Journal of Archaeological Science: Reports 71 (2026) 105733

Contents lists available at ScienceDirect

Journal of Archaeological Science: Reports

journal homepage: www.elsevier.com/locate/jasrep

Automatic detection of archaeological features from airborne laser scanning data using deep learning: Development and evaluation of the ADAF tool

Nejc Čož [a,*], Anthony Corns [b], Susan Curran [b], Dragi Kocev [c,d], Žiga Kokalj [a,e,f]

[a] Research Centre of the Slovenian Academy of Sciences and Arts, Novi trg 2, SI-1000, Ljubljana, Slovenia
[b] Discovery Programme, Merchants House, 27-30 Merchants Quay, D08 K3KD, Dublin, Ireland
[c] Bias Variance Labs, Koprska ulica 72, SI-1000, Ljubljana, Slovenia
[d] Institut "Jožef Stefan", Jamova cesta 39, SI-1000, Ljubljana, Slovenia
[e] Postgraduate School ZRC SAZU, Novi trg 2, SI-1000, Ljubljana, Slovenia
[f] Department of Prehistoric and Historical Archaeology, University of Vienna, Franz-Klein-Gasse 1, A-1190, Wien, Austria

ARTICLE INFO

Dataset link: https://github.com/EarthObservation/adaf, https://doi.org/10.5281/zenodo.15848662

Keywords:
Lidar
Airborne laser scanning
Deep learning
CNN
Archaeological prospection
Automated detection

ABSTRACT

The need to use machine learning (ML) in archaeology is constantly increasing due to the growing availability of large amounts of high-quality airborne laser scanning data (ALS, lidar). The Automatic Detection of Archaeological Features (ADAF) tool was developed to provide an easily accessible deep learning (DL) software for automating the detection of archaeological features from ALS data. The software has a user interface that requires minimal interaction and no prior knowledge of ML techniques, greatly improving accessibility for the archaeological community. The underlying ML models (MLMs) have been trained on an extensive archive of ALS datasets in Ireland, which have been labelled by experts with three types of archaeological features, namely enclosures, ringforts and barrows. The core components of the tool are the Relief Visualisation Toolbox (RVT) for processing the ALS data into ML-friendly visualisations and the Artificial Intelligence Toolbox for Earth Observation (AiTLAS), which provides access to the DL models, both of which are actively used in the field of aerial archaeology. Although developed for detecting selected archaeological monuments in Ireland, ADAF supports retraining with data from other regions and other monument types, enabling its use in archaeological contexts worldwide. The practical application of the software was tested on an active infrastructure development corridor in Ireland. The results showed a promising 84 % recall rate for known archaeological sites and outperformed human detection by discovering 116 potential new archaeological sites that had not been noticed during manual inspection.

1. Introduction

Improving the detection of archaeological sites and remains with deep learning algorithms applied to remote sensing data represents the most advanced and effective approach to identifying potential new archaeological areas not included in the existing archaeological and cultural record (Kadhim and Abed, 2023). However, the application of deep learning techniques in archaeological prospection is still not widely adopted in the archaeological community. This highlights the need for rigorous evaluation and refinement of these advanced methods, as well as the development of accessible and user-friendly frameworks that can be seamlessly integrated into standard archaeological workflows.

The use of ALS (Airborne Laser Scanning) data for archaeological surveys has become indispensable in landscape archaeology since its introduction more than two decades ago (Challis and Howard, 2006; Challis et al., 2008; Crow et al., 2007; Harmon et al., 2006; Crutchley, 2009). ALS has allowed a rapid acquisition of extensive landscape data, which has significantly improved the identification and documentation of archaeological sites and monuments, particularly in advance of new infrastructure developments. Today, consulting ALS data is a common practice, but it is still mostly limited to manual documentation and analysis of archaeology at the site-level, leaving it largely underutilised for large-scale analyses (Verschoof-van der Vaart and Lambers, 2021; Bennett et al., 2014; Bevan, 2015). While ALS data allow for a broader and more comprehensive analysis of landscapes, large amounts of data can quickly outpace the human ability to manually analyse it. Furthermore, manual analysis is biased towards the expectations, experience, knowledge and observational skills of the interpreter(s) and carries the risk of overlooking or dismissing relevant objects (Cowley, 2012; Halliday, 2013; Cowley, 2016). To capitalise on this surge of

* Corresponding author.
E-mail address: nejc.coz@zrc-sazu.si (N. Čož).

https://doi.org/10.1016/j.jasrep.2026.105733
Received 6 November 2025; Received in revised form 27 March 2026; Accepted 28 March 2026
Available online 31 March 2026

data, a paradigm shift in conventional techniques for the analysis and interpretation of ALS data for archaeology is required. Key to this is a greater incorporation of computer-aided feature detection methods as an effective approach for analysing airborne and satellite data (Bennett et al., 2014).

This research develops a streamlined method for analysing ALS data using machine learning (ML), in particular convolutional neural networks (CNNs), for archaeological applications. These methods address the challenge of processing large volumes of remote sensing data, i.e. at a scale that makes manual analysis impractical, especially with limited budgets. By learning to generalise from large sets of labelled images rather than relying on manually defined parameters or classification rules, ML approaches can reduce both the time and cost of traditional processing methods while often providing more accurate results. Despite the abundance of remote sensing data from optical and radar satellites as well as airborne laser scanning, the availability of accurate archaeological reference data remains a significant bottleneck (Guyot et al., 2021). Such data is essential for the effective training and application of ML algorithms. However, the processes involved in collecting, publishing and maintaining labelled archaeological datasets are complex and existing inventories still have their limitations when it comes to providing large-scale, high-quality references.

In recent years, ML workflows for the automatic detection of archaeological objects from ALS data have been increasingly adopted in research projects. Various studies have explored different CNN architectures for classification (Somrak et al., 2020), object detection (Verschoof-van der Vaart et al., 2020; Trier et al., 2021; Orengo et al., 2020; Bonhage et al., 2021; Trier et al., 2018) and semantic segmentation approaches (Bundzel et al., 2020; Banasiak et al., 2022; Kocev et al., 2022; Wang et al., 2024). A notable limitation of these studies is that the CNN methods are usually provided as computer code, which is often not user-friendly and therefore not accessible to archaeologists with limited or no programming skills. Another drawback is that these workflows are often tailored to specific research projects, making it difficult to adapt them to new or different datasets. For computer code to be reusable, it needs to be easily accessible for download and accompanied by comprehensive documentation. This includes clear instructions on how to download, install and deploy the code, as well as access to the source code itself.

In the paper we present the design process and rationale behind the development of The Automatic Detection of Archaeological Features (ADAF) tool. The ADAF software was developed through a collaboration between the Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU), the Discovery Programme, the Bias Variance Labs and Transport Infrastructure Ireland (TII). Although the underlying ML models rely heavily on a high-quality training dataset, this dataset is not described or analysed in detail here. Rather, the creation and comprehensive documentation of this dataset will be the subject of a dedicated data paper that is currently in preparation. In this paper, we focus on the development of the ADAF tool itself and the rationale behind the design decisions. We make the entire ADAF workflow (source code) openly accessible in line with the principles of Open Science and FAIR (Findable, Accessible, Interoperable, Reusable) data management. By addressing both technical and user-friendly challenges in the application of ML to archaeological data, this work contributes to the standardisation of detection methodologies and lowers the barriers to the broader adoption of ML technologies in archaeological practice.

2. Background

Airborne laser scanning (ALS) has become a pivotal source of high-resolution topographic data in archaeology, enabling detailed documentation of landscapes and the identification of archaeological features with unprecedented precision (Bennett et al., 2025). However, the growing volume and complexity of ALS data demand scalable, consistent and efficient analytical methods (Verschoof-van der Vaart and Lambers, 2019). Deep learning, a subfield of machine learning, has emerged as a transformative tool for analysis of ALS data, with growing application in archaeological research (Fiorucci et al., 2020; Kadhim and Abed, 2023). Among the various ML techniques, convolutional neural networks (CNNs) have demonstrated particular effectiveness for spatial and visual analysis of ALS-derived data (Ball et al., 2017). Yet, despite their potential, CNNs are still not commonly used in archaeological practice (Kadhim and Abed, 2023).

To ensure wider adoption of these promising techniques, it is essential to address several key challenges that currently hinder their integration into archaeological practice. These include the preparation of machine learning-ready datasets, the technical complexity of configuring and training ML models and the need for accessible, user-friendly tools that can be easily learned and integrated into existing archaeological workflows. Overcoming these barriers requires the development of tailored software solutions designed specifically for archaeological applications.

Firstly, preparing input data suitable for ML workflows is a complex challenge (Kokalj and Somrak, 2019). It includes converting datasets such as ALS point clouds or digital elevation models (DEM) into formats that ML algorithms can process effectively. This requires pre-processing steps such as changing the range of pixel values, handling missing or unavailable data and ensuring consistent DEM quality across different datasets. In addition, archaeological datasets often lack comprehensive labelling, necessitating the creation of reference data or annotations for training of ML models. Although labour-intensive, this step is critical for supervised learning approaches.

Secondly, the development of suitable ML models involves navigating a range of technical tasks, from selecting and implementing appropriate CNN architectures to optimising hyperparameters, managing limited training data and mitigating overfitting (Ball et al., 2017). In most cases, end users do not have the programming or ML expertise to address these tasks independently. As a result, the effective application of ML in archaeology remains largely confined to collaborations with ML specialists or custom-made solutions developed within individual research projects.

Finally, there is currently no free, specialised software that supports end-to-end machine learning workflows tailored to archaeological research. Existing geographic information systems (GIS), such as ArcGIS Pro, offer built-in deep learning tools for geospatial analysis that can be integrated into archaeological workflows (Fylaktos et al., 2025). While ArcGIS Pro is user-friendly, it provides only a limited range of machine learning methods and still requires users to independently prepare data, select appropriate models and fine-tune hyperparameters. Moreover, being proprietary software with expensive licence costs can restrict accessibility for researchers and hinder the reproducibility of results, presenting challenges for academic and open science contexts.

