

Semi-automated LiDAR Vegetation Classification for Mediterranean Archaeology: Designing a Pipeline Leveraging a Multi-Layer Stacked Ensemble Approach

Nicola Lercari¹, Aaron Fandrei², Zubin Zellmann², Yiming DU¹, Mina Yacoub¹, Dario Calderone¹,
Rodolfo Brancato⁴, Saverio Scerra⁵, Davide Tanasi⁶, Rosa Lanteri⁷, David Rügamer^{2,3}

¹ LMU Munich, Institute for Digital Cultural Heritage Studies, 80539 Munich, Germany- n.lercari@lmu.de
yiming.du@campus.lmu.de; mina.yacoub@campus.lmu.de; dario.calderone@lmu.de

² LMU Munich, Institute for Statistics, 80539 Munich, Germany- a.fandrei@campus.lmu.de; z.zellmann@campus.lmu.de

³ Munich Center for Machine Learning (MCML), 80539 Munich, Germany- david.ruegamer@stat.uni-muenchen.de

⁴ University of Naples Federico II, 80133 Naples, Italy- rodolfo.brancato@unina.it

⁵ Soprintendenza per i Beni Culturali e Ambientali di Ragusa, 97100 Ragusa, Italy- saverio.scerra@regione.sicilia.it

⁶ University of South Florida, Institute for Digital Exploration, 33620 Tampa, USA- dtanasi@usf.edu

⁷ Parco Archeologico e Paesagg. di Siracusa, Eloro, Villa del Tellaro e Akrai, 96100 Syracuse, Italy- rosa.lanteri@regione.sicilia.it

Keywords: Archaeological LiDAR Classification, Machine Learning, RandLA-Net, Stacking Classifier, XGBoost/CatBoost

Abstract

Dense Mediterranean vegetation often conceals archaeological features in LiDAR data, posing a significant challenge for archaeological analysis. This paper presents a novel machine learning pipeline for semi-automated vegetation classification in drone-based archaeological LiDAR point clouds, which were captured to survey the Mediterranean landscape of Sicily, Italy. Our approach integrates an extensive feature engineering stage with a multi-layer stacked ensemble classifier and a RandLA-Net deep learning model. The pipeline was trained on a semantically annotated drone-based LiDAR dataset from the site of Kamarina. It achieved high accuracy in distinguishing vegetation from ground points (0.99 overall accuracy, weighted macro F1 $\approx$ 0.93). To evaluate generalizability, we tested the model on a secondary site (Heloros) with different vegetation characteristics, obtaining an F1 of $\sim$ 0.70. Qualitative inspection of results confirms that our model effectively removes vegetation while preserving archaeological structures. Our results demonstrate the potential of ensemble learning and 3D deep neural networks in archaeological remote sensing, enabling more efficient visualization and mapping of hidden archaeological features.

1. Introduction

1.1 Archaeological LiDAR Challenges

In archaeological remote sensing, precisely identifying and removing vegetation from airborne laser scanning (ALS), commonly known as Light Detection and Ranging (LiDAR), datasets is crucial for revealing archaeological features that are obscured by vegetation. Despite its costs and logistical complexity, ALS has revolutionized archaeological reconnaissance in heavily forested regions of Mesoamerica, Northern and Central Europe, and Southeast Asia (Cifani, Opitz, and Stoddart, 2007; Chase et al. 2011; Štular et al., 2012; Evans et al., 2013; Risbøl et al., 2020). However, its application to Sicily's Mediterranean landscape remains underexplored. Mapping human-made structures in a 'Mediterranean scrub' environment using LiDAR is particularly challenging, as dense vegetation comprising small trees, thorny shrubs, and bushes obscures archaeological remains significantly impacted by centuries of human activity.

