Model 1 with the RandLA-Net architecture. The analyses are performed using the validation reference dataset.

Class	Accuracy	Precision	Recall	F1 Score	Elapsed Time [hh:mm:ss]
Macro average	0.92	0.75	0.86	0.79	06:59:36
2—Ground	0.91	0.77	0.91	0.83	
5—High Vegetation	0.91	0.98	0.88	0.93	
6—Building	0.95	0.50	0.79	0.61	

Considering the comparisons between Ground Truth, the predictions of each area, and the results of individual classes [Table 7], the sensitivity (Recall) of the ground class appears lower compared to the vegetation class. At the same time, the specificity (Precision) is higher, contrary to what happened with the use of unsupervised filters (Table 6). As for class 6 (buildings), the results did not meet the desired expectations. This is probably due to the limited number of points belonging to this class, and although part of the training dataset is related to an urban area, this led to an imbalance in the training data.

The DLM 2 (derived from the PointCNN architecture) finished the training in approximately 27 h at the end of the 25th epoch. Like DLM 1, DLM 2 was then evaluated on the validation dataset (Figure 12) and tested on four different areas (Figure A2). The performances were statistically analyzed regarding the overall and individual class evaluation metrics (Table 8).

Table 8. Accuracy metric assessment of the trained DLM 2 with the PointCNN architecture. The analyses were performed using the validation reference dataset.

Class	Accuracy	Precision	Recall	F1 Score	Elapsed Time [hh:mm:ss]
Macro average	0.92	0.87	0.80	0.83	26:39:36
2—Ground	0.92	0.94	0.95	0.95	
5—High Vegetation	0.96	0.53	0.62	0.57	
6—Building	0.92	0.87	0.80	0.83	

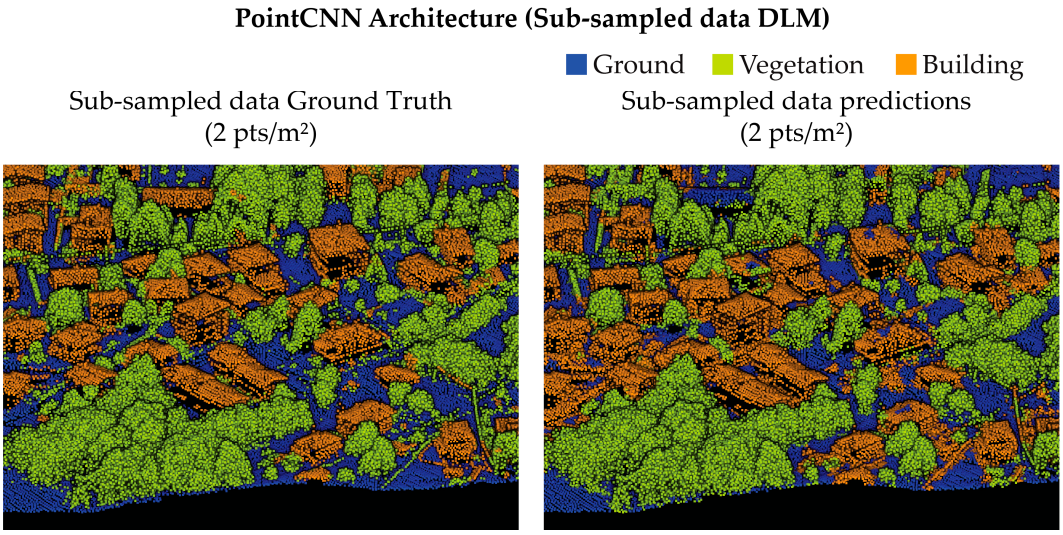


Figure 12. Ground Truth comparison from prediction results on the trained DLM 2 validation dataset with the PointCNN architecture.

The evaluation of Model 2 was carried out following the same methodology used for Model 1 (RandLA-Net classification model). From the evaluation metrics reported in Table 8, the ground class is more sensitive to over-predictions, unlike class 5 (vegetation), which is generally under-predicted. The metrics observed for the "building" class are consistent with the results obtained with Model 1, highlighting that this class is generally not predicted adequately in both cases. This is evident from the evaluation metrics—the lowest among the considered classes in the case of both models' predictions.

Considering the achieved outcomes of unsupervised and supervised approaches for the DEM generation, a further discussion related to the comparison of the filtering algorithm and the validation of the trained DL semantic segmentation models is proposed in Section 4.1.

3.2. Evaluation of MCLs for Grid-Based DTM Classification

This section reports the results relative to the methods developed in Section 2.3.

In order to assess the most effective classification technique for detecting the anthropogenic forms of the terrain, the MLC outcomes were analyzed for each prepared geomorphological layer. According to the error-test pipeline, this step was crucial to assess the suitability of the GMLs and to determine which one was the most effective as primary data for classification. Moreover, another significant aim was to evaluate the effectiveness of combining various geomorphological analyses as an alternative to the use of individual bands. An initial assessment of the evaluation metrics was carried out for the RF classifier, and the results are presented in Table 9, where it is possible to observe that Hillshade GML is the best-performing one. Yet, it is the only one to reach sufficient metrics (all the observed values are approximately $\geq$50%). However, as observable from the comparison between the evaluation metrics reported in Tables 9 and 10, the SVM classifier outperformed the RF approach. However, the RF classifier demonstrates a slight training time superiority over the SVM. The elapsed time for RF is usually below 10 min, while SVM models are trained in 13 to 15 min.

Table 9. RF classifier: accuracy metric assessment for each GLM, tested following the training labels provided in Section 2.

Analysis	Class	Accuracy	Precision	Recall	F1 Score
Aspect	Terrain	0.75	0.98	0.75	0.85
	Anthropogenic forms	0.75	0.12	0.75	0.21
Flow Direction	Terrain	0.95	0.97	0.97	0.97
	Anthropogenic forms	0.95	0.41	0.40	0.40
Hillshade	Terrain	0.97	0.98	0.99	0.98
	Anthropogenic forms	0.97	0.65	0.48	0.55
Magnetic north	Terrain	0.80	0.98	0.80	0.88
	Anthropogenic forms	0.80	0.14	0.71	0.23
Roughness	Terrain	0.56	0.97	0.56	0.71
	Anthropogenic forms	0.56	0.06	0.64	0.11
Slope	Terrain	0.58	0.98	0.57	0.72
	Anthropogenic forms	0.58	0.07	0.71	0.13

Table 10. SVM classifier: accuracy metric assessment for each GLM, tested following the training labels provided in Section 2.

Analysis	Class	Accuracy	Precision	Recall	F1 Score
Aspect	Terrain	0.97	0.98	0.99	0.98
	Anthropogenic forms	0.97	0.65	0.48	0.55
Flow Direction	Terrain	0.97	0.98	0.99	0.99
	Anthropogenic forms	0.97	0.74	0.56	0.64
Hillshade	Terrain	0.97	0.98	0.99	0.98
	Anthropogenic forms	0.97	0.70	0.49	0.58
Magnetic north	Terrain	0.95	0.97	0.98	0.98
	Anthropogenic forms	0.95	0.46	0.45	0.46
Roughness	Terrain	0.95	0.96	0.99	0.97
	Anthropogenic forms	0.95	0.30	0.14	0.19
Slope	Terrain	0.93	0.97	0.96	0.96
	Anthropogenic forms	0.93	0.23	0.28	0.25

Since the superior effectiveness of the SVM over the RF classifier has been confirmed, further analyses were addressed. In particular, it can be noted from the SVM results that several GMLs reached sufficient metrics and, therefore, can be considered suitable for this kind of classification (Aspect, Flow Direction, Hillshade).

Similar to the analyses carried out in the previous paragraphs, the evaluation metrics were also compared to evaluate the employed classifiers and, specifically, which composite raster performs better for the aimed classification task. Considering the evaluation metrics reported in Tables 11 and 12, it is evident that the SVM classification results outperformed again the RF results. While the Recall results are comparable, a significant improvement in terms of Accuracy (+16%), Precision (+17%), and F1 score (+19%) was observed.

Table 11. Accuracy metric assessment of RF classifier executed on each composite geomorphological raster, tested following the training labels prepared in Section 2.