ML improves traditional workflows by automating the manual (expensive and time-consuming) detection of archaeological features in remote sensing data. However, despite their transformative potential, ML results must be considered as approximations. As Cowley (2012) noted, archaeology needs to move away from a fixation on the correctness of individual identifications and begin to view overall patterns as descriptive. Accepting a higher rate of false positives can prevent the oversight of important archaeological features (Cowley et al., 2020). Canuto et al. (2018) found that human experts missed approximately fifteen per cent of objects in a study of ancient Maya settlements. Similarly, Verschoof-van der Vaart and Lambers (2021) reported that when reviewing sites based on ML results, additional archaeological features were discovered that had been missed by manual and automated analyses. These findings underscore the complementary nature of ML and human expertise, especially when it comes to addressing areas of uncertainty and leveraging ML results for improved manual analysis.

To address these challenges, the archaeological community needs to develop open access tools that are consistent with the principles of open

Fig. 1. Schematic representation of the ADAF workflow.

science and FAIR data management. These tools should (i) provide automated workflows for preparation of ML-ready data for archaeology, (ii) supply training datasets and pretrained ML models for detecting various archaeological features in landscapes from remote sensing data (e.g., ALS, satellite imagery, UAV-based surveys) and (iii) strive to provide user-friendly interfaces for non-experts. Furthermore, there is a pressing need for standardised methods and the creation of high-quality datasets. While ML methods are not a complete replacement for human expertise, they are becoming a powerful addition to the archaeologist's toolbox.

3. ADAF toolbox

3.1. Overall philosophy and design choices

The ADAF software has been developed with a particular focus on user accessibility. It offers not only an open source set of programming code and functions, but also a practical tool with a graphical user interface (GUI) (Čož et al., 2025a). This GUI captures all required input parameters and executes the chosen processing method, making ADAF a more intuitive alternative to many existing CNN-based ALS analysis workflows in archaeology. While studies such as Verschoof-van der Vaart et al. (2020), Verschoof-van der Vaart and Landauer (2021) offer similar workflows, their models usually require a familiarity with programming and a basic understanding of ML methods. Although their approach is effective for research purposes, such software often fails to integrate into practical workflows outside of academic environments. In contrast, the user-friendly design of ADAF aims to bridge this gap and make a tool based on CNNs accessible for wider application in archaeology.

Automatic detection of land features in ALS data using ML is a complex task that requires the development of a robust workflow. This workflow must be both user-friendly and capable of delivering accurate and reliable results. To create a state of the art product, many different software libraries, processing methods and quality controls had to be integrated into a single program, with a carefully designed user interface that unifies these components into a coherent and seamless processing tool.

ADAF was developed in Python 3 and relies on two Python libraries: the Relief Visualisation Toolbox Python library (RVT_py) (Kokalj and Somrak, 2019; Kokalj et al., 2024) and the Artificial Intelligence Toolbox for Earth Observation (AiTLAS) (Dimitrovski et al., 2023). The tool follows a workflow similar to those commonly used in landscape archaeology, such as Verschoof-van der Vaart et al. (2020), Bonhage et al. (2021), Maxwell et al. (2020), Altaweel et al. (2022), Küçükdemirci et al. (2023), Soroush et al. (2020) and progresses from the input of ALS data to the generation of detection polygons (Fig. 1). The first step involves pre-processing the ALS data to ensure compatibility with machine learning models (MLM). In the second step, inference of the selected MLM creates probability maps that highlight potential archaeological objects. In the final step, post-processing vectorises the results and minimises false-positive detections by removing objects that do not meet the user-defined constraints (see Section 3.4 for a more detailed description).

The main purpose of ADAF is to enable finding new archaeological sites with object detection and semantic segmentation MLMs. Although the software offers the possibility to retrain these models, retraining was intentionally excluded from the graphical user interface (GUI) to keep it simple. The ADAF GitHub repository provides dedicated Jupyter notebooks that facilitate the retraining process.

Two different approaches to retraining MLMs are supported. The first is to update an existing MLM, i.e. to improve an already trained ADAF model by incorporating additional training data and thus increasing its predictive performance. The second approach is to train a completely new MLM. Although this method does not use existing ADAF models, the training process does not start entirely from scratch either. Instead, the pretrained weights, such as ImageNet (Simon et al., 2016), available in the AiTLAS toolbox are used to accelerate and optimise the initial stages of training.

Detailed procedures and instructions for the retraining process fall beyond the scope of this article and are documented in detail in the software manual (Čož et al., 2025b). It is important to emphasise that the configuration of training the models for new areas requires both a considerable amount of expertise in ML and access to high quality, accurately labelled training data (e.g. archaeological DEMs with annotated archaeological sites), both of which can be difficult to obtain. In addition, the computational demands of training MLMs require computers equipped with high-end GPUs.

ADAF is designed to run efficiently on standard Windows 64-bit systems and is therefore accessible to the majority of users. It operates optimally on machines running Windows 10 or 11 and requires the Anaconda platform with Python 3.9 distribution. A minimum of 16 GB RAM and an NVMe SSD drive are recommended for smooth data handling and faster read/write operations. Although ADAF can run without a dedicated GPUs, an NVIDIA CUDA-enabled GPU is preferred for executing large-scale analyses. For example, using a standard laptop without GPU, the tool performs inference over one square kilometre of data at 0.5 m resolution in under one minute.

3.2. Inputs and outputs

The software works with any type of a DEM as input, but for best results the user is advised to prepare the ALS data as an archaeology-specific DEM, the so-called Digital Feature Model (DFM) (Pingel et al., 2015). It combines terrain data with archaeologically relevant micro-relief features such as standing walls and stones, roads, canals, earthworks and includes modern buildings for contextual interpretation (Štular et al., 2021). As DFMs are created in a specialised, computationally intensive way, a good alternative is to create a DEM by rasterising point clouds with existing classification information for terrain and buildings.

The code has been prepared to handle small and large DEMs and allows the user to select any file in a Tagged Image File Format (TIFF) or GDAL Virtual Format (VRT) - standard file formats for rasterised elevation data. The best results are achieved with input files that have a similar raster quality to the datasets used in training, i.e. DFMs with a spatial resolution of 0.5 m when using the default ADAF MLMs. If more than one file is selected, the software assumes that each file is a separate dataset and performs the inference for each file individually. If the files are in fact tiles of the same dataset, the user has the option to activate the on-the-fly creation of a virtual mosaic (VRT). The final results generated by ADAF are polygons representing the detected archaeological objects, stored in a GeoPackage format (an open, non-proprietary, platform-independent and standards-based vector file) compatible with most geographic information systems (GIS). When multiple classes of archaeological objects (e.g. any combination of barrows, enclosures, ringforts and all archaeology) are selected for processing, the resulting

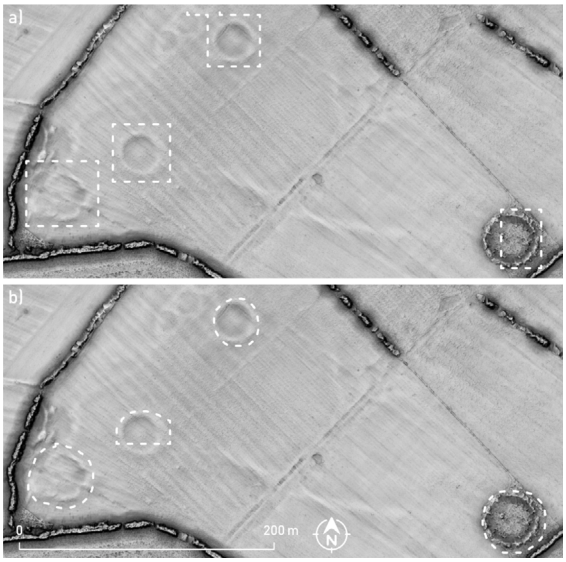

Fig. 2. Comparison of results from an object detection model (top) and a semantic segmentation model (bottom) applied to ALS-derived data for archaeological feature detection. The background image is a Combined Visualisation for Archaeological Topography (VAT) generated using the rvt_py Python library.

polygons are consolidated into a single vector file. Within this file, a label attribute is included to differentiate between the different archaeological classes, so that the user can easily distinguish between the different object types within the same dataset. The output has the same format regardless of the ML method selected. The only difference is that object detection creates rectangular bounding boxes around the detected objects, while semantic segmentation provides detailed outlines of the objects (Fig. 2).

3.3. ML-ready ALS data for Irish archaeology

The ML-ready ALS data for Irish archaeology (hereafter referred to as ML-ALS-IE dataset) was a cornerstone of the Automatic Detection of Archaeological Features Machine Learning Project, the aim of which was to facilitate the development of MLMs for automated archaeological detection. This dataset leverages an extensive archive of point cloud data provided by Transport Infrastructure Ireland (TII) and combines it with comprehensive archaeological reference data to create a robust resource for ML training. The training dataset consists of more than 23,000 samples in total. These samples, known as patches, consist of an image tile with labelled data, including segmentation masks and bounding boxes for object detection.

The TII point cloud dataset consists of over 200 ALS and photogrammetry-derived DFMs, ranging in file size from 100 MB to 50 GB and covering different parts of Ireland. Fig. 3 shows the geographical distribution of acquisitions across Ireland and illustrates the spatial extent of the dataset. The individual datasets varied considerably in terms of quality and spatial resolution, so a rigorous standardisation process was required to ensure consistency for ML applications. The differences in acquisition methods, point density, and data processing approaches meant that the point cloud data had to be carefully reprocessed and rasterised. To achieve this, datasets were homogenised and rasterised to a standard resolution of 0.5 m to ensure consistency in all subsequent analyses. The DFMs were created using the Triangular Irregular Network (TIN) method, which is particularly suited to the gently sloping landscapes of Ireland. This approach ensures compatibility between existing and potentially newly acquired ALS datasets and provides a uniform structure that facilitates ML training and analysis.