Our previous work has demonstrated that drone-based ALS offers a transformative solution. At the Greek settlement of Kamarina in coastal Southern Sicily, we successfully used ALS to penetrate dense vegetation along the Ippari River, capturing high-density point clouds with a RIEGL VUX-1UAV²² sensor equipped with an APPLANIX APX-20 IMU/GNSS system (Fig. 1). With a measured point distance of approximately 5 cm and a spatial resolution of 400 pts/m², our ALS data surpass publicly available LiDAR in detecting subtle archaeological

features (Brancato et al., 2024; Lercari et al., 2025). At Heloros, another Greek settlement we scanned in Southeastern Sicily, we utilized the same ALS instrument to successfully detect archaeological features beneath the canopy (Calderone et al. 2025). However, distinguishing LiDAR vegetation returns from other point classes (such as ground or building) has traditionally relied on off-the-shelf classifiers (e.g., RIEGL's "Off-Terrain" filter, LAsTools lasclassify, FUSION/LDV), which we found performed inconsistently on Kamarina's data, often misclassifying the site's dense shrubs and even labelling archaeological structures as vegetation, thus complicating manual filtering and further analysis.



Figure 1. The RIEGL VUX-1UAV²² aerial laser scanner and AceCore Noa drone platform we deployed to Kamarina and Heloros to capture the ALS data used in this study.

This underscores the need for a tailored solution. Therefore, the primary objective of our research is to develop a customized machine learning pipeline that integrates advanced feature engineering with a stacked ensemble of classifiers and a 3D deep learning model. We aim to train it on the unique vegetative characteristics of the Mediterranean landscapes of Sicily, to enhance vegetation classification accuracy and facilitate more precise point filtering for visual archaeological analysis and interpretation. In this paper, we present a semi-automated classification framework trained on our Kamarina site's drone-based LiDAR data, which achieved good results in minimizing occlusion and improving vegetation classification accuracy. This ultimately allows us to filter out trees and shrubs, enhancing the manual mapping of hidden archaeological structures in a Geographic Information System (GIS). The framework integrates advanced feature engineering with a stacked ensemble of classifiers and a 3D deep learning model.

2. Methods

2.1 Data Collection and Annotation Workflow

Our study utilizes high spatial resolution drone-based ALS data collected at Kamarina as the primary dataset for training and evaluation. In September 2024, we scanned a 15-hectare area of the ancient Kamarina urban core (Agora and surroundings) using a RIEGL VUX-UAV22 laser scanner mounted on an AceCore Noa uncrewed aerial vehicle (UAV), integrated with a 64 MP RGB camera. This UAV-based ALS survey produced a dense point cloud (with an average point spacing of approximately 5 cm) containing rich geometric and radiometric information. To create ground-truth labels for model training, we developed a structured manual annotation workflow in CloudCompare v2.14 (CloudCompare, Development Team, 2025), following well-established processes (Mazzacca et al. 2022; Cirigliano et al., 2025) (Fig. 2). The raw point cloud was initially segmented into smaller spatial units (by features and clusters). These were then carefully labelled into the following classes: Ground and Vegetation. Vegetation included grasses, shrubs, bushes, and trees. Buildings, archaeological structures, and other man-made objects were grouped as Ground (non-vegetation) points. Labelling was conducted iteratively by multiple human annotators, with cross-checks to ensure consistency. We employed a hierarchical subset approach during labelling—each vegetation patch was isolated into a separate subset to ensure mutually exclusive categories and complete coverage of the area. This strategy enhanced annotation precision and allowed targeted quality control adjustments. External reference data, including archaeological maps, site plans, and satellite imagery, were utilized to guide and validate the manual classifications. After labelling, the Kamarina dataset contained approximately 116 million points, of which about 18 million (15%) were labelled as vegetation and 98 million as ground (including archaeological remains).