Analysis	Class	Accuracy	Precision	Recall	F1 Score
Composite a	Terrain	0.87	0.98	0.87	0.93
	Anthropogenic forms	0.87	0.20	0.67	0.31
Composite b	Terrain	0.72	0.99	0.72	0.83
	Anthropogenic forms	0.72	0.12	0.83	0.21
Composite c	Terrain	0.85	0.99	0.85	0.91
	Anthropogenic forms	0.85	0.19	0.76	0.31
Composite d	Terrain	0.68	0.99	0.67	0.80
	Anthropogenic forms	0.68	0.11	0.85	0.19
Composite e	Terrain	0.83	0.99	0.84	0.91
	Anthropogenic forms	0.83	0.18	0.78	0.29

Table 12. Accuracy metric assessment of SVM classifier executed on each composite geomorphological raster, tested following the training labels prepared in Section 2.

Analysis	Class	Accuracy	Precision	Recall	F1 Score
Composite a	Terrain	0.95	0.98	0.96	0.97
	Anthropogenic forms	0.95	0.44	0.61	0.51
Composite b	Terrain	0.95	0.98	0.96	0.97
	Anthropogenic forms	0.95	0.45	0.65	0.53
Composite c	Terrain	0.95	0.98	0.97	0.97
	Anthropogenic forms	0.95	0.46	0.61	0.52
Composite d	Terrain	0.96	0.98	0.98	0.98
	Anthropogenic forms	0.96	0.58	0.60	0.59
Composite e	Terrain	0.96	0.98	0.98	0.98
	Anthropogenic forms	0.96	0.58	0.62	0.60

From the evaluation metrics reported in Table 11—and from a visual inspection, as observable in Figure 13—it is evident that the RF classifier applied on the composite raster datasets has not been effective for detecting the pixels belonging to the Trail class.

In fact, the classification of composite images achieved results characterized by approximately the same order of magnitude. None of these results outperformed the others, but overall, the metrics are insufficient. In particular, the Trails class is consistently over-predicted, resulting in a high level of noise after the classification task.

Regarding the SVM approach, a slight under-prediction can be observed, as indicated by the relatively low Precision value, which is always the lowest metric across all the other values. Generally, the data that led to the best results adopting the SVM approach are composite e (Hillshade, Hillshade, Aspect). This result is consistent with what was observed while analyzing the classification task on the individual GMLs, where the bands that enabled the achievement of better generalization were Hillshade and Aspect.

Contrarily, this does not apply regarding the RF approach, as the classification carried out on the Hillshade GML produced only adequate results, effectively outperforming the strategy based on the use of composite images.

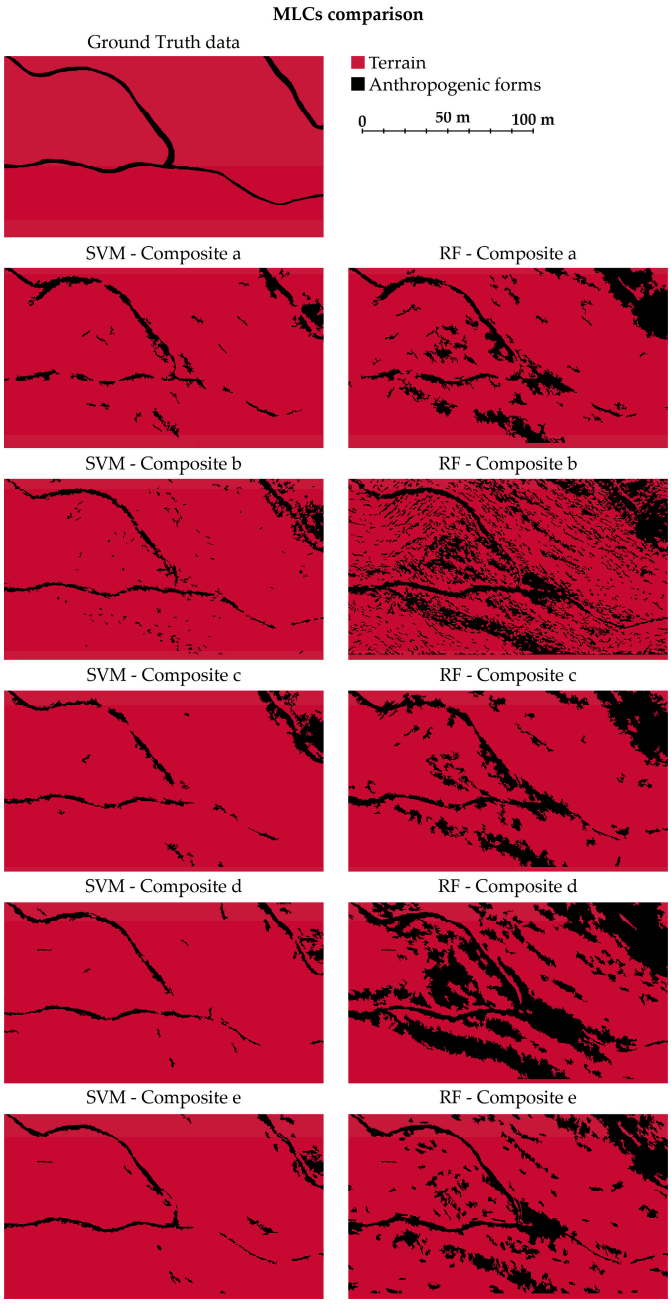


Figure 13. Qualitative comparison of the MLC results of the training area, using composite geomorphological raster, evidencing the higher effectiveness of the SVM approach due to evident cases of noise or overprediction.

Considering the achieved outcomes of the MLCs on the training area, a further discussion related to the validation of the trained ML predictive models is proposed in Section 4.2.

4. Discussion

4.1. Supervised and Unsupervised Approaches for Point Cloud Semantic Segmentation: Validation and Discussion

Considering what has been analyzed so far and the achieved results, as illustrated in the Section 3.1, two approaches have been carried out.

- Supervised Deep Learning models: DLM 1 (RandLA-Net) and DLM 2 (PointCNN)
- Unsupervised geometric filtering: integrated filter based on SMRF and geometric features calculation.

The trained DLM 1 and DLM 2 were specifically tested on four areas (100 m × 100 m test dataset) (Figure 14) using both original resolution and sub-sampled data in order to evaluate the classification in different morphological and urban contexts. In fact, these specific test areas have been chosen to analyze the model's behavior in different regions of the area, which accurately represent two main characteristics of the Como area: the dense urban fabric areas that embed the park and the forested, less anthropized areas of the park itself.

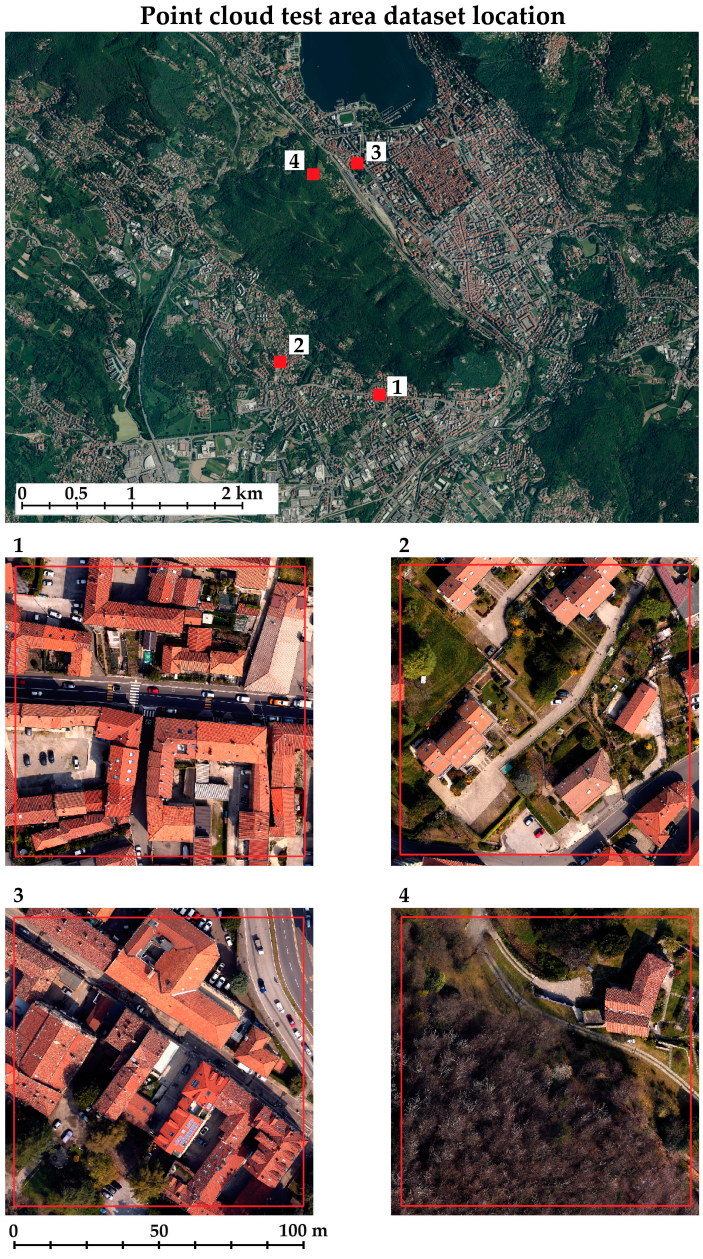


Figure 14. Map of the test areas' locations. The presence of heterogeneous features and morphology characterizes the considered areas.

