SEEING AMONG FOLIAGE WITH LIDAR AND MACHINE LEARNING: TOWARDS A TRANSFERABLE ARCHAEOLOGICAL PIPELINE

G. Mazzacca1,2, E. Grilli1, G.P. Cirigliano3, F. Remondino1, S. Campana3

1 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Web: http://3dom.fbk.eu  Email: <grilli><gmazzacca><remondino>@fbk.eu
2 Centre of Geotechnologies, University of Siena, Italy
3 Department of History and Cultural Heritage, University of Siena, Italy
Email: cirigliano@lapetlab.it stefano.campana@unisi.it

Commission II

KEY WORDS: point cloud classification, machine learning, remote sensing, landscape archaeology

ABSTRACT:
Airborne LiDAR technology has become an essential tool in archaeology during the last two decades since it allows archaeologists to measure and map items or structures that would otherwise be hidden under vegetation. In order to detect and characterise the archaeological evidence, it is a common practice to extract accurate digital terrain models (DTM) by filtering out the vegetation from Airborne Laser Scanning (ALS) datasets. Although previous approaches have performed well in ALS filtration, they are still subject to several variables (flight height, forest cover, type of sensors utilised, etc.) and are frequently integrated into expensive commercial software or customised for specific locations. This study presents a workflow for treating ALS archaeological datasets using machine learning algorithms for both filtering the vegetation and detecting hidden structures. The workflow is applied to two different archaeological environments (in terms of structures, vegetation, landscape, point density), and results demonstrate that the pipeline is rapid and accurate, and the prediction model is transferable.

1. INTRODUCTION

The use of Airborne Laser Scanning (ALS) and Light Detection And Ranging (LiDAR) technologies is nowadays a well-established practice for identifying and mapping archaeological evidence (Chen et al., 2017; Opitz, 2016; Zhang et al., 2016). The ability to filter the returning signal created by the hit vegetation makes it an essential instrument in areas with dense forest or shrub cover. At the same time, the rising of new point cloud classification methods to identify structures on the ground or separate surface and terrain data is offering various advantages to archaeologists to see through the foliage (Chen et al., 2017; Historic England, 2018; Štular and Lozić, 2020; Štular et al., 2021). Numerous studies in the field use Digital Terrain Models (DTMs) to identify anomalies of archaeological interest at the topographical and geomorphological levels (Barbour et al., 2019; Thompson, 2020). However, the generation of an accurate DTM from LiDAR data relies on many factors, such as flight height, forest cover and density, used pulse, season conditions, terrain slope, etc. (Hyppölä et al., 2004). Consequently, the creation of a DTM becomes a complex procedure that requires numerous assumptions and decisions during the project planning, data acquisition and subsequent analytic workflow (Doneus et al., 2020).

1.2 Aim of the work

The most requested output from a LiDAR survey in an archaeological context under vegetation is to map both archaeological elements above ground and anomalies of potential historical interest at ground level (i.e., shallow walls covered by vegetation). The task is challenging for different reasons:

- dataset dimensions, generally spanning hectares and billions of points;
- the possible presence of steep slopes;
- the amount and type of vegetation that covers the archaeological areas;
- the possible types of archaeological structures present under the vegetation.

This work aims to outline an optimal, fast and transferable pipeline that works in different archaeological environments to distinguish vegetated and non-vegetated areas in LiDAR datasets.

1.3 Datasets

This study considers two different archaeological sites surveyed with UAV-based LiDAR sensor (RIEGL miniVUX-2 / miniVUX-3). The two sites are particularly complex in terms of extension, forest cover and terrain roughness:

- **San Martino ai Campi** archaeological area (Riva del Garda, Trento, Italy) – Dataset_1: it is a complex, multiperiod settlement of military and religious activities dated from the late Iron Age to the Middle Ages (Belloni et al., 2013). The alpine area is characterised by steep slopes, rocky walls, and dense forest cover, primarily of tall trees, located on high land. All these factors make automatic archaeological evidence recognition extremely challenging. The available LiDAR data occupies an area of ca 4 ha and features about 500pt/sqm density.

- **Roselle** archaeological area (Grosseto, Italy) – Dataset_2: the Etruscan and Roman city of *Rosellae* has been relatively well explored archaeologically, with a site history that spans the eighth/ninth centuries BC to the early twelfth century AD, with well-defined phases of foundation, elaboration, and
decline (Nicosia et Poggesi, 1998; Campana, 2017). Unlike the archaeological area of San Martino, Roselle is set in a more Mediterranean and hilly context, characterised by the presence of tall vegetation but also dense shrubbery, a characteristic that makes challenging the extraction of a free-of-noise DTM. The available LiDAR data occupies an area of ca 800 ha and features approximately a 300 pt/sqm density.