Beyond Kamarina, a second site, Heloros, was used to test the model's generalization. The dataset captured in the Marianelli area, located along the rugged coast south of Heloros, is smaller (approximately 0.03 km², containing around 10 million points). The labeled Heloros-Marianelli dataset exhibits a different class balance, with approximately 6.5 million vegetation points compared to 3.7 million ground points (Fig. 7b). The vegetation in this dataset varies in both density and type, offering a valuable contrast to Kamarina for evaluating the robustness of our pipeline. Ground truth for Heloros was obtained through a similar manual process (with the Ground class including archaeological structures as before). No Heloros data were

employed in training; the Heloros/Marianelli dataset serves purely as an independent test to assess how well a model trained on Kamarina can transfer to new Mediterranean landscapes and archaeological sites.

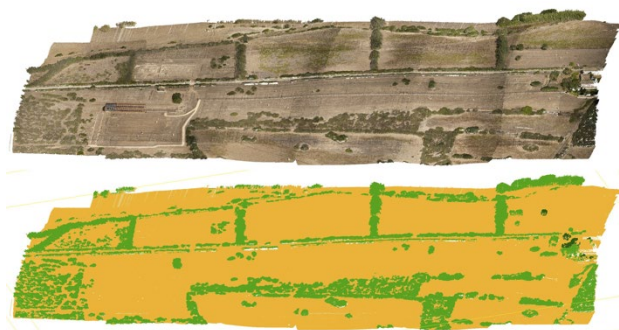


Figure 2. View of a portion of the Kamarina point cloud before (top) and after (bottom) labelling. This figure illustrates how dense vegetation (green points) can obscure the underlying terrain and archaeological structures (yellow points).

2.2 A Semi-automated Pipeline for Semantic Segmentation of Archaeological LiDAR Data using Machine Learning

We developed a semi-automated classification pipeline that combines three key components: (a) extensive feature preprocessing, (b) feature selection with mixed methods, and (c) a multi-layer stacked ensemble model for the final semantic segmentation.

In the first stage (a), the labeled point cloud is spatially partitioned and downsampled to facilitate efficient feature computation and reduce spatial autocorrelation. We divided the dataset into 50m × 50m grid tiles and processed each tile independently. Within each tile, we applied voxel grid filtering to downsample the point density, utilizing a voxel size optimized to balance detail and efficiency. To voxelize our dataset, we employed cubic voxels with sides measuring 15 centimeters. This substantially reduced the total point count while maintaining the overall geometry of vegetation clusters and ground surfaces (Fig. 3). For each remaining point (voxel centroid) in a tile, we computed a rich set of features that capture local geometry and point attributes (Table 1). Following the approach of Thomas et al. (2018), we defined an approximately spherical neighborhood (using a BallTree structure) to collect the 16 nearest neighboring points around each point.



Figure 3: Bird's-eye view of Partitioned Kamarina Grid with 50m x 50m tiles: (Blue) Holdout Test Set and (Pink) Training Set.

In stage (b) of our pipeline, from this neighborhood, we derived various geometric descriptors: point density, local height above ground, normal vector components, and eigenvalue-based shape features such as planarity, sphericity, and verticality. These

features quantify the 3D structure of the local neighborhood, for instance, planar surfaces versus scattered points, which can help distinguish vegetation (often more spatially dispersed or canopy-like) from solid ground or structural surfaces. Additionally, we utilized all available LiDAR return attributes (ASPR Point Data Record Format 7 in LAS), including each point's intensity, RGB color values, return number, and number of returns. We also created a few composite features that combine geometric and spectral information (e.g., mean green sphericity, indicating how strongly the points in a neighborhood form a spherical shape and how green they are on average), hypothesizing that these might be especially indicative of leafy vegetation (Table 1).