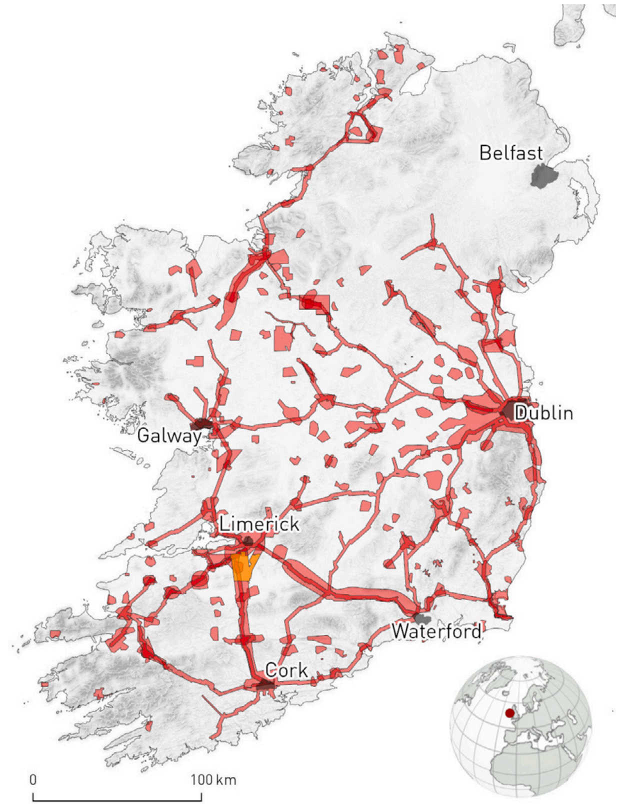

Fig. 3. Geographical extent of ALS datasets provided by TII (red). The test area for the practical application is just south of Limerick (yellow).

The digitisation of the relevant archaeological feature types (barrows, enclosures and ringforts) was anchored in the known monuments from the Irish Sites and Monuments Record (SMR) — a database curated by the Archaeological Survey of Ireland (ASI) that contains details of more than 138,800 archaeological monuments in Ireland and is regularly updated to reflect new discoveries and revisions (DCHG, 2023). The three monument types were selected based on a set of defined criteria (Fig. 4) which took into account their clearly defined morphological characteristics, their frequency in the Irish landscape and their suitability for ALS-based detection. Enclosures and ringforts are among the most numerous monument types in Ireland, providing a large enough dataset for ML training. Barrows, which are less common, were augmented by including barrows from the UK to create a more representative dataset (Historic England, 2025). The digitisation process focused on the visible archaeological features evident in the DFMs. Polygons were carefully delineated to capture the full extent of each monument. Several attributes for each polygon were also captured including: the classification of the quality of the underlying DFM data, if the monument features displayed evidence of truncation and if the monuments are multivallate in form. These attributes would enable the filtering during ML training. This resulted in a total of 10,718 segmentation polygons.

Prior to tiling, the DFMs were converted to an ML-compatible image format to improve feature visibility and standardise the data for training. The dataset was then divided into patches by placing a uniform grid over the DFMs, with grid offsets applied to ensure a diverse set of training examples. This produced samples with both full and partial monuments, serving as a form of data augmentation, providing different spatial contexts and fragmentations of the same feature types. The digitised polygons were used to generate labelled data for each patch. The tile size was set at 512 pixels (256 m × 256 m) to achieve an optimal balance between archaeological feature representation and

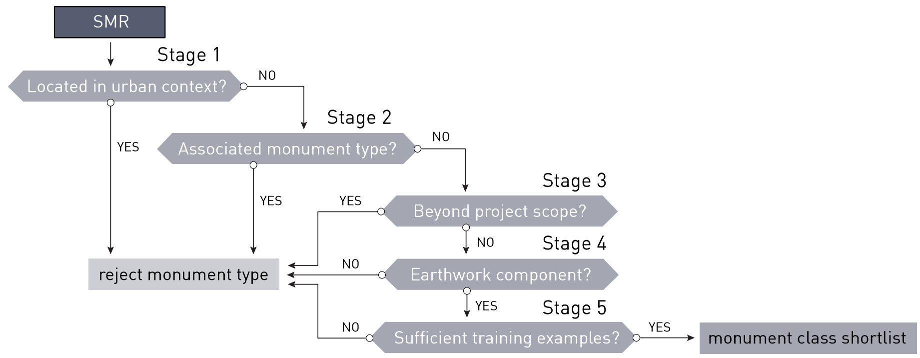

Fig. 4. Decision tree outlining the criteria used to select monument types suitable for machine learning-based detection from ALS data. The associated monument type criterion refers to cases where multiple discrete records of the same monument type are present, e.g. 'Cross' and 'Church'.

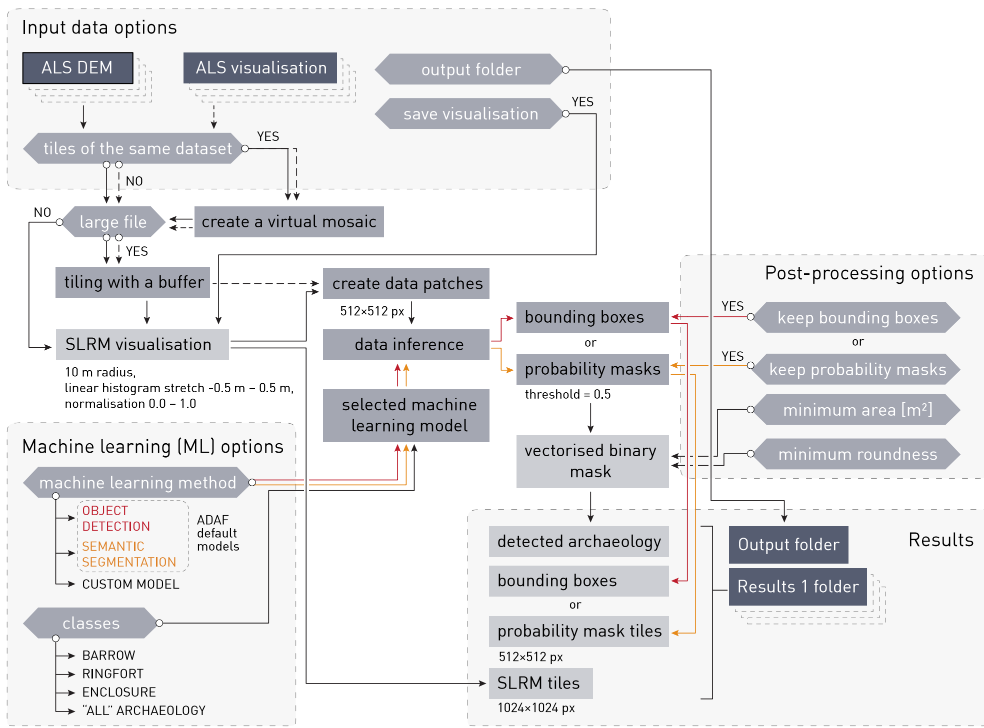

Fig. 5. Overview of the ADAF processing workflow, showing the three main stages: pre-processing (visualisation and tiling of ALS data), machine learning inference (detection of archaeological features) and post-processing (vectorisation and filtering of results).

computational efficiency. This tile size allowed for complete capture of the selected archaeological monument types while retaining enough contextual information for ML training.

3.4. User-friendly machine learning process

In the context of ADAF software, the processing workflow (Fig. 5) is divided into three main specialised steps, each with several important sub-steps:

Step 1. Pre-processing, which includes computing visualisations from DEMs and creating ML-ready data patches for inference (i.e. detection).
Step 2. Inference, which utilises the developed MLMs for detection of archaeological features.
Step 3. Post-processing, which includes methods to improve detection rates and remove obvious false positive detections.

3.4.1. Pre-processing

The task of the pre-processing module is to convert the input data into the correct format for the ML module. This includes calculating the required visualisation from the DEM, adapting the pixel values to the expected range for ML (e.g. normalisation or linear stretching) and dividing the data set into ML-ready tiles.

Although the primary type of input data are DEMs, the MLMs were trained on a DEM derivative called Simple Local Relief Model (SLRM), a commonly used type of relief visualisation in landscape archaeology (Fig. 6). The SLRM removes large-scale morphological elements (e.g. hills, valleys) from the data, leaving only small-scale features (Hesse, 2010). Other visualisations were also tested during the development of ADAF. However, the SLRM provided consistently robust results across different MLMs and had the added benefit of requiring minimal computational effort, thus reducing both processing time and computer memory usage.

The SLRM images are generated with RVT_py, running the SLRM visualisation function (Kokalj et al., 2024). The radius for trend assessment is set to 10 m, a value that is effective at isolating the small-scale

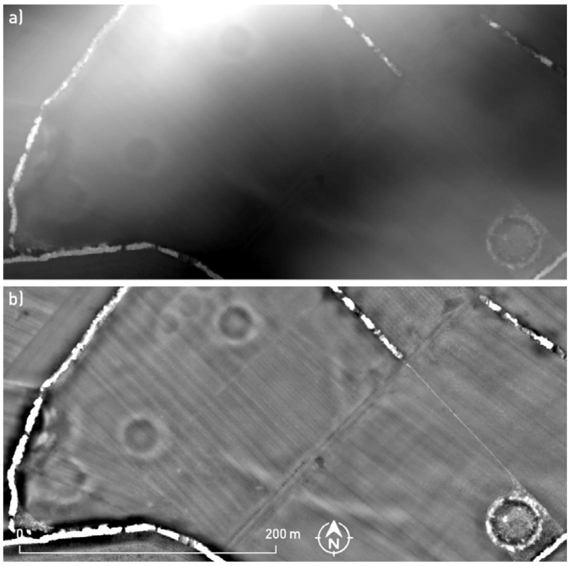

Fig. 6. Comparison between DEM with local dynamic range adjustments and 1 percent min/max clipping (top) and SLRM visualisation for ADAF (bottom).