2. RELATED WORKS

The processing and interpretation of aerial point clouds for archaeological research resulted in various methodologies tailored to the reference context, aiming to assess the archaeological potential of an area and map its elements (Doneus and Briese, 2011; Fernandez-Díaz et al., 2014; Doneus et al., 2020). Most studies rely on visualisation techniques (VTs) such as hillshade or slope, generated from DTMs, to improve the interpretation of the archaeological areas (Grammer et al., 2017; Thompson, 2020). It is essential to specify that in the archaeological context, when a DTM is used to identify archaeological structures, then standing stones, walls, roads, canals, earthworks, and similar should overcome any filtering and be reflected in the final DTM (Doneus et al., 2020). Although a large body of algorithms has been proposed for ALS filtering (Pfeifer et al., 2014; Chen et al., 2017), the DTM generation for archaeological purposes itself is, most of the time, an “unspoken” topic. Indeed, in most articles, this process is carried out using licensed software and a series of manual filtering operations. However, as Doneus et al. (2020) pointed out, ground point filtering within commercial software is a black-box, and the risk of removing archaeological information is significant.

Different researchers have already focused on archaeological evidence detection at a raster/2D level (Figure 1), adopting machine learning (Freeland et al., 2016; Guyot et al., 2018; Niculiță, 2020; Rom et al., 2020) and deep learning (Albrecht et al., 2019; Trier et al., 2021) strategies, whereas only a few studies aimed at automatic filtering/segmenting point clouds (Hu and Yuan, 2016; Geveart et al., 2018; Bulatov et al., 2021).

![Figure 1. A typical pipeline for processing ALS datasets. The filtered 2.5D DTM is used with visualisation techniques (VTs) and machine learning (ML) methods to detect archaeological evidence.](Image)

Contrary to previous works, this paper proposes a pipeline (Figure 2) based on artificial intelligence (AI) algorithms for both filtering out the vegetation and detecting the archaeological structures directly processing LiDAR point clouds. Generalisation and transferability capabilities are also evaluated. Following this state-of-the-art section, the article describes the proposed approach for processing archaeological ALS datasets (Section 3), reporting the experiments and outcomes for two separate datasets (Section 4). Finally, a discussion is given summarising the acquired knowledge (Section 5).

3. METHODOLOGY

The developed pipeline for processing and interpreting archaeological ALS datasets (Figure 2) can be summarised in three main steps:

1. multi-level multi-resolution (MLMR) point cloud semantic segmentation (Section 3.1);
2. bare-ground DTM generation;
3. VTs for anomaly detection (Section 3.2).

![Figure 2. The proposed MLMR pipeline for processing an archaeological ALS dataset and extract hidden man-made structures under vegetation.](Image)

3.1 MLMR semantic segmentation

The 3D dataset is firstly classified using a multi-level multi-resolution (MLMR) procedure (Teruggi et al., 2020), involving the generation of semantic segmentation models at two different levels of geometric resolution. The first level aims at filtering out the vegetation from the rest of the datasets. The second level focuses on the identification of the archaeological structures, and it is typically performed at a higher resolution. The predictive models are developed for both levels by training a Random Forest (RF) algorithm (Breiman, 2001) with reduced manually annotated portions of the datasets, including known elements and the most discriminative features. The predictive models are then validated on a manually classified portion of the dataset (validation set), using standard machine learning metrics and finally used to extend the classification to the whole dataset (Figure 3). In both levels, three main groups of features are used: LiDAR features, geometric features and height-based features. The geometric features selected for the task are, in particular, linearity, planarity, sphericity, surface variation, and verticality extracted in a multi-scale approach (Weinmann et al., 2015). For the purpose of the classification, the introduction of accurate height-based features is fundamental to recognise and correctly classify structures above and at ground level.
Among these features, the "heigh-above" and "heigh-below" are particularly effective in distinguishing between structure and ground classes. In addition, it is beneficial the calculation of the absolute distance between the point cloud and a mesh (C2M), representing an approximation of the bare ground. The last one is generated using the Cloth Simulation Filter (CSF) plugin (Zhang et al., 2016) within the CloudCompare software. The CSF algorithm is chosen as it allows the selection of different rigidness parameters according to the slope degrees. Between the first (L1) and the second (L2) level of the classification procedure, the vegetation class is furtherly checked to verify the presence of possible "lost" structures and to correct the point cloud classification accordingly. Geometric features such as linearity and surface variation are calculated to accomplish this task. After the second classification level is accomplished, a DTM is generated using the class ground, interpolating the minimum values on the z-axis. The DTM resolution can vary according to the aims of the subsequent research.