Point Attributes	Geometric Properties	Contextual Features
Return Nr.	Sphericity/Planarity	Mean Green Value
RGB Values	Verticality	Mean Sphericity
	Normals (z)	Mean Green
	Density	Sphericity
	Curvature Variance	Mean Local Height
	Local Height	
	Omnivariance	

Table 1: Selected features categorized by type

Given the large initial feature vector, a feature selection step was performed to identify the most informative features and reduce dimensionality. We employed both filter-based criteria and wrapper methods to rank features based on their importance. More specifically, we evaluated the correlation of each feature with the target (vegetation vs. ground). We used permutation-based importance measures (Breiman, 2001) from ensemble tree models, specifically a Random Forest (Breiman, 2001) and an XGBoost model (Chen and Guestrin, 2016), to assess their contribution to prediction accuracy. Features that consistently ranked low were dropped, resulting in a subset of top-performing features for the model. A final manual selection ensured that only the most relevant geometric and radiometric features were retained for classification (Table 1), thereby improving learning efficiency and mitigating overfitting.

2.2.1 Classification Ensemble: In stage (c) of our pipeline, we developed a two-layer stacked ensemble classifier, which represent the core of our classification approach (Fig. 4). In the first layer, we train a heterogeneous ensemble of three base learners in parallel: (1) a shallow Multi-Layer Perceptron (MLP) neural network with three hidden layers, (2) an XGBoost gradient boosting decision tree model, which we used with a default configuration and (3) a CatBoost gradient boosting model (Dorogush et al., 2018), which we also used with a default configuration.

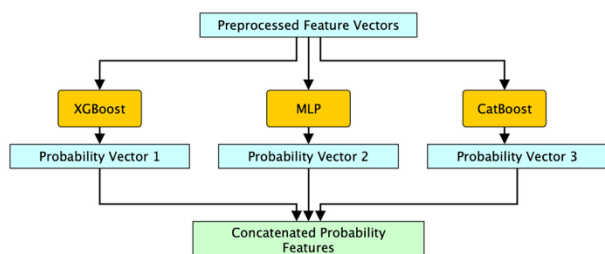


Figure 4. Workflow of Stacking Classifier.

Each base model in the ensemble is trained to predict a binary label (Vegetation vs Ground) for each point using the selected features. Rather than committing to a single model's prediction, we leveraged all three. The probability outputs from each of the three models were fed into a meta-classifier, a logistic regression that learns how to optimally combine them. This stacking approach produces a more robust classification decision for each point.

2.2.2 Integration of Deep Learning (RandLA-Net): A key innovation of our pipeline is the incorporation of a point cloud deep neural network in the second layer of the ensemble to further refine classification. We used the Randomly Subsampled Local Aggregation Network (RandLA-Net) model by Hu et al. (2021), adapted by Gaydon (2022) for urban and terrain aerial LiDAR pointwise semantic segmentation. RandLA-Net is a spatially aware classifier that can learn local context directly from point coordinates and features. While Gaydon (2022) used RandLA-Net as a standalone model, our approach utilizes it as a final classifier, incorporating additional input features from the preprocessing stage and the class probabilities returned by our stacking ensemble. By providing additional signals from our feature engineering and the stacking ensemble, we enable a lightweight RandLA-Net to accurately classify points despite having limited training data (Fig. 5).

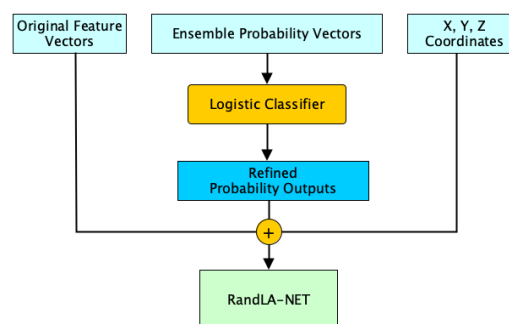


Figure 5. RandLA-Net uses all originally selected features, class probabilities from the stacking ensemble, and coordinates for each LiDAR point as input.

The RandLA-Net architecture comprises two main parts: an encoder and a decoder. The encoder is tasked with downsampling the point cloud to compute dense localized feature representations. Meanwhile, the decoder manages the upsampling process to return to the original resolution. The ensemble predictions are included as additional input features allowing RandLA-Net to leverage the signal from the ensemble while still enabling it to learn complementary geometric patterns from the point cloud. The architecture downsamples and aggregates features through a series of set abstraction layers to capture a broader context, and subsequently upsamples to generate point-wise predictions (Hu et al., 2021). We trained RandLA-NET on the Kamarina data for a limited number of epochs due to the relatively small training area, leveraging the ensemble predictions as additional features. RandLA-Net produces the final output of the pipeline, classifying each voxelized point as either vegetation or ground. After the classification of the voxelized points, the results are projected back onto the original full-resolution point cloud.