Table 1
Parameters used in the pre-processing step for generating ML-ready Simple Local Relief Model (SLRM) visualisation. These include settings for trend assessment, value clipping, histogram stretching and handling of missing data.

Radius for trend assessment (m)	10
Linear histogram stretch (m)	[−0.5, 0.5]
Normalisation	[0, 1]
Value of "nodata" and "nan" pixels	0

morphological signals relevant for detection of archaeological features in the ML-ALS-IE dataset. A linear histogram stretch and normalisation of the SLRM are then applied, and nodata values are set to zero, as required by the ML algorithm. The pre-processing methods and their parameter values are listed in Table 1.

The dataset is automatically divided into smaller tiles to optimise processing efficiency. The tiling routine calculates the maximum buffer required for seamless results and trims the raster back to the original size after processing. The data is then divided into square tiles of 512 pixels. Although ML processing can support any tile size based on powers of two (e.g. 64, 128, 256, 512, 1024), our analysis has shown that a tile size of 512 pixels provides the best balance between memory usage, processing speed and detection performance.

3.4.2. Machine learning models

The core functionality of the ADAF software relies on trained MLMs developed through a rigorous evaluation and refinement process using the AiTLAS framework (Dimitrovski et al., 2023). The aim was to identify the best performing models for the automatic detection of archaeological features from ALS-derived data.

The development process involved exploring different visualisations for training samples, filtering samples based on their quality, applying data augmentation techniques and adjusting image sizes. A comprehensive analysis of different Deep Neural Network (DNN) architectures was conducted. Several state-of-the-art deep learning models for semantic segmentation and object detection were evaluated under identical training conditions. As a result of testing, the HRNet architecture (Wang et al., 2020) was selected for semantic segmentation and Faster R-CNN (Ren et al., 2015) was chosen for object detection. For each architecture, a dedicated model was trained for each of the four detection classes (barrows, enclosures, ringforts and all combined), resulting in a total of eight individual machine learning models.

Table 2
Percentage of samples in the validation and test subsets.

	Train	Validation	Test
Barrows	65	18	17
Enclosures	63	17	20
Ringforts	64	16	20

Table 3
Mean Average Precision (mAP) performance metric for the default ADAF model for object detection (i.e., Faster R-CNN) and Intersection-over-Union (IoU) performance metric of the default ADAF model for semantic segmentation (i.e., HRNet).

	Faster R-CNN	HRNet
Enclosures	0.325	0.282
Ringforts	0.715	0.575
Barrows	0.594	0.504
"All archaeology"	0.613	0.584

A set of 23,100 labelled image patches was used in training of MLMs, split approximately 60/20/20 (actually 64/17/19) into training, validation and test datasets (see Table 2). The division of the data was based on the geographical location of the samples to ensure spatial separation (Simidjievski et al., 2021) and maximise the generalisation ability of the MLMs (Bundzel et al., 2020). Each image patch consisted of a 512 × 512 px tile derived from SLRM visualisations. The training data was augmented with random flips and rotations to improve robustness. All models were trained for 50 epochs with a learning rate of 1×10^{-4}, using ImageNet pretrained weights (Simon et al., 2016) provided by AiTLAS. This transfer learning approach significantly accelerated training and improved convergence (Kocev et al., 2022).

Table 3 summarises the performance of the four final models for semantic segmentation and object detection. Semantic segmentation performance is reported using Intersection over Union (IoU), a standard ML metric that measures the area of overlap between the predicted and reference regions relative to their combined area, with values closer to 1 indicating better agreement. Object detection performance is reported using mean Average Precision (mAP), a standard metric based on the precision–recall relationship, with higher values indicating better detection performance. Although the standard implementation of IoU does not capture all aspects of model performance relevant to archaeological prospection (Fiorucci et al., 2022), it was used here as a comparable evaluation measure, consistent with previous archaeological deep-learning studies based on ALS data (Bundzel et al., 2020; Banasiak et al., 2022).

In summary, the MLMs that drive ADAF are the result of a carefully constructed and annotated Irish ALS dataset, a methodical evaluation of ML architectures using AiTLAS, extensive testing under controlled conditions and rigorous model selection. These models now form the basis for ADAF's ability to automatically identify and delineate archaeological objects in large ALS datasets, supporting both semantic segmentation and object detection methods.

3.4.3. Post-processing

The primary task of the post-processing module is to convert the results into an easily useable output format. This includes the vectorisation of bounding boxes as a result of the object detection method or the vectorisation of probability masks created with the semantic segmentation method. All polygons are collected into one vector layer and their respective detected class is listed in the attribute table.

When using methods for automatic detection of archaeological features it is better to have more false positives than to miss potential discoveries through false negatives (Verschoof-van der Vaart et al., 2020). With this in mind, we have found that there are always a number of obvious false positives that can be removed automatically without the risk of loosing potential new detections. Their characteristics vary from dataset to dataset, so the parameters for post-processing should be adjusted on a case-by-case basis.

We have included two parameters for removing known false positive examples from the vector layer — the area size of the polygon and its geometric shape. Archaeological features of the selected Irish monument types tend to be circular and no smaller than $40\,\mathrm{m}^2$, which guided us in the selection of post-processing parameters and their respective default values as described below. The filtering can always be done manually within a GIS, but including it in the automated workflow reduces steps in the manual analysis. Both parameters are written in the attribute table of the output vector layer.

Filtering by area

To filter out small segmentation polygons, the user can set the minimum area slider to remove detections that are smaller than the set threshold. The threshold should be increased if the results show an excess of small polygons that do not correspond to archaeological features. Conversely, it can be lowered or set to zero if smaller, clearly identifiable archaeological features are detected, but removed in post-processing.

The default setting is $40\,\mathrm{m}^2$ and the slider can be moved in increments of $5\,\mathrm{m}^2$ between the values of $0\,\mathrm{m}^2$ and $100\,\mathrm{m}^2$. The default value was determined from the geometric statistics of the polygons used to train the MLMs. Fig. 7a shows the number of polygons smaller than $500\,\mathrm{m}^2$ from the ML-ALS-IE dataset for different archaeological objects — there are many more larger objects that are not relevant for setting the minimum area threshold and are therefore not included in the figure. The only two objects smaller than $30\,\mathrm{m}^2$ were two barrows from the United Kingdom. The known barrows in Ireland are all larger than $50\,\mathrm{m}^2$, while enclosures and ringforts are not smaller than $90\,\mathrm{m}^2$ and $120\,\mathrm{m}^2$ respectively, so the user can be less conservative in selecting the threshold when detecting one of the latter.

Filtering by roundness

The ADAF has been trained to detect archaeological features that typically have a circle-shaped footprint. Removing elongated shapes from the detected features can help to reduce the number of false positives, particularly caused by the detection of hedges and modern field boundaries. As a threshold, we introduced the minimum roundness index of a polygon. The roundness parameter is defined following the formulation of parameter K by Cox (1927). It is calculated as the area of the polygon divided by the square of the perimeter of its convex hull and normalised by multiplying the result with 4π (Eq. (1)). Rounder objects have a roundness value closer to 1, while longer, thin objects have a value closer to 0.

$$roundness = \frac{4\pi * area}{(convex\ perimeter)^2} \tag{1}$$

The default value of 0.5 was derived from the geometric properties of the polygons used to train the MLMs (Fig. 7b). Only two ringforts and five enclosures, fewer than 0.001 % of the polygons, have roundness values just below 0.5 (Table 4). All of these exhibit deviations from a circular footprint due to the presence of modern structures, such as roads and buildings that cut through or have been constructed over the original extent of the monuments. A less stringent threshold may be appropriate for barrows, as the majority of them exhibit roundness values closer to 1 (Fig. 7b), reflecting their typically well-preserved circular morphology.

It is important to note that this parameter is only applicable when post-processing the semantic segmentation results, as it is not relevant for object detection, which inherently generates rectangular bounding boxes.

Table 4
Summary statistics of object roundness based on the polygons of archaeological features from ML-ALS-IE dataset, showing minimum, median and standard deviation values for each monument class, along with the number of objects falling below the post-processing threshold of 0.5.

	Roundness			Instances	
	min.	median	std. dev.	total	under threshold
Barrows	0.663	0.975	0.03	522	0
Barrows (UK)	0.513	0.999	0.04	1939	0
Ringforts	0.494	0.969	0.05	5933	2
Enclosures	0.490	0.966	0.07	2324	5

3.5. Graphical user interface

Creating MLMs is still a task for computer science experts, but as ML methods become more widely used, it is important to make them accessible to non-experts. The first step in lowering the entry barrier is to provide an intuitive user interface that simplifies the workflow and allows users to focus solely on setting inputs and outputs without the need for programming or technical expertise.

The design process of the ADAF GUI followed an iterative approach that required close collaboration between designers, remote sensing specialists, ML experts and archaeologists. User testing was conducted through a combination of (i) online software demonstration involving TII archaeological consultants and external archaeologists, (ii) dedicated workshop with archaeological practitioners and (iii) individual consultations. Feedback gathered during these activities directly informed the refinement of the GUI. As a result, the interface was streamlined to restrict user interaction to essential parameters only: selection of input files, definition of output location and choice of ML method. All other processing parameters were predefined by the development team based on optimal results derived from a systematic testing of the ML-ALS-IE dataset.

The GUI uses the Jupyter Notebook Widgets for a clear and intuitive interface. All parameters are selected by interacting with buttons, dialogue boxes, selection boxes and sliders (Fig. 8). It is designed to provide an intuitive workflow through its three main components: Input data options, Machine learning options and Post-processing options.