In this case, the aim was to evaluate both the generalization capability of the models and the performance in classifying data characterized by different resolutions and densities compared to the training data. It becomes evident that in dense urban areas—characterized by a balanced class distribution—the classification results are more uniform. In contrast, the classification model has not performed a valuable generalization in areas with a more non-

balanced class distribution. Of course, better performance is correlated to the majority class. These considerations were summarized while visually inspecting DLM 1 result (Figure A1) and using the correlated detailed metrics analysis reported in Table A1. Additionally, regarding DLM 2 a lower generalization capability was observed from a visual inspection (Figure A2) of the classified point clouds for the specified test areas and the metric analysis (Table A2).

Finally, an overall comparison of the results derived from the adopted approaches (supervised and unsupervised) (Table 13) shows that the unsupervised filter method slightly outperformed the two DLMs in terms of Accuracy and F1 score. This corresponds to the preliminary expectations, as the state of the art has shown that the effectiveness of geometric-based methods is currently more efficient for this type of task than ML-based methods [19]. Moreover, the computational cost of the unsupervised filtering is more effective in terms of elapsed time, as well as human and machine resources.

Table 13. Point cloud classification models comparison on the validation dataset.

Unsupervised Geometric Filter (Original Resolution Data)		**RandLA-Net (Original Resolution Data)**		**PointCNN (Sub-Sample Data)**	
Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score
0.93	0.82	0.92	0.79	0.93	0.78

Furthermore, all the classification approaches have been tested using both original resolution and sub-sampled datasets. The aim was to assess how differences in terms of spatial resolution could affect the final classification results and, additionally, evaluate the effectiveness of a DLM on a less dense point cloud. In fact, open LiDAR datasets from public administration's landscape-scale survey campaigns are generally less dense than the current point cloud for Como Park. For this reason, it is thus necessary to conduct further evaluations on sparser datasets to validate this method's suitability and adaptability on a different case study where it is not possible to acquire ad hoc datasets (as for the current dataset). The authors underline that the significant results achieved using both original resolution and sub-sampled data are characterized by the same order of magnitude, confirming how this methodology is suitable for classifying point clouds with different densities (see Appendix A).

Although the state of the art typically employs more eigenvalues for defining class thresholds, the authors decided to use only the λ_3 eigenvalue (normal) and the sphericity geometric value. This approach was followed to evaluate the effectiveness of a classification based on the proposed method while simultaneously reducing computational effort, considering the significant extent of the study area. From the evaluation metrics, it is possible to observe the effectiveness of this type of approach and witness the correct behavior of the predictive algorithm. However, it should be underlined that the ML-based approach also led to adequate results. In particular, the model derived from RandLA-Net outperformed the one trained from PointCNN architecture, as seen in Section 3.1.2 and Table 13.

Based on the critical analysis of the point cloud approach, the following section presents final considerations regarding the validation of the image-based strategy.

4.2. Machine Learning Classification Models Validation

Considering what has been analyzed so far and the achieved results, as illustrated in the Section 3.2, two final models have been trained.

- Model RF: Random Forest algorithm. This model was applied to Hillshade GML.
- Model SVM: Support Vector Machines algorithm. This model was applied to composite e raster (Hillshade, Hillshade, Aspect).

This section aimed to evaluate the effectiveness of the MLCs' trained models in three test areas (Area A, Area B, and Area C (Figure 15)) that differ from the area where the training labels were extracted.

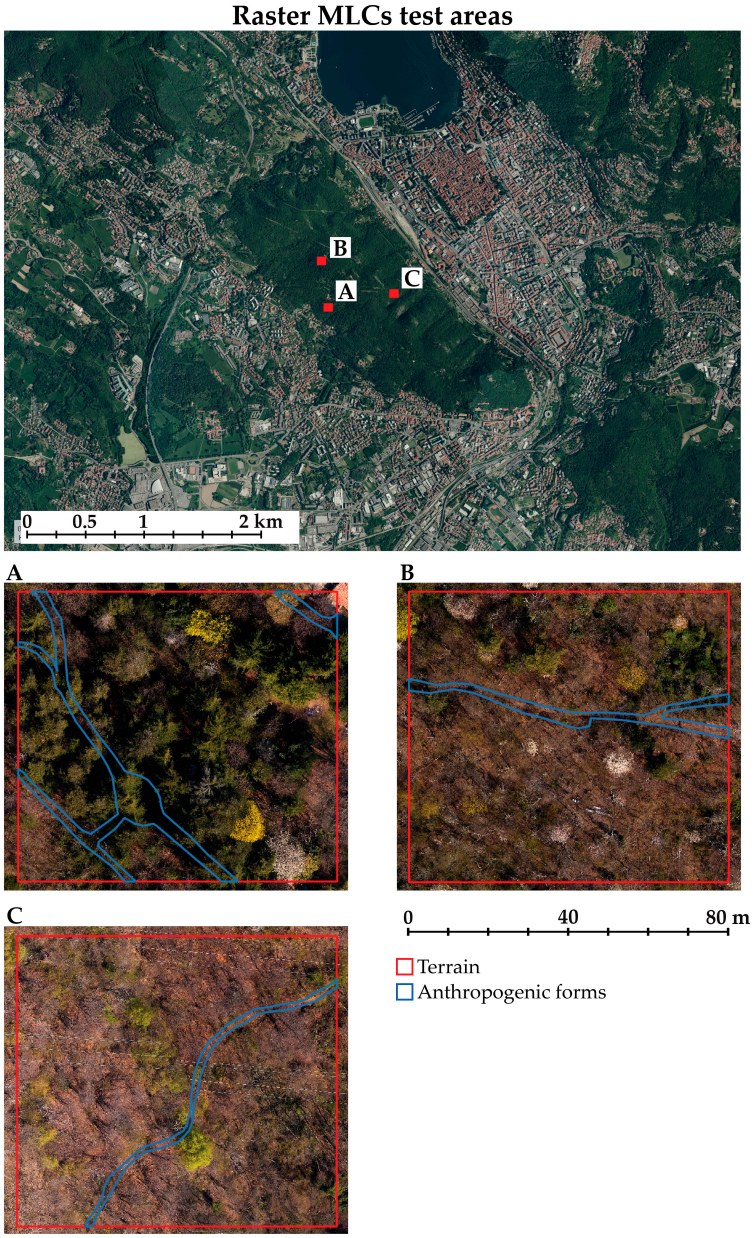


Figure 15. Map of the test areas (A, B, C) location. The presence of heterogeneous features and morphology characterizes the considered areas.

Moreover, in each area, manual labels were generated by a human operator to provide Ground Truth data for assessing the classification and calculating the evaluation metrics. Subsequently, the input data were classified using both SVM and RF classification models.

From Figure 16 and the analysis of the evaluation metrics (Table 14), it is evident that the model exhibits a good generalization capability, as the results are coherent with those achieved in the area where the training data were extracted.

Table 14. Accuracy metric assessment of the RF classification model executed on three test areas. The low Precision indicates a tendency of the model to mispredict the Trails class and has a generally noisy outcome. However, the main anthropogenic elements (Trails) have been adequately predicted.