2.2.3 Interpolator: Since voxelization removed some points in phase (a) of our pipeline, we needed a method to assign labels to all original points. Various approaches exist for this task, ranging from simple clustering methods to more advanced machine learning techniques. Given the limited redundancy ensured by our voxelization approach, we employed an Inverse Distance Weighted (IDW) nearest neighbor method. For each point, we identified the three nearest neighbors and assigned a label based on their class probabilities. This method was applied to both, the voxelized points already labeled by the model and unlabeled points.

2.2.4 Experimental Setup: For classification, we represented ground or non-vegetation as class 0 and vegetation as class 1. We evaluated performance using accuracy, precision, recall, and F1 score, with particular attention to correctly identifying vegetation. We analyzed the performance of both the stacking ensemble alone and the full pipeline (stacking ensemble plus RandLA-Net). We tested our approach on two datasets: 1) A Kamarina holdout set consisting of 10 random batches of partitioned data (using only the voxelized points for performance metric computation), which included 990,000 ground points and 19,000 vegetation points; 2) The Heloros-Marianelli dataset to assess the model's generalization capabilities to a different site.

3. Results and Analysis

3.1 Performance on Training Site (Kamarina)

We first evaluated the performance of our model on the holdout test set from Kamarina. Looking at the performance on the voxelized points (i.e., not considering the interpolated points), we observed the following results (Table 2)

Ensemble				Ensemble + RandLA-Net		
Class	Prec.	Recall	F1-score	Prec.	Recall	F1-score
Ground	1.00	0.96	0.98	1.00	1.00	1.00
Veg.	0.34	0.96	0.51	0.79	0.94	0.86
Accuracy	0.96		0.99			
Macro	0.74		0.93			
Avg F1						

Table 2. Classification performance on the Kamarina dataset (hold-out test set). The model achieves very high accuracy and F1-score on both classes, especially Ground (non-vegetation).

The heterogeneous ensemble achieved good overall accuracy (0.96), however this metric is not a reliable indicator of true performance due to the class imbalance in the labeled dataset. The precision and F1 score for the vegetation class were poor (0.34 and 0.51, respectively), indicating a high number of false positives.

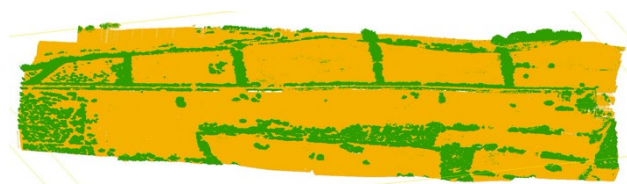


Figure 6. Train and Holdout predictions for Kamarina.

Adding the RandLA-Net significantly improved performance, achieving a precision of 0.79 and an F1 score of 0.86 for vegetation, while also improving overall accuracy to 0.99. The macro-average F1 score improved from 0.74 with just the stacking ensemble to 0.93 with the full pipeline. Thus, the model successfully identified vegetation patterns, including smaller shrubs that may have been missed in manual labeling. Importantly, the model did not misclassify buildings or archaeological structures as vegetation. A visualization of these results shows a close match between our model's predictions and the ground truth labels (Fig. 6 and Fig. 2).