The first section of the GUI, **Input Data Options**, facilitates the selection of input files and provides pre-processing capabilities. The user can select one or more files via a standard file dialogue, which is opened via the *Select file* button. Each selected file is treated as an individual dataset and processed independently. For datasets consisting of multiple tiles (e.g. tiled rasters), the GUI supports the virtual mosaic, where the tiles are treated as a single dataset. By default, the output files are saved in a subfolder of the same directory as the input files, but users can specify a different destination using the *Change output folder* button.

Pre-processing prepares the data for ML. An unprocessed DEM file is expected as a default input, which is then automatically preprocessed to generate visualisations and adjust pixel values for ML interpretability. There is an option for users to skip this preprocessing step in case they have already prepared the appropriate visualisations themselves. To enable this, a radio button switches between the *DEM* and *Visualisation* input modes. Preprocessed files can be created externally using tools such as the RVT plugin for QGIS, by applying the parameter settings listed in Table 1. In addition, users can select the *Save visualisations* checkbox to save preprocessed files for reuse in subsequent runs, streamlining repeated analysis of the same dataset.

The second section, **Machine learning options**, provides control over the selection of MLMs. The user can choose between two ML approaches - *semantic segmentation* and *object detection*. A drop-down menu allows the selection of pretrained models and currently includes the default *ADAF model* trained on ML-ALS-IE dataset or a *Custom model*

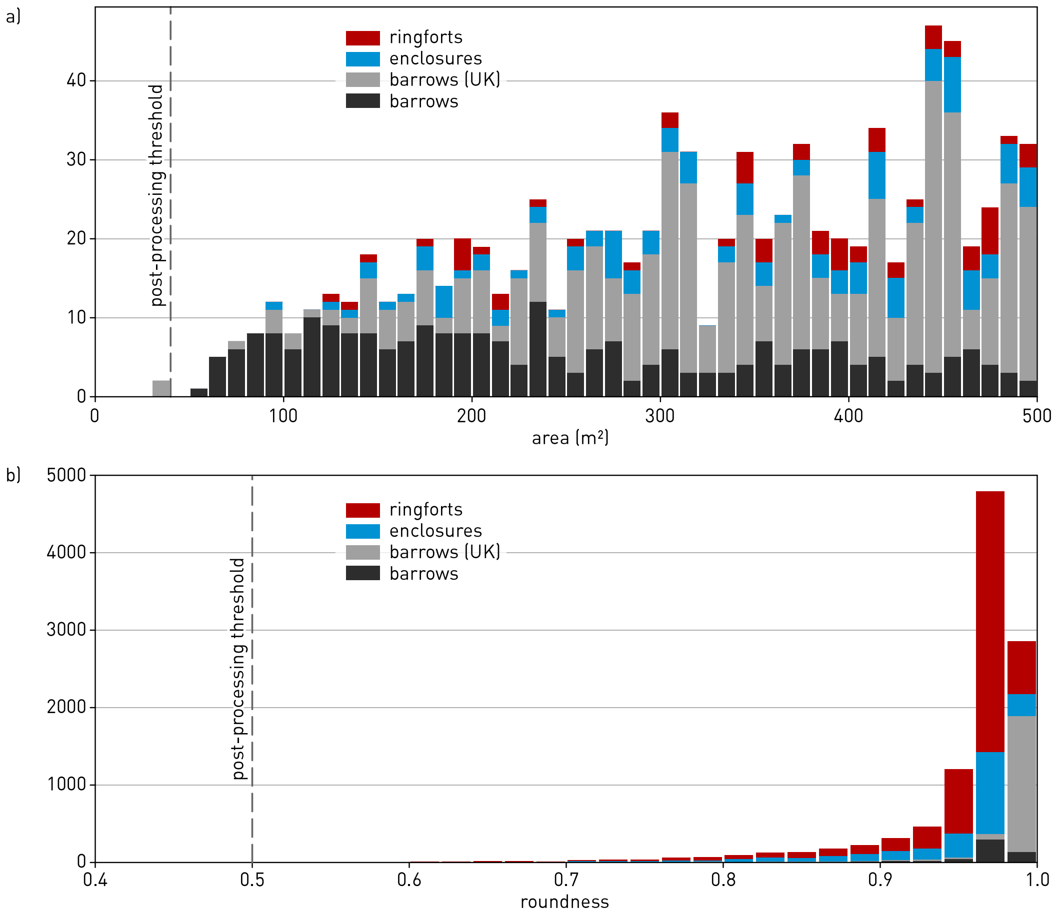

Fig. 7. (a) Distribution of archaeological objects in the ML dataset by area (m^2), limited to objects smaller than $500\,m^2$ and (b) distribution by roundness. The dashed line marks the default thresholds used in post-processing to exclude (a) small and (b) elongated detections.

that the user has either provided or trained independently. The custom models have to adhere to ADAF model specifications and require expertise in MLM training. Detailed instructions for creating custom models can be found in the software manual (Čož et al., 2025b).

When using the *ADAF model*, users can specify archaeological object classes for detection via the *Select classes for inference* options. Available classes include *barrows*, *ringforts*, *enclosures* and an aggregated *All archaeology* category. Any combination of these classes can be selected. Detected objects will be labelled with the corresponding class attributes in the output vector file.

The last section, **Post-processing options**, provides tools to refine the detection results. To exclude insignificant detections, the user can adjust the thresholds for the size and shape of the polygons using sliders. They can also save intermediate results such as probability masks. These are raw raster outputs of the MLM that indicate the probability of detection per pixel and are useful for evaluating the performance of the model or for customised post-processing.

4. Practical application — comparing human and machine performance

To evaluate the performance of the ADAF tool in a real-world setting, we conducted a case study using the ALS dataset of the TII N/M20 motorway project, an active infrastructure development corridor between Cork and Limerick in Ireland (see Fig. 3). The dataset covers a total area of $197\,km^2$ and provided a robust testing environment as it contains various archaeological contexts as well as a baseline inventory of 621 known archaeological sites recorded in the Sites and Monuments Record (SMR) database. Of these, 253 are categorised as likely (or at least possible) to be found in the ALS data. The remainder are sites that have been destroyed by recent construction, upstanding monuments or sites with no recognisable surface expression. The study employed the ADAF semantic segmentation model, using the broader "All Archaeology' detection class to maximise feature coverage. The post-processing parameters were configured with a minimum area threshold of $30\,m^2$ (very small) and a roundness value of 0.4 to help reduce the very obvious false positives while preserving archaeologically interesting features.

The analysis aimed to answer two key research questions. First, how effectively does the model identify already documented archaeological sites? To answer this question, we compared the ADAF results with a subset of the aforementioned 253 SMR records. Second, we assessed the model's potential to discover previously unrecorded sites by comparing its results with those of expert visual inspection of the ALS data. Human inspection identified 343 features of potential archaeological interest, 217 of which received a higher confidence score indicating a high probability of archaeological origin.

In total, the ADAF tool detected 772 potential archaeological features, which are summarised in Table 5. An initial review of the detection results shows that 213 match known SMR sites. Considering that the dataset contains 253 known SMR sites, this yields a recall value of 0.84. Here, SMR sites were treated as a single group and were not further subdivided by preservation state, management status, or degree of disturbance. As both the training and test datasets include sites with varying levels of visibility in ALS visualisations, the achieved recall is considered encouraging. The percentage of detected known sites is high, considering ADAF models were trained on 3 out of 18 monument type categories.

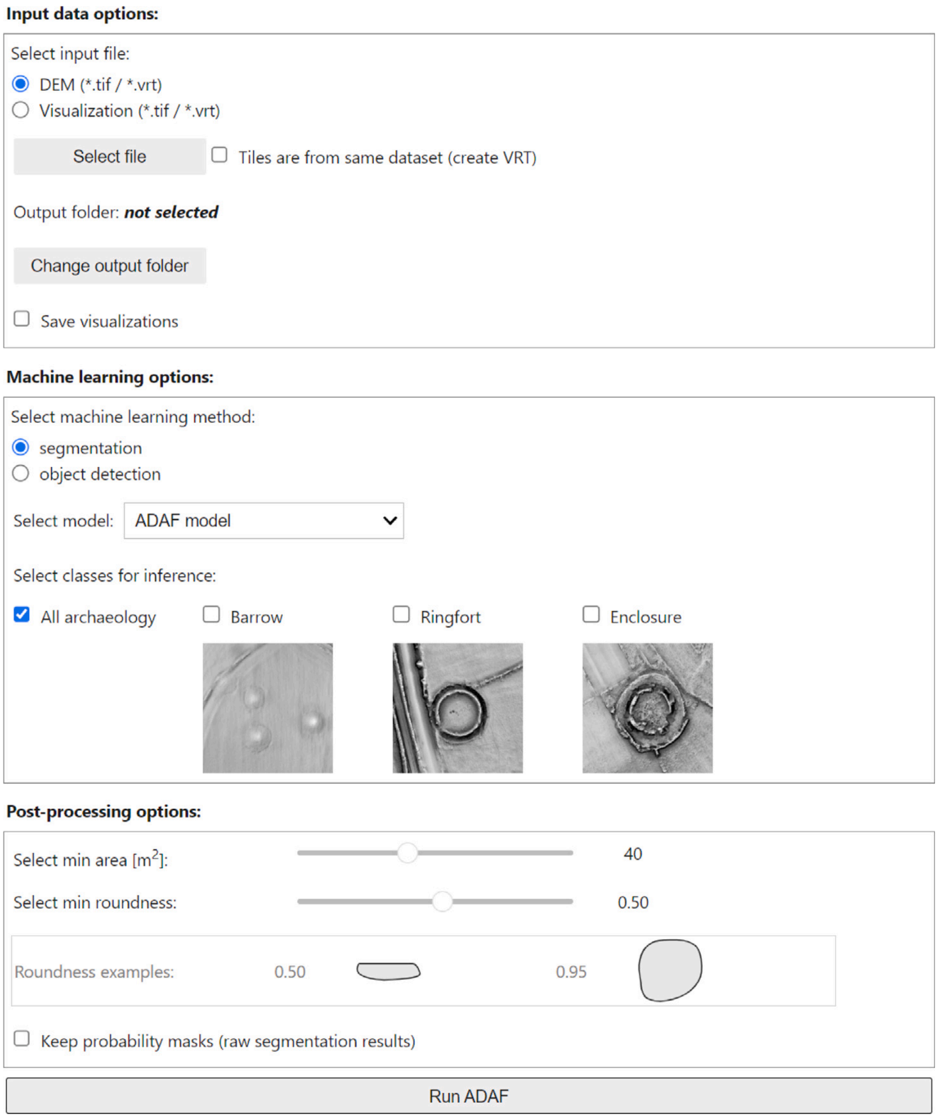

Fig. 8. The ADAF GUI is designed to provide an intuitive workflow through its three main components: Input data options, Machine learning options and Post-processing options.