Analysis	Area	Class	Accuracy	Precision	Recall	F1 Score
Hillshade	A	Terrain	0.82	0.96	0.84	0.90
		Anthropogenic forms	0.82	0.26	0.59	0.36
	B	Terrain	0.91	0.99	0.92	0.95
		Anthropogenic forms	0.91	0.27	0.76	0.40
	C	Terrain	0.93	0.99	0.93	0.96
		Anthropogenic forms	0.93	0.21	0.74	0.33

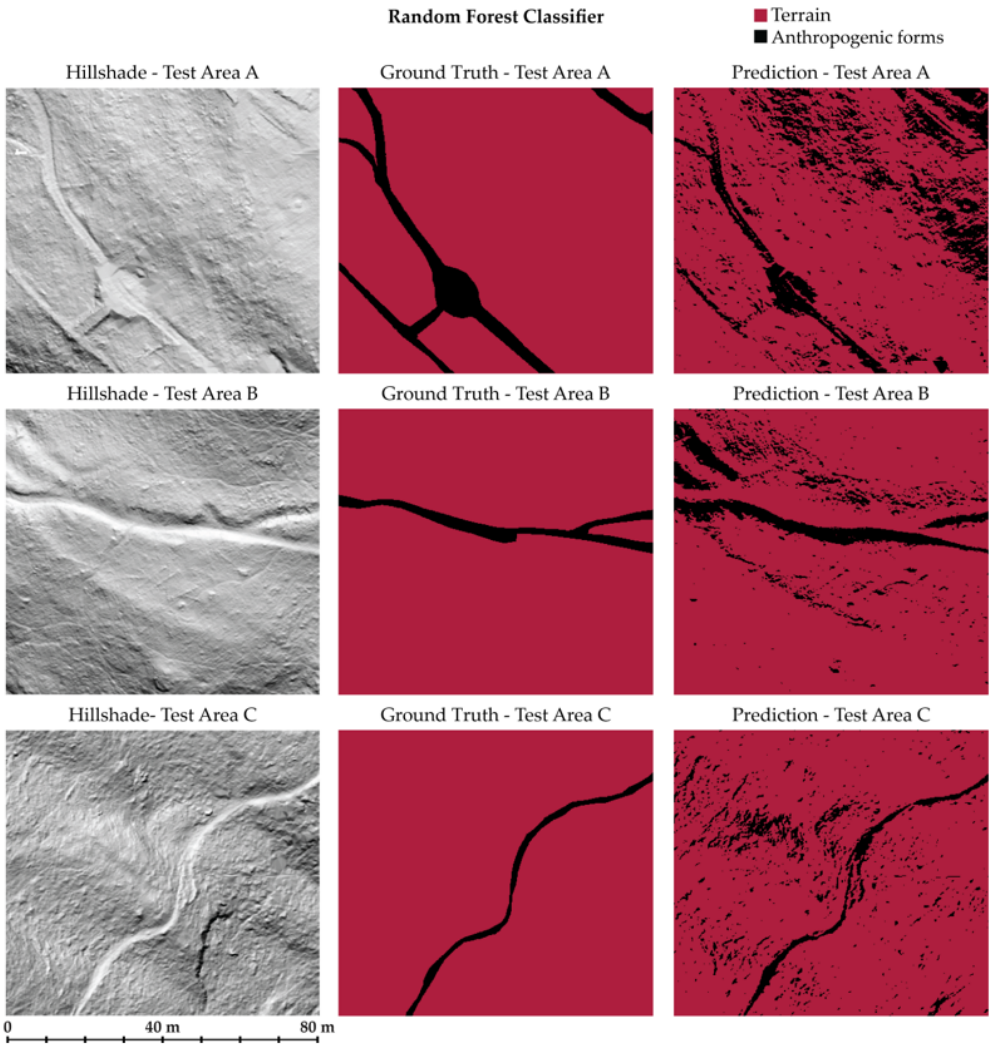


Figure 16. Prediction results of trained random forest model for composite e raster on three test areas.

Concerning the model's performance derived from the SVM model, even in this case, the results are comparable to the predictions achieved in the training area (Table 15). Figure 17 reports the classified images related to the three considered areas and, despite a slight under-prediction (in Area A) and relatively high nose level—especially in Area C, where the interpretation of the anthropogenic forms is more challenging due to the high

steepness and roughness of the terrain—the Trails have been detected with an adequate level of accuracy.

Table 15. Accuracy metric assessment of the SVM classification model executed on three test areas.

Analysis	Area	Class	Accuracy	Precision	Recall	F1 Score
Composite e	A	Terrain	0.91	0.93	0.97	0.95
		Anthropogenic forms	0.91	0.40	0.19	0.32
	B	Terrain	0.96	0.99	0.97	0.98
		Anthropogenic forms	0.96	0.47	0.68	0.55
	C	Terrain	0.90	0.99	0.90	0.95
		Anthropogenic forms	0.90	0.15	0.68	0.24

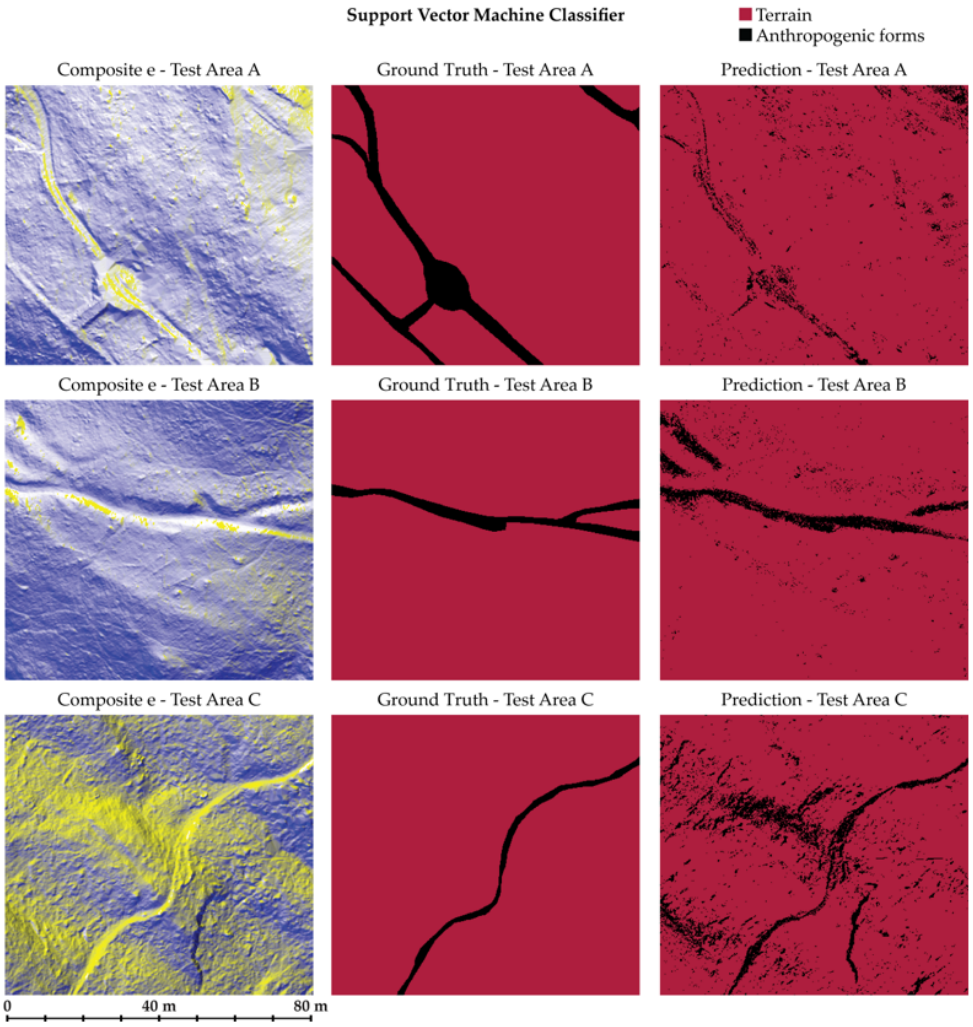


Figure 17. Prediction results of trained support vector machine model for Composite e raster (Red-shaded visualization; Green-shaded visualization; Blue-Aspect) on the three test areas.

Throughout the pipeline observed previously, particularly regarding the two steps preceding the training of models SVM/RF, the evaluation metrics underlined a general superiority of the SVM approach over the RF method. While evaluating the performance of the Trail class in each step of the pipeline, it can be noted from Table 16 that all the observed F1 score values are generally superior in the case of the SVM classification.

Table 16. MLC model comparison. Despite the SVM method outperforming RF in the first two steps, the results of the final predictive models are substantially comparable.

	F1 Score Average		F1 Score Max		F1 Score Min	
	SVM	**RF**	**SVM**	**RF**	**SVM**	**RF**
Single geomorphologic analyses *	**0.45**	0.27	**0.64**	0.55	**0.19**	0.11
Geomorphologic composite *	**0.54**	0.26	**0.59**	0.31	**0.51**	0.19
Classification models **	**0.37**	0.36	**0.55**	0.40	0.24	**0.33**

* F1 score comparisons are calculated only from MLC results from the Trail class. ** The classification models are trained using the best resultant raster for each classifier. SVM: composite e; RF: shaded visualization DTM.