3.2 Generalization to a New Site (Heloros)

To assess the pipeline's generalizability, we applied the trained model (without retraining) to an ALS dataset captured at Heloros in the Marianelli area. This dataset presents a more challenging test due to its distinct vegetation profile and a higher proportion of vegetation points. As shown in Table 3, the results display minimal differences between the ensemble and the full pipeline (with RandLA-Net), with both achieving an F1 score and accuracy of approximately 0.70. While these metrics might initially suggest poor generalization quantitatively, despite the similarity between sites, a visual inspection of the Heloros-Marianelli results revealed a different story. When examining their visualization, we observed that our model did a good job of identifying the main vegetation clusters while correctly ignoring the points underneath the vegetation (Fig. 7).

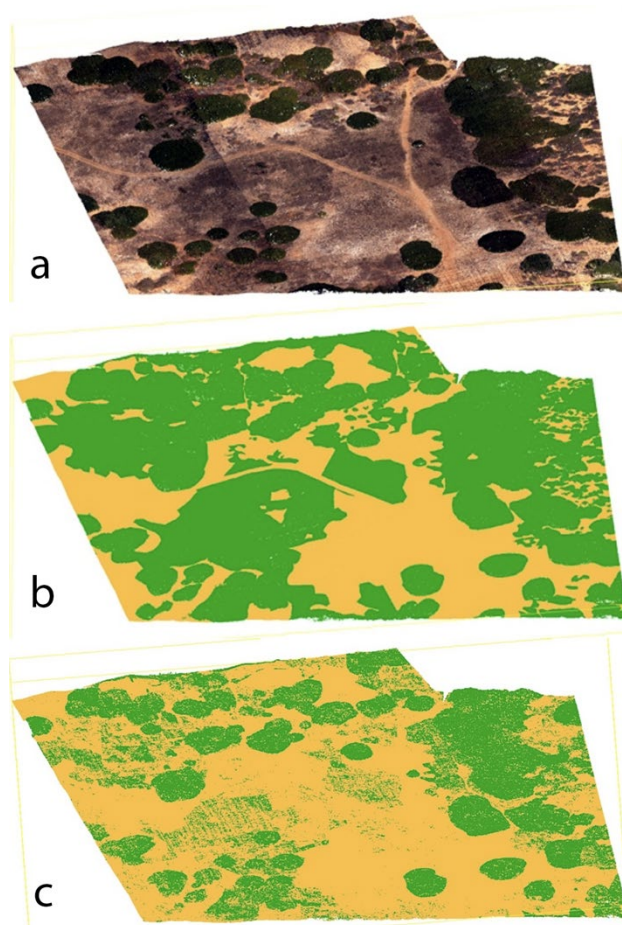


Figure 7. a) View of the Heloros-Marianelli RGB point cloud; b) manual labels; c) predictions.

Ensemble				Ensemble + RandLA-Net		
Class	Prec.	Recall	F1-score	Prec.	Recall	F1-score
Ground	0.45	0.99	0.62	0.55	0.99	0.71
Veg.	0.99	0.61	0.76	0.99	0.53	0.69
Accuracy	0.70		0.70			
Macro	0.69		0.70			
Avg F1						

Table 3. Classification performance comparison on the Heloros-Marianelli dataset.

4. Limitations

4.1 Label Quality

The comparison between predicted labels and ground truth for Heloros revealed significant issues with the quality of manual labelling. The ground truth labels do not account for ground points underneath vegetation or in transition areas, leading to misleading performance metrics. More precise and detailed labeling would be necessary for improving model training, as it helps reduce the noise-to-signal ratio and enables a more accurate evaluation of model performance.

4.2 Training Data Diversity

Although we scanned much larger areas of both Kamarina and Heloros, due to time constraints, we decided to focus on designing and developing the pipeline and models instead of spending hundreds of additional hours labelling all the available datasets. As a result, we had only one labelled dataset from Kamarina for training our model. Including additional training areas from Kamarina and Heloros, as well as new sites with diverse vegetation patterns and archaeological features, would likely enhance the model's generalization capabilities. This will be the focus of a new paper that we plan to finalize after the CIPA 2025 Conference (more details are provided in the Conclusions).