Table 5
Performance comparison between the ADAF tool and human experts in detecting archaeological features from ALS data. The results summarise detection accuracy, recall and identification of new sites within the test area.

Category	Human inspection	ADAF
Total detections	343	772
Sites in SMR	–	213
Total new features of potential archaeological interest	343	429
Potential unknown sites (high confidence)	*217*	*197*
Potential unknown sites (low confidence)	*126*	*362*
Detected by both methods	81	81
Detected by a single method	262	348

This means that the tool identified 429 potential sites. A comprehensive manual review revealed that 362 of those were false positives, consisting primarily of modern or natural landscape features such as palaeochannels, golf courses, quarries, modern agricultural embankments and urban (embanked) infrastructure (Fig. 9). Similar types of false positives have also been reported in previous research (Verschoof-van der Vaart et al., 2020). Among the remaining detections, 81 features were detected by both the ADAF tool and human interpretation. However, the tool identified a total of 116 potential sites that had been missed during the human assessment (Fig. 10).

The tool's detection rate of 84 % for known SMR entries is a promising result for a fully automated method. However, the direct comparison with human inspection shows the limitations of both approaches. Of the features identified by human inspection, only 24 % were also detected by the ADAF tool, indicating a relatively modest overlap between manual and automated interpretations. Conversely, only 18 % of the features detected by ADAF were also recognised by human analysts, meaning that 82 % of the machine-detected features were not identified in the manual assessment. This asymmetry suggests that ADAF is capable of detecting a considerable number of features that may be missed by human interpretation. A likely explanation for this is that human analysts are able to identify a wider and more diverse range of monument types, including those not represented in the ML-ALS-IE dataset (e.g. field systems, upstanding structures). At the same time, the high number of purely machine detections indicates the sensitivity of the tool in identifying subtle surface anomalies that are easily overlooked by the human eye, especially in large and complex landscapes (see Fig. 10). These results highlight the complementary strengths of human and machine approaches and illustrate the potential benefits of

Fig. 9. Examples of false positive detections. (a) a round pen for training horses, (b) a cul-de-sac, (c) a cemetery, (d) a part of a golf course, (e) a moated castle hill, (f) circular tractor tracks, (g) a quarry stockpile, (h) natural deposits. National High Resolution Imagery shows the same area as the lidar data. CYAL50478200 1' Tailte Éireann — Surveying.

extending the training data to a wider variety of archaeological site types, which is consistent with previous studies that also stressed the value of combining automated detection with human interpretation and the importance of broader training datasets (Verschoof-van der Vaart and Lambers, 2021; Verschoof-van der Vaart, 2022).

This comparative analysis highlights the complementary strengths of human expertise and machine inference. While ADAF demonstrated strong performance in detecting known site types included in the training data – achieving an 84% match rate against the SMR dataset – its identification of previously undocumented sites underscores its utility as a discovery tool in archaeological prospection. At the same time, the divergence in results between human and machine detection illustrates the importance of integrating both approaches to achieve comprehensive landscape interpretation.

5. Discussion

The performance of ADAF in the TII N/M20 motorway project case study shows that the tool does not replace expert interpretation, but is a complementary method. While ADAF identified a large number of probable archaeological features, a considerable number of which were not detected during the initial human assessment, human expertise remains essential for the interpretation of ambiguous features, especially those not included in the training dataset. In this sense, the strength of ADAF lies in its ability to quickly and consistently highlight areas of potential interest in large datasets, reducing the cognitive and logistical burden on human analysts.

The tool also demonstrated a high level of expertise in accurately delineating the location and extent of features. In several cases, ADAF provided spatially more precise outlines of already recorded monuments compared to existing SMR records, which are usually represented by centroids or approximate boundaries. This higher spatial precision increases the usefulness of the results for follow-up investigations and decision-making in development projects.

The detection of 362 false positives is considerable, but not necessarily a limitation when view in light of the wider context of archaeological prospection. Many of these false positives were associated with ambiguous landforms or modern anthropogenic features that have similar morphological characteristics to archaeological structures. As discussed by Cowley et al. (2020), some degree of over-detection in

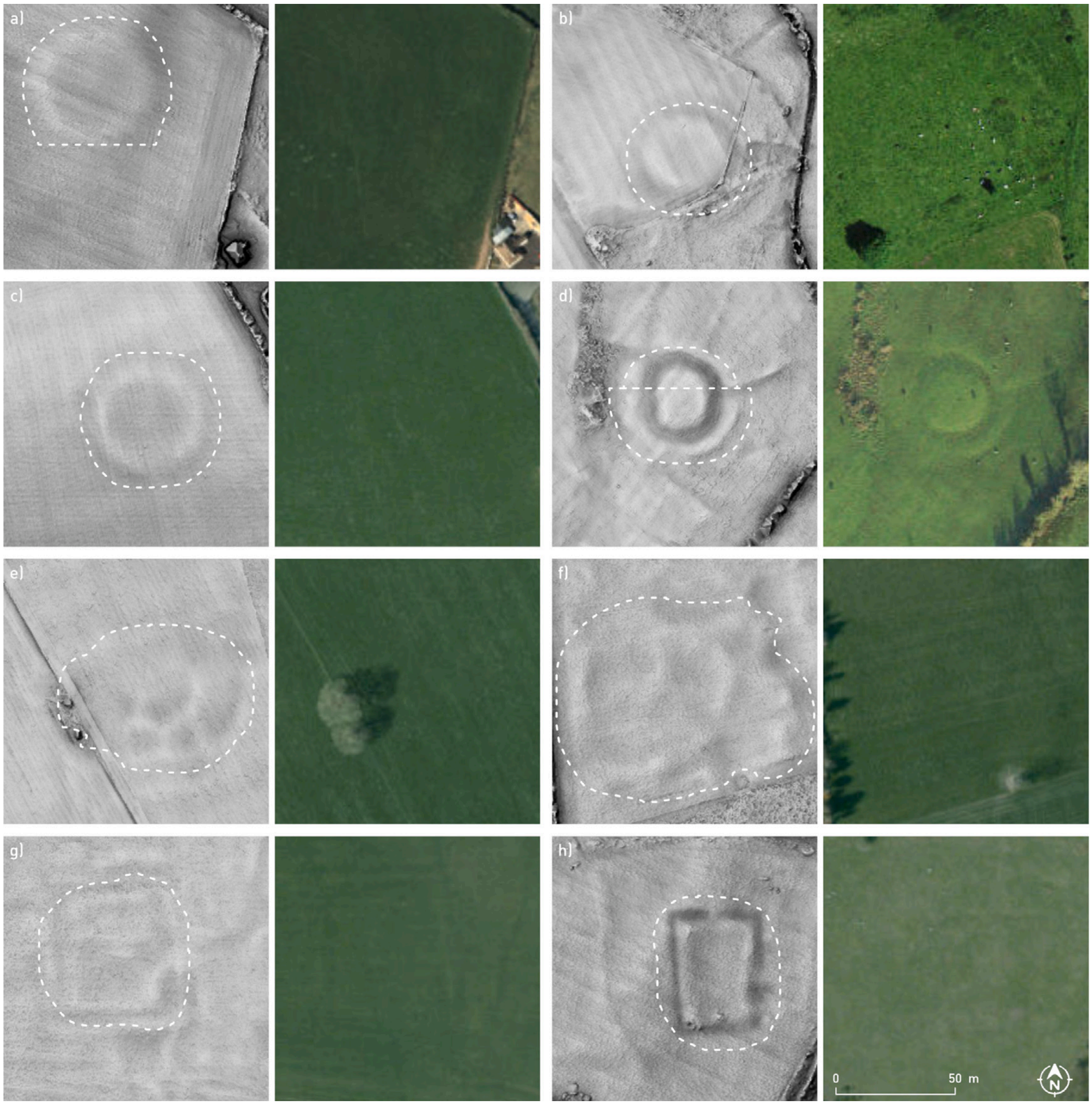

Fig. 10. Examples of detected features that potentially represent unrecorded archaeological sites. National High Resolution Imagery shows the same area as the lidar data. CYAL50478200 1' Tailte Éireann – Surveying.

initial screenings is not only acceptable but desirable. The aim is not only to achieve a perfect classification rate, but also to ensure that the possibility of overlooking genuine archaeological features is minimised.

In practical terms, the outputs of ADAF can be treated similarly to predictive archaeological maps, i.e. as guides for further verification and not as final classifications. In this capacity, the high recall rate of ADAF becomes a valuable asset when it comes to flagging potential features in advance of fieldwork, especially in contexts such as infrastructure planning where time and resources are limited.

An important consideration is the extent to which the performance of ADAF on the TII's Irish ALS dataset can be generalised to other projects. The model was trained on a high quality and well-labelled training dataset focusing on three common Irish monument types (barrows, ringforts and enclosures). The model's detection performance is therefore strongest when applied to areas and sites that reflect the characteristics of the training data. In regions with other monument typologies or landscapes with a different morphology, re-training or adaptation of the model is required.