5. Conclusions

Finally, consideration can be given to the lessons learned and the opportunities offered by integrating the various proposed methodologies and also some identifiable bottlenecks.

This study underscores the significance of integrating diverse approaches while critically assessing their performance and applicability. In this context, valuable insights can be addressed, contributing to advancements in remote sensing data analysis, environmental monitoring, land use planning, and archaeological research. Moreover, the flexibility of the proposed integrated methods can also be underlined. While the proposed research focuses on detecting historical trails through the generated predictive models, it should be noted that this strategy could yield comparable results for identifying different types of artifacts (e.g., terraces or other archaeological evidence). The versatility of the explored technologies could provide a powerful investigation tool for researchers working in the archaeological and landscape heritage framework.

In the proposed pipeline, the semantic enrichment of 3D unstructured data represents a valuable opportunity to enhance automation in the processes related to landscape and terrain morphology knowledge. Here, the results obtained from the DL approach highlight how ANN represents a valid and powerful alternative to the more consolidated geometric unsupervised filter. While the geometric-based filters are effective compared to the DL approach, they are strictly dependent on the topography and the vegetation of each considered site, as well as the urban and architectural morphology. Supervised methods, such as ANN-based methods, can achieve higher levels of abstraction and, consequently, higher generalization capability, highlighting the significant perspectives associated with using this technology.

Moreover, the supervised methods allow for less generalized semantic classes to be assigned, leading to more effective segmentation. However, it should be underlined that ANNs require significant structured data for the training. Thus, generating an adequately structured training dataset can be a time-consuming and demanding bottleneck. Recently, this issue has been addressed, and several solutions have been proposed to overcome the criticality represented by the generation of the training dataset; for example, generating an artificial training dataset [79,80], using open-source benchmark datasets [62,72], or exploiting pre-trained ANN for fine-tuning purposes [81,82]. Generally, this research direction represents one of the main perspectives for future application development.

Instead of considering the raster-based classification approach and evaluating the performances of the two MLC algorithms, as expected [38], the proposed methodology addressed the overall superiority of the SVM classifier compared to the RF classifier. However, it is worth mentioning that both the employed MLCs represent an adequate approach for developing automation solutions for the classification aims.

Moreover, it is worth emphasizing that the recognition of anthropogenic forms characterized by a high level of granularity represents a challenging task not only for supervised or unsupervised methods but also for traditional manual classification operations performed by a domain expert image analyst. In the framework of this type of landscape and urban-scale analysis, when a classification and detection task is required, a significant imbalance between classes is not uncommon due to the granularity of the features, as in the case of this present research. In this regard, it should be underlined that the effectiveness of unsupervised geometric

filtering methods is not affected by class imbalance, contrary to what happens in supervised methods such as ML and DL. Balanced classes are preferable for supervised approaches to obtain optimal results [44]. However, an imbalanced distribution of class population can lead to difficulties in the interpretation of metrics, as for classes composed of a limited number of samples, a few incorrect predictions can heavily influence the values of evaluation metrics. In fact, despite relatively low metric values ($\leq 60\%$), upon visually inspecting the classified images, it is worth underlining that the classification models adequately detected and classified the anthropogenic forms in the raster datasets.

Another aspect that deserves attention is the high costs of this type of sensor, i.e., the use of a full waveform ALS solution applied to the landscape heritage documentation framework. It enables the penetration of the sparse vegetation and thus its suitability in areas similar to the Spina Verde. Nevertheless, this limitation can be overcome by the recent development of low-cost UAS-based LiDAR commercial solutions (e.g., DJI Zenmuse L1). These systems have been increasingly used in recent years and can guarantee greater accessibility, ensuring sustainable data acquisition for geomorphological analysis purposes. However, the penetration capabilities of these more compact solutions are still being evaluated and validated in the literature.

Moreover, a reflection is crucial in training data preparation, to generate the minimum number of support vectors [38]. The main aim is to minimize operator manual activities while enhancing automation in classification processes. In this framework, the proposed methodology could represent a significant contribution.

Furthermore, regarding the future perspectives of this research (Figure 18), one of the most interesting directions is the possibility of applying the tested methodology to different case studies. These differ in terms of extension, density, and morphology. As evidenced in Section 4.1, the pipeline is suitable also for sparse low-scale LiDAR datasets.

Finally, considering the effectiveness of the trained classification models in detecting point classes related to anthropogenic forms and the necessity in the framework of heritage studies of enhancing automation processes, an interesting perspective is represented by the possibility of exploiting specific databases related to heritage and archaeological domains to train and validate predictive models for the automatic object detection—and semantic enrichment—of cultural and anthropogenic assets.

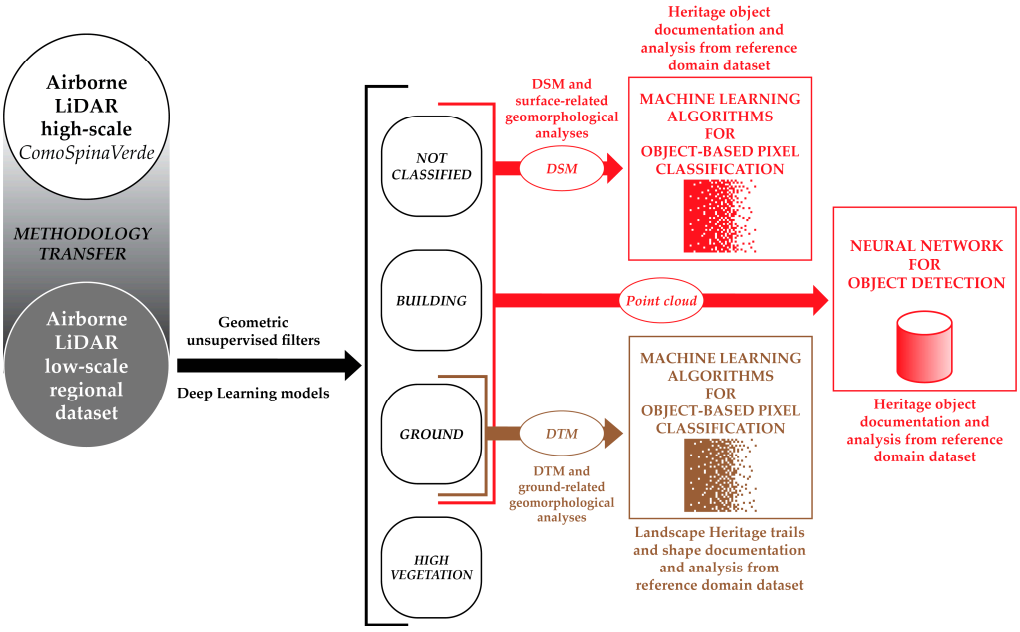


Figure 18. Future perspective schema aimed at transferring the proposed methodological pipeline for high-scale airborne data (Figure 2) to other available low-scale regional LiDAR datasets focusing on other landscape heritage domains.

Author Contributions: Conceptualization, M.C.; methodology, M.C., G.P. and A.S.; software, M.C. and M.B.; validation, M.C., G.P. and M.B.; formal analysis, M.C. and M.B.; investigation, M.C. and A.S.; resources, A.S. and M.B.; data curation, M.C., G.P. and M.B.; writing—original draft preparation, M.C.; writing—review and editing, M.C., G.P., G.S, M.B. and A.S.; visualization, M.C., G.P. and G.S.; supervision, A.S.; project administration, A.S.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: Parco Spina Verde, Como (IT) has funded Politecnico di Torino and CNR IRPI for the primary airborne and terrestrial data acquisition.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: Thanks to the Parco Spina Verde administration for support and funding.

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

Appendix A

The authors have decided to extensively report the metrics analysis of the DLMs related to the four test dataset areas in this supplementary section. The analysis demonstrates that the methodology, besides the improvement capability, is suitable for application with sparser ALS datasets (e.g., regional and national administration datasets).

Table A1. Accuracy metric assessment of the trained Model 1 with the RandLA-Net architecture, applied to four test areas. The metrics have been analyzed by comparing the adopted strategy for both original resolution and sub-sampled data.