4.3 Design Choices and Hyperparameters

Our implementation includes various design choices and hierarchical implementations that would benefit from systematic evaluation, such as combinations of neighborhood sizes with different features and models. The extensive hyperparameter space we utilized (but omitted here due to the maximum paper length) creates opportunities for further optimization but also poses computational challenges.

4.4 Computational Requirements

The requirement for neighborhood information calculations and the volume of point cloud data restrict the feasibility of training on low-end machines. This establishes a barrier to entry for those with limited computational resources.

5. Conclusions

This study presented a novel semi-automated pipeline for LiDAR vegetation classification, specifically designed for Mediterranean archaeological landscapes. Our machine learning framework effectively tackles the ongoing challenge of segmenting and removing vegetation returns that obscure

ground surfaces in agricultural and rewilding areas, enabling more precise identification of hidden archaeological structures in the data. By integrating an ensemble of classical machine learning models with a point-based deep neural network, our approach capitalizes on the strengths of both: high precision from engineered features and contextual spatial awareness from the neural network. The pipeline's preliminary results (Table 2 and Table 3) highlight the potential of specialized machine learning in advancing archaeological remote sensing while also facilitating the efficient processing of dense 3D data with minimal manual intervention.

While our model generalized moderately well to a second site (Heloros), the experiments also highlighted the challenges of applying specialized classification models across sites. In particular, differences in vegetation types and labeling quality can impact performance. Nonetheless, the model's ability to detect vegetation at Heloros without retraining is an encouraging step toward broader applicability. Moving forward, our research will focus on further enhancing the pipeline's generalizability and capabilities. We plan to incorporate additional sensor modalities, such as Long-Wave Infrared (LWIR) imagery collected at both Kamarina and Heloros, to provide the model with richer information for distinguishing vegetation. We also plan to expand the training dataset with more samples from diverse Mediterranean sites, including additional areas of Kamarina and Heloros, to enhance the model's robustness across various terrains and vegetative conditions. Another approach is to experiment with advanced deep learning architectures or to fine-tune foundation models pre-trained on large LiDAR datasets, such as IGNF's open RandLA-Net model for aerial LiDAR, to determine if they provide gains in our context. In summary, our work demonstrates a practical and accurate pipeline for automated vegetation removal in archaeological LiDAR and sets the foundation for increasingly general and powerful tools to support digital archaeological exploration in vegetated landscapes.

Acknowledgements

The ALS data used in the study were captured under research permits issued by the Soprintendenza per i Beni Culturali e Ambientali di Ragusa and the Parco Archeologico e Paesaggistico di Siracusa in Sicily, for which we are grateful for granting us access to the archaeological sites. We are also thankful to the Munich Center for Machine Learning (MCML) for providing technical expertise and insights into this research.

References

- American Society for Photogrammetry and Remote Sensing. LAS specification 1.4 - r15. Technical report, ASPRS, 2019. URL https://www.asprs.org/wp-content/uploads/2019/07/LAS_1_4_r15.pdf.
- Brancato, R., Giaccoppo, F., Luongo, G., Pianese, C., Roccuzzo, M. C., Scerra, S., 2024. Osservazioni archeologiche e topografiche sull'insediamento d'altura nel territorio ibleo (Sicilia sud-orientale): il caso di Monte Casasia, in: Cannata, A., Scaravilli, M. S., Frasca, M. (Eds.), *Hyblaica. Studi Di Archeologia e Topografia dell'Altopiano Ibleo 2*. Archaeopress, Oxford, pp. 49–66.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45(1), 5–32. DOI: 10.1023/A:1010933404324.