This implies that while ADAF provides an out-of-the-box solution for certain Irish contexts, wider applicability requires the availability of labelled reference data and technical expertise for model refinement.

6. Conclusions

This study presents ADAF, a purpose-built open-access software tool designed to facilitate the automatic detection of archaeological objects from ALS data using state-of-the-art deep learning techniques. By integrating semantic segmentation and object detection capabilities into a user-friendly, GUI-based platform, ADAF offers a notable advance for archaeological prospection.

As demonstrated in the practical application case study, ADAF works well for the intended purposes. The results of the detection on the N/M20 dataset showed a high recall rate of known archaeological sites and, more importantly, the model identified 116 potentially new archaeological sites that had previously been overlooked by expert visual analyses. This not only confirms the reliability of the tool in detecting known archaeological sites, but also highlights its ability to detect sites that are too subtle to be identified by a human.

False-positives remain unavoidable limitation of automatic detection methods. However, they can be effectively mitigated by both built-in post-processing filters and supplementary geospatial analyses (Verschoof-van der Vaart et al., 2020). The prevalence of false positives, often associated with urban features, golf courses, roundabouts and other modern landscape features, highlights the need for more sophisticated exclusion strategies. This is a good pointer for future development of the tool, which could focus on integration with external datasets, such as land cover and land use maps, to enable automatic filtering based on contextual geographic information. Such approach would improve the accuracy of results and reduce the need for manual review.

The open design of the tool, based on FAIR data principles, ensures that it is not only accessible but also extendable. The creation of a high quality, expert annotated training dataset of over 23,000 image patches (512 × 512 px) from ML-ALS-IE dataset was a cornerstone of the project and provides a solid foundation for retraining the models on other datasets. The data can also serve to build new models and combine them into ensembles to increase performance. However, wider adoption outside of Ireland will require the development of additional training data and expertise to ensure transferability of the model.

To facilitate retraining, ADAF provides code for fine-tuning existing models with new data and for training entirely new models from scratch. These options allow archaeologists and researchers to adapt the tool to different geographical contexts and monument types, provided that well-annotated reference datasets are available. While retraining requires a certain level of technical proficiency and access to suitable hardware, the modular structure of the tool and the open documentation help to lower the barrier for customising ADAF to specific needs.

Implementing ADAF into the development process of TII's infrastructure plans will ensure that the project has lasting value beyond the initial research and development stage. TII has incorporated its use into its technical guidelines for the impact assessment of TII National Roads and Greenway projects (TII, 2024). As part of the Preliminary Route Options Phase 2 Stage 1 activities, consultants are contracted to also use the tool to identify previously unrecorded archaeological sites that could be avoided by considering alternative route designs. We hope that this integration will also encourage major infrastructure planning and development organisations outside Ireland to use the tool in their workflows, as well as fund its maintenance and future development.

Beyond its immediate application, ADAF holds substantial enduring value. In research, it represents a major step forward in archaeological prospection and can catalyse further innovations in the automated interpretation of remote sensing data. For public and institutional stakeholders, it provides an additional tool for large-scale heritage planning and management. Finally, for the archaeological community, ADAF serves as a reproducible, scalable and transparent framework that contributes to the standardisation of ML-based methods for the detection of archaeological sites from ALS data.

Ultimately, ADAF is an example of how the integration of deep learning and open science principles can advance the field of archaeological prospection and provide new opportunities for the discovery, documentation and protection of our cultural heritage on a large scale.

CRediT authorship contribution statement

Nejc Čož: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Anthony Corns:** Writing – review & editing, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Susan Curran:** Writing – review & editing, Resources, Investigation. **Dragi Kocev:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Žiga Kokalj:** Writing – original draft, Visualization, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

The research was funded by Transport Infrastructure Ireland Open Research Call 2021 and by the Slovenian Research and Innovation Agency core funding Earth observation and geoinformatics (P2-0406). It was also funded by the European Union (ERC, STONE, 101089123). The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Ana Kostovska for performing the machine learning experiments.

Data availability

The source code for the ADAF tool is openly available on GitHub under the Apache License 2.0: https://github.com/EarthObservation/adaf. The trained machine learning models (weights) used in this study are archived and accessible via Zenodo: https://doi.org/10.5281/zenodo.15848662. Any additional data is available from the corresponding author upon reasonable request.

References

Altaweel, M., Khelifi, A., Li, Z., Squitieri, A., Basmaji, T., Ghazal, M., 2022. Automated archaeological feature detection using deep learning on optical UAV imagery: Preliminary results. Remote. Sens. 14 (3), 553. http://dx.doi.org/10.3390/rs14030553.

Ball, J.E., Anderson, D.T., Sr, C.S.C., 2017. Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community. J. Appl. Remote. Sens. 11 (4), 042609. http://dx.doi.org/10.1117/1.JRS.11.042609.

Banasiak, P.Z., Berezowski, P.L., Zapłata, R., Mielcarek, M., Duraj, K., Stereńczak, K., 2022. Semantic segmentation (U-Net) of archaeological features in airborne laser scanning—example of the białowieża forest. Remote. Sens. 14 (4), 995. http://dx.doi.org/10.3390/rs14040995.

Bennett, R., Cowley, D., Gaffney, C., Opitz, R., Rášová, A.B., Zerboni, A., Corns, A., Russell, A., Villarejo, A.J.O., Mann, B., Collaro, C., Novák, D., Mlekuž Vrhovnik, D., Rensink, E., fovet, e., Fontana, G., Kramer, I., Herzog, I., Streatfeild-James, J., Eogan, J., Zachar, J., de Kort, J.W., Waagen, J., Lambers, K., Challis, K., Teale, K., Killoran, L., Banaszek, Ł., Meyer-Heß, M.F., Rybska, M., Kostamovaara, M., Oakey, M., Doneus, M., Kecheva, N., Crabb, N., Anttiroiko, N., Trier, Ø.D., Risbøl, O., Crow, P., O'Keeffe, P., Evans, S., Popović, S., Crutchley, S., Davis, S., Zoldoske, T., Driver, T., Fildes, T., Verschoof-van der Vaart, W.B., Kokalj, Ž., 2025. Guidelines for the Use of Airborne Laser Scanning (Lidar) in Archaeology (EAC Guidelines 10). Technical Report, European Archaeological Council, http://dx.doi.org/10.5281/ZENODO.14609210.

Bennett, R., Cowley, D.C., Laet, V.D., 2014. The data explosion: Tackling the taboo of automatic feature recognition in airborne survey data. Antiquity 88 (341), 896–905. http://dx.doi.org/10.1017/S0003598X00050766.

Bevan, A., 2015. The data deluge. Antiquity 89 (348), 1473–1484. http://dx.doi.org/10.15184/aqy.2015.102.

Bonhage, A., Eltaher, M., Raab, T., Breuß, M., Raab, A., Schneider, A., 2021. A modified mask region-based convolutional neural network approach for the automated detection of archaeological sites on high-resolution light detection and ranging-derived digital elevation models in the North German Lowland. Archaeol. Prospect. 28 (2), 177–186. http://dx.doi.org/10.1002/arp.1806.

Bundzel, M., Jaščur, M., Kováč, M., Lieskovský, T., Sinčák, P., Tkáčik, T., 2020. Semantic segmentation of airborne LiDAR data in maya archaeology. Remote. Sens. 12 (22), 3685. http://dx.doi.org/10.3390/rs12223685.

Canuto, M.A., Estrada-Belli, F., Garrison, T.G., Houston, S.D., Acuña, M.J., Kováč, M., Marken, D., Nondédéo, P., Auld-Thomas, L., Castanet, C., Chatelain, D., Chiriboga, C.R., Drápela, T., Lieskovský, T., Tokovinine, A., Velasquez, A., Fernandez-Diaz, J.C., Shrestha, R., 2018. Ancient lowland maya complexity as revealed by airborne laser scanning of northern guatemala. Science 361 (6409), eaau0137. http://dx.doi.org/10.1126/science.aau0137.

Challis, K., Howard, A.J., 2006. A review of trends within archaeological remote sensing in alluvial environments. Archaeol. Prospect. 13 (4), 231–240. http://dx.doi.org/10.1002/arp.296.

Challis, K., Kokalj, Ž., Kincey, M., Moscrop, D., Howard, A.J., 2008. Airborne lidar and historic environment records. Antiquity 82 (318), 1055–1064. http://dx.doi.org/10.1017/S0003598X00097775.

Cowley, D.C., 2012. In with the new, out with the old? auto-extraction for remote sensing archaeology. In: Bostater, C.R., Mertikas, S.P., Neyt, X., Nichol, C., Cowley, D., Bruyant, J.-P. (Eds.), SPIE Proceedings 8532. Edinburgh, 853206. http://dx.doi.org/10.1117/12.981758.

Cowley, D.C., 2016. What do the patterns mean? archaeological distributions and bias in survey data. In: Forte, M., Campana, S. (Eds.), Digital Methods and Remote Sensing in Archaeology: Archaeology in the Age of Sensing. In: Quantitative Methods in the Humanities and Social Sciences, Springer International Publishing, Cham, pp. 147–170. http://dx.doi.org/10.1007/978-3-319-40658-9_7.

Cowley, D.C., Banaszek, Ł., Geddes, G., Gannon, A., Middleton, M., Millican, K., 2020. Making LiGHT Work of large area survey? developing approaches to rapid archaeological mapping and the creation of systematic national-scaled heritage data. J. Comput. Appl. Archaeol. 3 (1), 109–121. http://dx.doi.org/10.5334/jcaa.49.

Cox, E.P., 1927. A method of assigning numerical and percentage values to the degree of roundness of sand grains. J. Paleontol. 1 (3), 179–183, URL: https://www.jstor.org/stable/1298056.

Čož, N., Kokalj, Ž., Kostovska, A., 2025a. ADAF - automatic detection of archaeological features library v1.0.0. http://dx.doi.org/10.5281/ZENODO.15861628, Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU).