3.2 Visualisation techniques (VTs) for anomaly detection

According to the study’s needs, the obtained DTM is imported into a GIS environment for different VTs and highlight several features. For example, the use of shade or slope functions aid in the detection of previously unclassified structures or anomalies which appear only at the micro-relief level. Moreover, it is also possible to adopt automated solutions based on convolutional neural networks (CNN) for object detections (Kazimi et al., 2018; Verschoof-van der Vaart and Lambers, 2021). The network’s output would be a probability map for ground-level anomalies, helpful for the archaeologists because it allows them to go on-site and verify the presence of archaeological evidence starting from pre-acquired knowledge.

4. EXPERIMENTS AND RESULTS

4.1 San Martino ai Campi - Dataset_1

The pipeline above described was firstly tested on Dataset_1, i.e., an archaeological area located in an alpine context. This area is distinguished by the presence of mainly narrow (from 0.4m to 1m wide) and linear archaeological constructions, with the notable exception of the church’s apse and the buttress walls (circular and more than 1m wide). Many of these elements stand tall above the ground level, but a significant number, especially in the south-eastern area, are preserved only for a few centimetres above the ground surface. Therefore, within the first classification level (L1), at 0.5m resolution, geometric features were calculated at large radii in order to differentiate the vegetation from the rest of the dataset. At the second level (L2) instead, features at smaller radii were used to discriminate more precisely between ground and structure classes, working at a higher resolution (0.2m). The MLMR steps are summarised in Figure 4. The detection of shallow man-made structures remained a complex task to achieve, and the presence of many rocks at ground level (Figure 5c) added a layer of complexity to the procedure. Therefore, the accuracy results achieved for the class structure are lower if compared to ground and vegetation (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>L1 (OA = 98.42%)</th>
<th>L2 (OA = 93.72%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(%)</td>
<td>Non-veg.</td>
<td>Veg.</td>
</tr>
<tr>
<td>Prec.</td>
<td>96.38</td>
<td>98.91</td>
</tr>
<tr>
<td>Rec.</td>
<td>95.60</td>
<td>99.11</td>
</tr>
<tr>
<td>F1</td>
<td>95.99</td>
<td>99.01</td>
</tr>
</tbody>
</table>

Table 1. Overall Accuracy (OA), Precision, Recall and F1-score metrics achieved for Dataset_1 at the first (L1) and second (L2) levels of semantic segmentation.
However, through an on-field verification, the classification results prove to be successful in highlighting portions of structures which were hidden under the vegetation (Figure 5b). The subsequent application of different VTs has finally highlighted the presence of a high-plain region with an almost regular rectangular shape in the northeast area (Figure 6). The detection of linear landforms in this area could possibly hint towards the existence of buried structures. The semantic segmentation results furtherly confirm this observation. Indeed, as shown in Figure 6, in the highlighted red area, the algorithm has actually identified some structures (see purple patches).

### 4.2 Roselle Archaeological Area - Dataset_2

Dataset_2 depicts a different vegetated area with several archaeological structures, especially near the Roman forum. Most of these structures are narrow walls (0.7m to 1m wide), high or low above ground level, except for the city walls and Roman Amphitheatre, which have a more significant structure (the Roman cavea reach 6/7m in width). Instead, the remaining land is thickly vegetated, with only a few archaeological relics visible above ground. However, it is well known that the area is characterised by numerous Etruscan burial mounds (Campana, 2017). In this case, from a structural point of view, burial mounds can be divided into two categories: excavated and not excavated. The latter generally have a circular wall, without a cover and with an opening (the access corridor), showing a concavity in the data (Figure 8), whereas unexcavated mounds have the typical convex shape (Figure 9).

The two models (L1, L2) trained on Dataset_1 were directly applied to Dataset_2 to evaluate how knowledge from a different archaeological location could be directly used to predict heritage evidences in other datasets (Figure 7). The generalisation produced fairly good quality results (Table 2), even if the discrimination of the low-vegetation from the ground proved to be a challenging task. Despite the high accuracy achieved at the first level of segmentation (L1), a portion of low-vegetation that erroneously remained in the class non-vegetation resulted in some false positive structure recognition at L2.
In order to rapidly improve the L2 results, some small but significant portions of Dataset_2 were therefore included in the training set. The combined training \((T_{1+2})\) resulted in an almost 10\% improvement in the F1-scores for the ground class, as shown in Table 2.