- Calderone, D., Lercari, N., Tanasi, D., Busch, D., Hom, R., Lanteri, R., 2025. Tackling the Thorny Dilemma of Mapping Southeastern Sicily's Coastal Archaeology Beneath Dense Mediterranean Vegetation: A Drone-Based LiDAR Approach. *Archaeological Prospection* 31(3), 1–16. DOI: 10.1002/arp.1956.
- Chase, A. F., D. Z. Chase, J. F. Weishampel, et al. 2011. "Airborne LiDAR, Archaeology, and the Ancient Maya Landscape at Caracol, Belize." *Journal of Archaeological Science* 38, no. 2: 387–398. doi. org/10.1016/j.jas.2010.09.018.
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). ACM, New York, pp. 785–794. doi. org/10.1145/2939672.2939785.
- Cifani, G., R. Opitz, and S. Stoddart. 2007. Mapping the Ager Faliscus Road-System: The Contribution of LiDAR (Light Detection and Ranging) Survey. *J. of Roman Archaeology* 20, no. January: 165–176. doi.org/10.1017/S1047759400005353.
- Cirigliano, G. P., G. Mazzacca, F. Remondino, P. Liverani, G. Cantoro, H. Maschner, and S. Campana. 2025. Drone-Based High-Resolution LiDAR for Undercanopy Archaeology in Mediterranean Environment: Rusellae Case Study (Italy)." *Archaeological Prospection* n/a (n/a). (11 June 2025). doi.org/10.1002/arp.1980.
- CloudCompare, Development Team, 2025. (version 2.X) [GPL software]. Retrieved from <http://www.cloudcompare.org/> (11 June 2025).
- Dorogush, A. V., Ershov, V., Gulin, A., 2018. CatBoost: gradient boosting with categorical features support. arXiv:1810.11363. doi. org/10.48550/arXiv.1810.11363.
- Evans, D. H., R. J. Fletcher, C. Pottier, et al. 2013. Uncovering Archaeological Landscapes at Angkor Using Lidar. *Proceedings of the National Academy of Sciences* 110, no. 31: 12595–12600. doi.org/ 10.1073/pnas.1306539110.
- Gaydon, C., 2022. Myria3D: Deep Learning for the Semantic Segmentation of Aerial LiDAR Point Clouds. (Preprint, French National Mapping Agency IGNF). (11 June 2025)
- Hu, Qingyong, Bo Yang, Linhai Xie, Stefano Rosa, Yulan Guo, Zhihua Wang, Niki Trigoni, and Andrew Markham. 2022. Learning Semantic Segmentation of Large-Scale Point Clouds With Random Sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44 (11): 8338–54. doi.org/10.1109/TPAMI.2021.3083288.
- Lercari, N., Brancato, R., Calderone, D., Luongo, G., Scerra, S., 2025. Digitally Reconstructing an Invisible Polis: Airborne and Drone-based LiDAR Surveying and GIS Integration of Heterogeneous Data for Investigating the Archaeological Topography of Kamarina, Sicily. Presented at the 52nd CAA International Conference 2025 "Digital Horizons: Embracing Heritage in an Evolving World", Athens, Greece.
- Mazzacca, G., E. Grilli, G. P. Cirigliano, F. Remondino, and S. Campana. 2022. Seeing Among Foliage with Lidar and Machine Learning: Towards A Transferable Archaeological Pipeline. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* XLVI-2-W1-2022, 365–72. doi.org/10.5194/isprs-archives-XLVI-2-W1-2022-365-2022.
- Risbøl, O., D. Langhammer, E. S. Mauritsen, and O. Seitsonen. 2020. Employment, Utilization, and Development of Airborne Laser Scanning in Fenno-Scandinavian Archaeology—A Review. *Remote Sensing* 12, no. 9: 1411. doi.org/10.3390/rs12091411.
- Štular, B., Ž. Kokalj, K. Oštir, and L. Nuninger. 2012. Visualization of Lidar-Derived Relief Models for Detection of Archaeological Features. *Journal of Archaeological Science* 39, no. 11: 3354–3360. doi.org/10.1016/j.jas.2012.05.029.
- Thomas, H., Goulette, F., Deschaud, J. E., Marcotegui, B., LeGall, Y. 2018. Semantic classification of 3D point clouds with multiscale spherical neighborhoods. In *2018 International conference on 3D vision (3DV)*, 390–398. IEEE.