		Original Resolution Data				Sub-Sampled Data			
		Accuracy	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
1	Ground		0.69	0.98	0.81		0.65	0.97	0.78
	High Vegetation		0.48	0.47	0.47		0.49	0.48	0.48
	Building		0.96	0.68	0.80		0.94	0.68	0.79
	Macro average	0.78	0.71	0.71	0.69	0.75	0.69	0.71	0.68
2	Ground		0.95	0.57	0.71		0.88	0.54	0.67
	High Vegetation		0.57	0.76	0.65		0.74	0.78	0.76
	Building		0.37	0.76	0.50		0.39	0.75	0.51
	Macro average	0.64	0.63	0.70	0.62	0.67	0.67	0.69	0.65
3	Ground		0.46	0.54	0.50		0.48	0.56	0.52
	High Vegetation		0.47	0.93	0.63		0.58	0.92	0.72
	Building		0.86	0.31	0.45		0.84	0.31	0.46
	Macro average	0.53	0.60	0.59	0.53	0.59	0.64	0.60	0.56
4	Ground		0.48	0.56	0.52		0.90	0.15	0.25
	High Vegetation		0.58	0.92	0.72		0.85	0.94	0.89
	Building		0.84	0.31	0.46		0.03	0.24	0.05
	Macro average	0.59	0.64	0.60	0.56	0.78	0.59	0.44	0.40

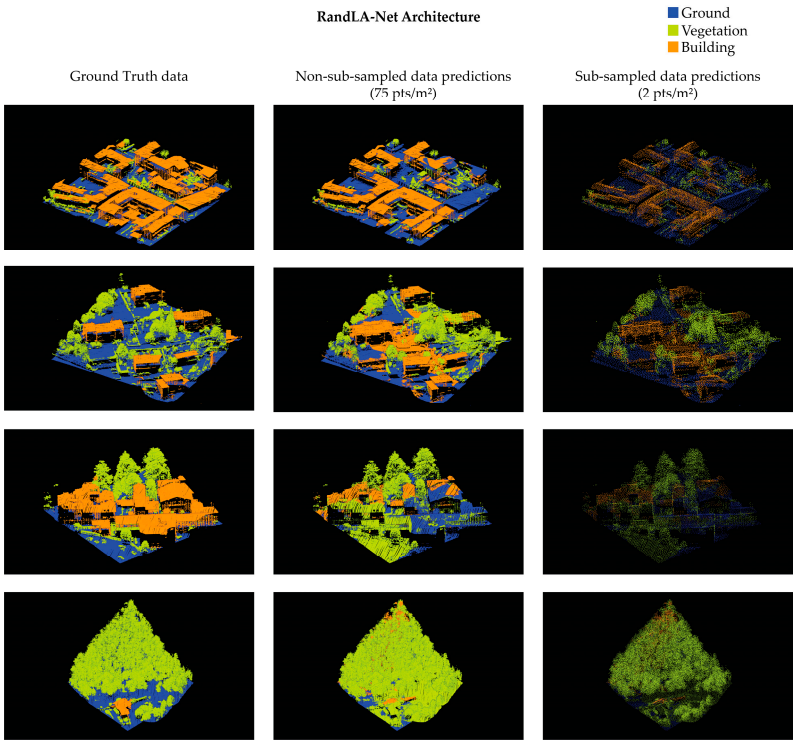


Figure A1. Prediction results of the trained model on four test datasets with RandLA-Net architecture.

Table A2. Accuracy metric assessment of the trained Model 2 with the PointCNN architecture, applied to four test areas. The metrics have been analyzed by comparing the adopted strategy for both original resolution and sub-sampled data.

		Original Resolution Data				Sub-Sampled Data			
		Accuracy	**Precision**	**Recall**	**F1 Score**	**Accuracy**	**Precision**	**Recall**	**F1 Score**
1	Ground		0.73	0.98	0.84		0.69	0.60	0.64
	High Vegetation		0.55	0.52	0.54		0.19	0.90	0.32
	Building		0.98	0.75	0.85		0.97	0.33	0.49
	Macro average	0.83	0.75	0.75	0.74	0.48	0.62	0.61	0.48
2	Ground		0.91	0.35	0.51		0.96	0.31	0.47
	High Vegetation		0.96	0.60	0.74		0.79	0.86	0.82
	Building		0.25	0.99	0.40		0.32	0.85	0.46
	Macro average	0.51	0.71	0.65	0.55	0.62	0.69	0.67	0.58
3	Ground		0.50	0.65	0.56		0.64	0.57	0.60
	High Vegetation		0.86	0.86	0.86		0.53	0.97	0.69
	Building		0.67	0.54	0.60		0.85	0.26	0.39
	Macro average	0.65	0.68	0.68	0.67	0.59	0.67	0.60	0.56
4	Ground		0.99	0.23	0.37		0.99	0.04	0.07
	High Vegetation		0.85	0.98	0.91		0.84	1.00	0.91
	Building		0.09	1.00	0.17		0.18	0.84	0.29
	Macro average	0.68	0.64	0.73	0.49	0.81	0.67	0.63	0.43

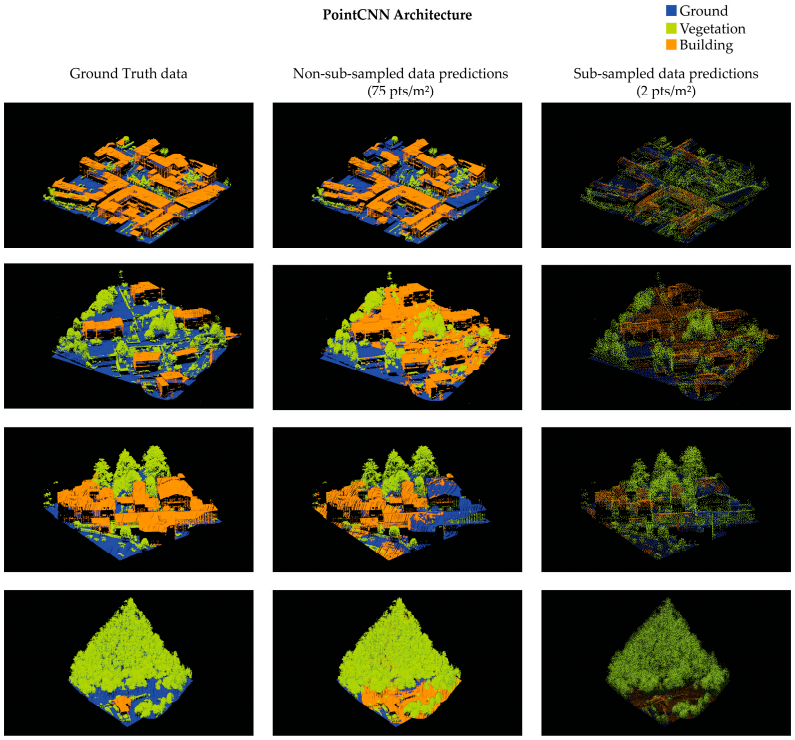


Figure A2. Prediction results of the trained model on test datasets with PointCNN architecture.