Čož, N., Kokalj, Ž., Kostovska, A., Kocev, D., Corns, A., Curran, S., 2025b. Automatic detection of archaeological features – ADAF user manual. http://dx.doi.org/10.5281/ZENODO.15848155, Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU).

Crow, P., Benham, S., Devereux, B.J., Amable, G.S., 2007. Woodland vegetation and its implications for archaeological survey using LiDAR. For.: An Int. J. For. Res. 80 (3), 241–252. http://dx.doi.org/10.1093/forestry/cpm018.

Crutchley, S., 2009. Ancient and modern: Combining different remote sensing techniques to interpret historic landscapes. J. Cult. Herit. 10, e65–e71. http://dx.doi.org/10.1016/j.culher.2009.09.005.

DCHG, 2023. Department of housing, local government and heritage: national monuments service - archaeological survey of Ireland. URL: https://data.gov.ie/dataset/national-monuments-service-archaeological-survey-of-ireland. (Accessed 18 April 2023).

Dimitrovski, I., Kitanovski, I., Panov, P., Kostovska, A., Simidjievski, N., Kocev, D., 2023. AiTLAS: artificial intelligence toolbox for earth observation. Remote. Sens. 15 (9), 2343. http://dx.doi.org/10.3390/rs15092343.

Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., James, S., 2020. Machine learning for cultural heritage: A survey. Pattern Recognit. Lett. 133, 102–108. http://dx.doi.org/10.1016/j.patrec.2020.02.017.

Fiorucci, M., Verschoof-van der Vaart, W.B., Soleni, P., Le Saux, B., Traviglia, A., 2022. Deep learning for archaeological object detection on LiDAR: New evaluation measures and insights. Remote. Sens. 14 (7), 1694. http://dx.doi.org/10.3390/rs14071694.

Fylaktos, P., Petropoulos, G.P., Lemesios, I., 2025. A deep learning method for the automated mapping of archaeological structures from geospatial data: A case study of delos island. ISPRS Int. J. Geo-Information 14 (6), 220. http://dx.doi.org/10.3390/ijgi14060220.

Guyot, A., Lennon, M., Lorho, T., Hubert-Moy, L., 2021. Combined detection and segmentation of archeological structures from LiDAR data using a deep learning approach. J. Comput. Appl. Archaeol. 4 (1), 1–19. http://dx.doi.org/10.5334/jcaa.64.

Halliday, S., 2013. I walked, i saw, i surveyed, but what did i see?…and what did i survey? In: Opitz, R.S., Cowley, D.C. (Eds.), Interpreting Archaeological Topography: Airborne Laser Scanning, 3D Data and Ground Observation. Oxbow Books, Oxford and Oakville, UK, pp. 63–75. http://dx.doi.org/10.2307/j.ctvh1dqdz.10.

Harmon, J.M., Leone, M.P., Prince, S.D., Snyder, M., 2006. LiDAR for archaeological landscape analysis: a case study of two eighteenth-century maryland plantation sites. Am. Antiq. 71 (4), 649–670. http://dx.doi.org/10.2307/40035883.

Hesse, R., 2010. LiDAR-derived local relief models – a new tool for archaeological prospection. Archaeol. Prospect. 17 (2), 67–72. http://dx.doi.org/10.1002/arp.374.

Historic England, 2025. Scheduled monuments GIS data. URL: https://www.data.gov.uk/dataset/640d5dab-4295-4b85-bc26-a8369abb73f6/scheduled-monuments-gis-data. (Accessed 18 April 2023).

Kadhim, I., Abed, F.M., 2023. A critical review of remote sensing approaches and deep learning techniques in archaeology. Sensors 23 (6), 2918. http://dx.doi.org/10.3390/s23062918.

Kocev, D., Simidjievski, N., Kostovska, A., Dimitrovski, I., Kokalj, Ž., 2022. Discover the mysteries of the maya: selected contributions from the machine learning challenge & the discovery challenge workshop at ECML PKDD 2021. http://dx.doi.org/10.48550/arXiv.2208.03163, arXiv:2208.03163.

Kokalj, Ž., Maroh, Ž., Oštir, K., Zakšek, K., Čož, N., 2024. Relief visualization toolbox python library. original-date: 9/7/2020, URL: https://github.com/EarthObservation/RVT_py.

Kokalj, Ž., Somrak, M., 2019. Why not a single image? combining visualizations to facilitate fieldwork and on-screen mapping. Remote. Sens. 11 (7), 747. http://dx.doi.org/10.3390/rs11070747.

Küçükdemirci, M., Landeschi, G., Ohlsson, M., Dell'Unto, N., 2023. Investigating ancient agricultural field systems in Sweden from airborne LIDAR data by using convolutional neural network. Archaeol. Prospect. 30 (2), 209–219. http://dx.doi.org/10.1002/arp.1886.

Maxwell, A.E., Pourmohammadi, P., Poyner, J.D., 2020. Mapping the topographic features of mining-related valley fills using mask R-CNN deep learning and digital elevation data. Remote. Sens. 12 (3), 547. http://dx.doi.org/10.3390/rs12030547.

Orengo, H.A., Conesa, F.C., Garcia-Molsosa, A., Lobo, A., Green, A.S., Madella, M., Petrie, C.A., 2020. Automated detection of archaeological mounds using machine-learning classification of multisensor and multitemporal satellite data. Proc. Natl. Acad. Sci. 117 (31), 18240–18250. http://dx.doi.org/10.1073/pnas.2005583117.

Pingel, T.J., Clarke, K., Ford, A., 2015. Bonemapping: A LiDAR processing and visualization technique in support of archaeology under the canopy. Cartogr. Geogr. Inf. Sci. 42 (Suppl. 1), 18–26. http://dx.doi.org/10.1080/15230406.2015.1059171.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. http://dx.doi.org/10.48550/ARXIV.1506.01497, Version Number: 3.

Simidjievski, N., Kocev, D., Kostovska, A., Dimitrovski, I., Kokalj, Ž., Saux, B.L., Draksler, A., Somrak, M., Veljanovski, T., Kitanovski, I., Stepišnik, T., Petković, M., Panov, P., Džeroski, S., Aparicio, S., Loekken, S., 2021. Discover the mysteries of the maya: ECML PKDD 2021 - discovery challenge. URL: https://biasvariancelabs.github.io/maya_challenge/about/. (Accessed 22 April 2021).

Simon, M., Rodner, E., Denzler, J., 2016. ImageNet pre-trained models with batch normalization. http://dx.doi.org/10.48550/ARXIV.1612.01452.

Somrak, M., Džeroski, S., Kokalj, Ž., 2020. Learning to classify structures in ALS-derived visualizations of ancient maya settlements with CNN. Remote. Sens. 12 (14), 2215. http://dx.doi.org/10.3390/rs12142215.

Soroush, M., Mehrtash, A., Khazraee, E., Ur, J.A., 2020. Deep learning in archaeological remote sensing: automated qanat detection in the kurdistan region of Iraq. Remote. Sens. 12 (3), 500. http://dx.doi.org/10.3390/rs12030500.

Štular, B., Eichert, S., Lozić, E., 2021. Airborne LiDAR point cloud processing for archaeology. pipeline and QGIS toolbox. Remote. Sens. 13 (16), 3225. http://dx.doi.org/10.3390/rs13163225.

TII, 2024. Guidelines for Cultural Heritage Impact Assessment of TII National Road and Greenway Projects PE ARC-02009. TII Publications, URL: https://cdn.tii.ie/publications/PE-ARC-02009-01.pdf. (Accessed 23 January 2025).

Trier, Ø.D., Cowley, D.C., Waldeland, A.U., 2018. Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. Archaeol. Prospect. 26 (2), 165–175. http://dx.doi.org/10.1002/arp.1731.

Trier, Ø.D., Reksten, J.H., Løseth, K., 2021. Automated mapping of cultural heritage in Norway from airborne lidar data using faster r-CNN. Int. J. Appl. Earth Obs. Geoinf. 95, 102241. http://dx.doi.org/10.1016/j.jag.2020.102241.

Verschoof-van der Vaart, W.B., 2022. Learning to Look at LiDAR: Combining CNN-based Object Detection and GIS for Archaeological Prospection in Remotely-Sensed Data (Ph.D. thesis). Leiden University, Leiden, arXiv:1887/3256824.

Verschoof-van der Vaart, W.B., Lambers, K., 2019. Learning to look at LiDAR: the use of r-CNN in the automated detection of archaeological objects in LiDAR Data from the Netherlands. J. Comput. Appl. Archaeol. 2 (1), 31–40. http://dx.doi.org/10.5334/jcaa.32.

Verschoof-van der Vaart, W.B., Lambers, K., 2021. Applying automated object detection in archaeological practice: A case study from the southern Netherlands. Archaeol. Prospect. 1–17. http://dx.doi.org/10.1002/arp.1833.

Verschoof-van der Vaart, W.B., Lambers, K., Kowalczyk, W., Bourgeois, Q.P.J., 2020. Combining deep learning and location-based ranking for large-scale archaeological prospection of Lidar data from The Netherlands. ISPRS Int. J. Geo-Information 9 (5), 293. http://dx.doi.org/10.3390/ijgi9050293.

Verschoof-van der Vaart, W.B., Landauer, J., 2021. Using CarcassonNet to automatically detect and trace hollow roads in LiDAR data from the Netherlands. J. Cult. Herit. 47, 143–154. http://dx.doi.org/10.1016/j.culher.2020.10.009.

Wang, S., Hu, Q., Wang, S., Ai, M., Zhao, P., 2024. Archaeological site segmentation of ancient city walls based on deep learning and LiDAR remote sensing. J. Cult. Herit. 66, 117–131. http://dx.doi.org/10.1016/j.culher.2023.11.005.

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., Xiao, B., 2020. Deep high-resolution representation learning for visual recognition. URL: https://arxiv.org/abs/1908.07919, arXiv:1908.07919.