$$\begin{array}{cccccc}
\text{T1} & \text{T1+2} \\
\hline
\text{L1} & \text{L2} & \text{L1} & \text{L2} \\
\text{(OA = 94.87\%)} & \text{(OA = 96.58\%)} & \text{(OA = 96.58\%)} & \text{(OA = 96.58\%)} \\
\text{Prec.} & 93.97 & 95.65 & 97.24 & 98.77 & 97.44 & 87.64 \\
\text{Rec.} & 94.92 & 94.82 & 99.18 & 56.71 & 99.46 & 65.23 \\
\text{F1} & 94.44 & 95.23 & 98.2 & 66.98 & 98.44 & 76.44 \\
\end{array}$$

Table 2. Results achieved for Dataset_2 at the first (L1) and second (L2) levels of segmentation/generalisation. In T1, the training set is equal to Dataset_1, while in T\((1+2)\), T1 is integrated with some portion of Dataset_2.

As for the previous experiment, the finally extracted ground was used to generate a DTM. Because of the presence of some noise in the point cloud (low-vegetation remaining), a resolution of 0.5m was chosen as the best compromise for a clean DTM generation. Different VTs were then adopted for enhancing the most useful terrain characteristics to identify the previously mentioned burial mounds.

In particular, using the Relief Visualisation Toolbox (RVT) (Zakšek et al., 2011; Kokalj and Somrak, 2019) within the QGIS environment (QGIS Development Team, 2022), a Simplified Local Relief Model (SLRM) was extracted from the DTM. In this way, we could enhance the visibility of small-scale surface variations. In particular, two separate image outputs were produced starting from the SLRM: (i) a shaded image which facilitates the visualisation of local relief (Figure 8) and (ii) a combination of hill-shaded terrain, slope, openness, and sky view factor suitable for the detection of “concave evidence” which can be interpreted as an excavated burial mound (Figure 9). As future work, we aim to start from these image enhancement results to train e.g. a RCNN network (Girshick et al., 2014) to extend the object detection to the whole area of Roselle.
Figure 8. On the left, the 3D point cloud of a typical concave (excavated) mound. On the right, the raster DTM, visualised with an *hillshaded* filter to enhance the excavated mounds (in yellow brackets).

Figure 9. On the left, the 3D point cloud of a not excavated mound. On the right, the raster DTM, visualised with a combination of *hillshaded terrain*, *slope*, *openness*, and *sky view factor* to highlight the not excavated mounds (in red brackets).

5. CONCLUSIONS

The use of machine learning algorithms demonstrated significant advantages in processing large LiDAR datasets, facilitating the otherwise difficult manual identification of hidden heritage evidence, saving time in the process execution.

The nature of the investigated territories, densely forested or shrub-covered, as well as the presence of significant steep slopes, made terrain filtering and automatic recognition tasks of considerable complexity. In this context, the proposed pipeline was designed in such a way that it could spot for archaeological artifacts both above and below ground in multiple steps and working entirely on the LiDAR point cloud. Moreover, the MLMR approach is designed to facilitate and speed-up the computation process (Teruggi et al., 2020): once the features are extracted, the time for training and testing is in the range of few minutes.

Applying the pipeline on both datasets allowed to verify the ability of the proposed methodology, highlighting both pros and critical issues.

Pros:
- Fast processing of large datasets through a supervised machine learning approach.
- Accurate vegetation filtering in complex environments.
- Detection and mapping of above-ground structures directly on the 3D point cloud.
- Generation of high-resolution DTMs for anomalies and shallow structures detection.
- Output easily transferrable to a GIS environment for further data processing.
- Possibility to generalise and transfer the classification models to different environments.

Critical issues:
The presence of low-vegetation "noise" can induce the false positive detection of structures. Steep slopes and rocks at ground level can be easily misclassified as structures. The geometric features need to be adapted to the context (i.e., wide walls need big-scale features). The implementation of a CNN for anomaly detection will represent the last stage of our proposed pipeline, whose goal is to have an accurate, rapid and adaptive solution, based on open-source tools, for archaeological evidence detection and mapping in complex environments.

ACKNOWLEDGEMENTS

Part of this research paper would have not been possible without the financial support of Global Digital Heritage (https://globaldigitalheritage.org) and Dr. Herbert Maschner covering the cost of data acquisition, processing and interpretation. Special thanks are also addressed to our partner Alto Drones (https://www.alto-drones.com/en/) for both LiDAR acquisitions and to Dr. Nicoletta Pisu (Archaeological Heritage Office of the Department for Cultural Heritage, Trento, Italy) for her valuable discussions and support for the San Martino ai Campi archaeological area.

REFERENCES