References

1. Robinson, H. *Morphology and Landscape*, 3rd ed.; University Tutorial Press: York, UK, 1977; ISBN 0723107564/9780723107569.
2. Shukla, T.; Tang, W.; Trettin, C.C.; Chen, G.; Chen, S.; Allan, C. Quantification of Microtopography in Natural Ecosystems Using Close-Range Remote Sensing. *Remote Sens.* **2023**, *15*, 2387. [CrossRef]
3. Turitto, O.; Baldo, M.; Audisio, C.; Lollino, G. A LiDAR Application to Assess Long-Term Bed-Level Changes in a Cobble-Bed River: The Case of the Orco River (North-Western Italy). *Geogr. Fis. Din. Quat.* **2010**, *33*, 61–76.
4. Hemingway, H.; Opalach, D. Integrating Lidar Canopy Height Models with Satellite-Assisted Inventory Methods: A Comparison of Inventory Estimates. *For. Sci.* **2024**, *70*, 2–13. [CrossRef]
5. Kerr, J.T.; Ostrovsky, M. From Space to Species: Ecological Applications for Remote Sensing. *Trends Ecol. Evol.* **2003**, *18*, 299–305. [CrossRef]
6. Mancini, F.; Dubbini, M.; Gattelli, M.; Stecchi, F.; Fabbri, S.; Gabbianelli, G. Using Unmanned Aerial Vehicles (UAV) for High-Resolution Reconstruction of Topography: The Structure from Motion Approach on Coastal Environments. *Remote Sens.* **2013**, *5*, 6880–6898. [CrossRef]
7. Corte, E.; Ajmar, A.; Camporeale, C.; Cina, A.; Coviello, V.; Giulio Tonolo, F.; Godio, A.; Macelloni, M.M.; Tamea, S.; Vergnano, A. Multitemporal Characterisation of a Proglacial System: A Multidisciplinary Approach. *Earth Syst. Sci. Data Discuss.* **2023**, *16*, 3283–3306. [CrossRef]
8. Petschko, H.; Bell, R.; Glade, T. Effectiveness of Visually Analyzing LiDAR DTM Derivatives for Earth and Debris Slide Inventory Mapping for Statistical Susceptibility Modeling. *Landslides* **2016**, *13*, 857–872. [CrossRef]
9. Scaioni, M.; Longoni, L.; Melillo, V.; Papini, M. Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives. *Remote Sens.* **2014**, *6*, 9600–9652. [CrossRef]
10. Forte, M.; Campana, S. (Eds.) *Digital Methods and Remote Sensing in Archaeology*; Springer: Berlin/Heidelberg, Germany, 2016. [CrossRef]
11. Colomina, I.; Molina, P. Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [CrossRef]
12. Sammartano, G.; Chiabrando, F.; Spanò, A. Oblique Images and Direct Photogrammetry with a Fixed Wing Platform: First Test and Results in Hierapolis of Phrygia (TK). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B2-2020*, 75–82. [CrossRef]
13. Vavulin, M.V.; Chugunov, K.V.; Zaitceva, O.V.; Vodyasov, E.V.; Pushkarev, A.A. UAV-Based Photogrammetry: Assessing the Application Potential and Effectiveness for Archaeological Monitoring and Surveying in the Research on the 'Valley of the Kings' (Tuva, Russia). *Digit. Appl. Archaeol. Cult. Herit.* **2021**, *20*, e00172. [CrossRef]
14. Lovitt, J.; Rahman, M.M.; McDermid, G.J. Assessing the Value of UAV Photogrammetry for Characterizing Terrain in Complex Peatlands. *Remote Sens.* **2017**, *9*, 715. [CrossRef]

15. Kalacska, M.; Arroyo-Mora, J.P.; Lucanus, O. Comparing UAS LiDAR and Structure-from-Motion Photogrammetry for Peatland Mapping and Virtual Reality (VR) Visualization. *Drones* **2021**, *5*, 36. [CrossRef]
16. Diara, F.; Roggero, M. Quality Assessment of DJI Zenmuse L1 and P1 LiDAR and Photogrammetric Systems: Metric and Statistics Analysis with the Integration of Trimble SX10 Data. *Geomatics* **2022**, *2*, 254–281. [CrossRef]
17. Mazzacca, G.; Grilli, E.; Cirigliano, G.P.; Remondino, F.; Campana, S. Seeing among Foliage with LiDaR and Machine Learning: Towards a Transferable Archaeological Pipeline. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLVI-2-W1-2022*, 365–372. [CrossRef]
18. Doneus, M.; Briese, C.; Fera, M.; Janner, M. Archaeological Prospection of Forested Areas Using Full-Waveform Airborne Laser Scanning. *J. Archaeol. Sci.* **2008**, *35*, 882–893. [CrossRef]
19. Poux, F.; Billen, R. Voxel-Based 3D Point Cloud Semantic Segmentation: Unsupervised Geometric and Relationship Featuring vs Deep Learning Methods. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 213. [CrossRef]
20. Zhao, F.; Xiong, L.Y.; Wang, C.; Wang, H.R.; Wei, H.; Tang, G.A. Terraces Mapping by Using Deep Learning Approach from Remote Sensing Images and Digital Elevation Models. *Trans. GIS* **2021**, *25*, 2438–2454. [CrossRef]
21. Cappellazzo, M.; Baldo, M.; Sammartano, G.; Spano, A. Integrated Airborne LiDAR-UAV Methods for Archaeological Mapping in Vegetation-Covered Areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-2–2023*, 357–364. [CrossRef]
22. Albrecht, C.M.; Fisher, C.; Freitag, M.; Hamann, H.F.; Pankanti, S.; Pezzutti, F.; Rossi, F. Learning and Recognizing Archeological Features from LiDAR Data. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; IEEE: New York, NY, USA, 2020; pp. 5630–5636. [CrossRef]
23. Sithole, G.; Vosselman, G. Comparison of Filtering Algorithms. *Proc. ISPRS Work. Group III/3 Workshop* **2003**, *34*, 71–78.
24. Graham, L. LAS 1.4 Specification. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 93–102.
25. Maas, H.-G.; Vosselman, G. *Airborne and Terrestrial Laser Scanning*; Whittles Publishing: Caithness, UK, 2010; ISBN 9781904445876.
26. Kobler, A.; Pfeifer, N.; Ogrinc, P.; Todorovski, L.; Oštir, K.; Džeroski, S. Repetitive Interpolation: A Robust Algorithm for DTM Generation from Aerial Laser Scanner Data in Forested Terrain. *Remote Sens. Environ.* **2007**, *108*, 9–23. [CrossRef]
27. Chen, Z.; Gao, B.; Devereux, B. State-of-the-Art: DTM Generation Using Airborne LIDAR Data. *Sensors* **2017**, *17*, 150. [CrossRef] [PubMed]
28. Mokarram, M.; Sathyamoorthy, D. A Review of Landform Classification Methods. *Spat. Inf. Res.* **2018**, *26*, 647–660. [CrossRef]
29. Colucci, E.; Xing, X.; Kokla, M.; Mostafavi, M.A.; Noardo, F.; Spanò, A. Ontology-Based Semantic Conceptualisation of Historical Built Heritage to Generate Parametric Structured Models from Point Clouds. *Appl. Sci.* **2021**, *11*, 2813. [CrossRef]
30. Cattaneo, N.; Di Maria, F.; Guzzetti, F.; Privitera, A.; Righetti, G. Lidar Mapping Technology to Populate Green Areas GIS. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2011**, *XXXVIII-4/C21*, 79–82. [CrossRef]
31. Masini, N.; Coluzzi, R.; Lasaponara, R.; Masini, N.; Coluzzi, R.; Lasaponara, R. On the Airborne Lidar Contribution in Archaeology: From Site Identification to Landscape Investigation. In *Laser Scanning, Theory and Applications*; BoD–Books on Demand: Norderstedt, Germany, 2011. [CrossRef]
32. Doneus, M.; Neubauer, W. 3D Laser Scanners on Archaeological Excavations. In Proceedings of the CIPA 2005 XX International Symposium 2005: International Cooperation to Save the World's Cultural Heritage, Torino, Italy, 26 September–1 October 2005; pp. 1–6.
33. Doneus, M.; Mandlburger, G.; Doneus, N. Archaeological Ground Point Filtering of Airborne Laser Scan Derived Point-Clouds in a Difficult Mediterranean Environment. *J. Comput. Appl. Archaeol.* **2020**, *3*, 92–108. [CrossRef]
34. Fernandez-Diaz, J.; Carter, W.; Shrestha, R.; Glennie, C. Now You See It... Now You Don't: Understanding Airborne Mapping LiDAR Collection and Data Product Generation for Archaeological Research in Mesoamerica. *Remote Sens.* **2014**, *6*, 9951–10001. [CrossRef]
35. Haryuatmanto, G. Analysis of Airborne LiDAR Data for Archaeology Study Case: Sriwijaya Muaro Jambi Site. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1127*, 012012. [CrossRef]
36. Golden, C.; Scherer, A.K.; Schroder, W.; Murtha, T.; Morell-Hart, S.; Fernandez Diaz, J.C.; Jiménez Álvarez, S.D.P.; Alcover Firpi, O.; Agostini, M.; Bazarsky, A.; et al. Airborne Lidar Survey, Density-Based Clustering, and Ancient Maya Settlement in the Upper Usumacinta River Region of Mexico and Guatemala. *Remote Sens.* **2021**, *13*, 4109. [CrossRef]
37. Lu, D.; Weng, Q. A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [CrossRef]
38. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [CrossRef]
39. Alifu, H.; Vuillaume, J.F.; Johnson, B.A.; Hirabayashi, Y. Machine-Learning Classification of Debris-Covered Glaciers Using a Combination of Sentinel-1/-2 (SAR/Optical), Landsat 8 (Thermal) and Digital Elevation Data. *Geomorphology* **2020**, *369*, 107365. [CrossRef]
40. Maxwell, A.E.; Warner, T.A. Differentiating Mine-Reclaimed Grasslands from Spectrally Similar Land Cover Using Terrain Variables and Object-Based Machine Learning Classification. *Int. J. Remote Sens.* **2015**, *36*, 4384–4410. [CrossRef]
41. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
42. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychol. Rev.* **1958**, *65*, 386–408. [CrossRef]

43. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
44. Vargas, R.; Mosavi, A.; Ruiz, R. Deep Learning: A Review. *Preprints* **2018**, 2018100218. [CrossRef]
45. Cortes, C.; Vapnik, V.; Saitta, L. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
46. Maity, A. Supervised Classification of RADARSAT-2 Polarimetric Data for Different Land Features. *arXiv* **2016**, arXiv:1608.00501.
47. Zhou, F.; Zou, L.; Liu, X.; Zhang, Y.; Meng, F.; Xie, C.; Zhang, S. Microlandform Classification Method for Grid DEMs Based on Support Vector Machine. *Arab. J. Geosci.* **2021**, *14*, 1269. [CrossRef]
48. Gawior, D.; Rutkiewicz, P.; Malik, I.; Wistuba, M. Contribution to Understanding the Post-Mining Landscape—Application of Airborn LiDAR and Historical Maps at the Example from Silesian Upland (Poland). *AIP Conf. Proc.* **2017**, *1906*, 170017. [CrossRef]
49. Mukherjee, I.; Singh, U.K. Delineation of Groundwater Potential Zones in a Drought-Prone Semi-Arid Region of East India Using GIS and Analytical Hierarchical Process Techniques. *Catena* **2020**, *194*, 104681. [CrossRef]
50. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.Y.; et al. Segment Anything. In Proceedings of the IEEE International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 3992–4003. [CrossRef]
51. Brusco, D.; Belcore, E.; Piras, M. Popillia Japonica Newman Detection Through Remote Sensing and AI Computer Vision. In Proceedings of the 2023 IEEE Conference on AgriFood Electronics (CAFE 2023), Torino, Italy, 25–27 September 2023; pp. 50–54. [CrossRef]
52. Patrucco, G.; Setragno, F. Multiclass Semantic Segmentation for Digitisation of Movable Heritage Using Deep Learning Techniques. *Virtual Archaeol. Rev.* **2021**, *12*, 85–98. [CrossRef]
53. Felicetti, A.; Paolanti, M.; Zingarettia, P.; Pierdicca, R.; Malinverni, E.S. Mo.Se.: Mosaic Image Segmentation Based on Deep Cascading Learning. *Virtual Archaeol. Rev.* **2021**, *12*, 25–38. [CrossRef]
54. Zhang, Y.; Zhao, W.; Sun, B.; Zhang, Y.; Wen, W. Point Cloud Upsampling Algorithm: A Systematic Review. *Algorithms* **2022**, *15*, 124. [CrossRef]
55. Xu, S.; Vosselman, G.; Oude Elberink, S. Multiple-Entity Based Classification of Airborne Laser Scanning Data in Urban Areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 1–15. [CrossRef]
56. Sarker, S.; Sarker, P.; Stone, G.; Gorman, R.; Tavakkoli, A.; Bebis, G.; Sattarvand, J. A Comprehensive Overview of Deep Learning Techniques for 3D Point Cloud Classification and Semantic Segmentation. *Mach. Vis. Appl.* **2024**, *35*, 67. [CrossRef]
57. Matrone, F.; Grilli, E.; Martini, M.; Paolanti, M.; Pierdicca, R.; Remondino, F. Comparing Machine and Deep Learning Methods for Large 3D Heritage Semantic Segmentation. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 535. [CrossRef]
58. Yang, S.; Hou, M.; Li, S. Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review. *Remote Sens.* **2023**, *15*, 548. [CrossRef]
59. Pingel, T.J.; Clarke, K.C.; McBride, W.A. An Improved Simple Morphological Filter for the Terrain Classification of Airborne LIDAR Data. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 21–30. [CrossRef]
60. Hackel, T.; Wegner, J.D.; Schindler, K. Contour Detection in Unstructured 3D Point Clouds. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1610–1618. [CrossRef]
61. Nurunnabi, A.; Teferle, N.; Balado, J.; Chen, M.; Poux, F.; Sun, C. Robust Techniques for Building Footprint Extraction in Aerial Laser Scanning 3D Point Clouds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLVIII-3/W2-2022*, 43–50. [CrossRef]
62. Kölle, M.; Laupheimer, D.; Schmohl, S.; Haala, N.; Rottensteiner, F.; Wegner, J.D.; Ledoux, H. The Hessigheim 3D (H3D) Benchmark on Semantic Segmentation of High-Resolution 3D Point Clouds and Textured Meshes from UAV LiDAR and Multi-View-Stereo. *ISPRS Open J. Photogramm. Remote Sens.* **2021**, *1*, 100001. [CrossRef]
63. Zhang, X.; Xue, R.; Soergel, U. A TWO-STAGE APPROACH FOR RARE CLASS SEGMENTATION IN LARGE-SCALE URBAN POINT CLOUDS. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.—ISPRS Arch.* **2022**, *43*, 329–334. [CrossRef]
64. Laupheimer, D.; Haala, N. Multi-Modal Semantic Mesh Segmentation in Urban Scenes. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *5*, 267–274. [CrossRef]
65. Nong, X.; Bai, W.; Liu, G. Airborne LiDAR Point Cloud Classification Using PointNet++ Network with Full Neighborhood Features. *PLoS ONE* **2023**, *18*, e0280346. [CrossRef]
66. Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. PointCNN: Convolution On X-Transformed Points. *arXiv* **2018**, arXiv:1801.07791.
67. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 11105–11114. [CrossRef]
68. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [CrossRef]
69. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. Learning Semantic Segmentation of Large-Scale Point Clouds with Random Sampling. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 8338–8354. [CrossRef]
70. Ballouch, Z.; Hajji, R.; Poux, F.; Kharroubi, A.; Billen, R. A Prior Level Fusion Approach for the Semantic Segmentation of 3D Point Clouds Using Deep Learning. *Remote Sens.* **2022**, *14*, 3415. [CrossRef]
71. Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; Bennamoun, M. Deep Learning for 3D Point Clouds: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 4338–4364. [CrossRef]

72. Varney, N.; Asari, V.K.; Graehling, Q. DALES: A Large-Scale Aerial LiDAR Data Set for Semantic Segmentation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 717–726. [CrossRef]
73. Matrone, F.; Lingua, A.; Pierdicca, R.; Malinverni, E.S.; Paolanti, M.; Grilli, E.; Remondino, F.; Murtiyoso, A.; Landes, T. A Benchmark for Large-Scale Heritage Point Cloud Semantic Segmentation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B2-2020*, 1419–1426. [CrossRef]
74. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Contextual Classification of Lidar Data and Building Object Detection in Urban Areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 152–165. [CrossRef]
75. Jenson, S.K.; Domingue, J.O. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1593–1600.
76. Pomerleau, F.; Liu, M.; Colas, F.; Siegwart, R. Challenging Data Sets for Point Cloud Registration Algorithms. *Int. J. Robot. Res.* **2012**, *31*, 1705–1711. [CrossRef]
77. Pal, M.; Mather, P.M. Support Vector Machines for Classification in Remote Sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011. [CrossRef]
78. Tharwat, A. Classification Assessment Methods. *Appl. Comput. Inform.* **2018**, *17*, 168–192. [CrossRef]
79. Pellis, E.; Masiero, A.; Grussenmeyer, P.; Betti, M.; Tucci, G. Synthetic Data Generation and Testing for the Semantic Segmentation of Heritage Buildings. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-2–2023*, 1189–1196. [CrossRef]
80. Patrucco, G.; Setragno, F. Enhancing Automation of Heritage Processes: Generation of Artificial Training Datasets from Photogrammetric 3D Models. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2023**, *XLVIII-M-2–2023*, 1181–1187. [CrossRef]
81. Iman, M.; Arabnia, H.R.; Rasheed, K. A Review of Deep Transfer Learning and Recent Advancements. *Technologies* **2023**, *11*, 40. [CrossRef]
82. Han, X.; Zhang, Z.; Ding, N.; Gu, Y.; Liu, X.; Huo, Y.; Qiu, J.; Yao, Y.; Zhang, A.; Zhang, L.; et al. Pre-Trained Models: Past, Present and Future. *AI Open* **2021**, *2*, 225–250. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.