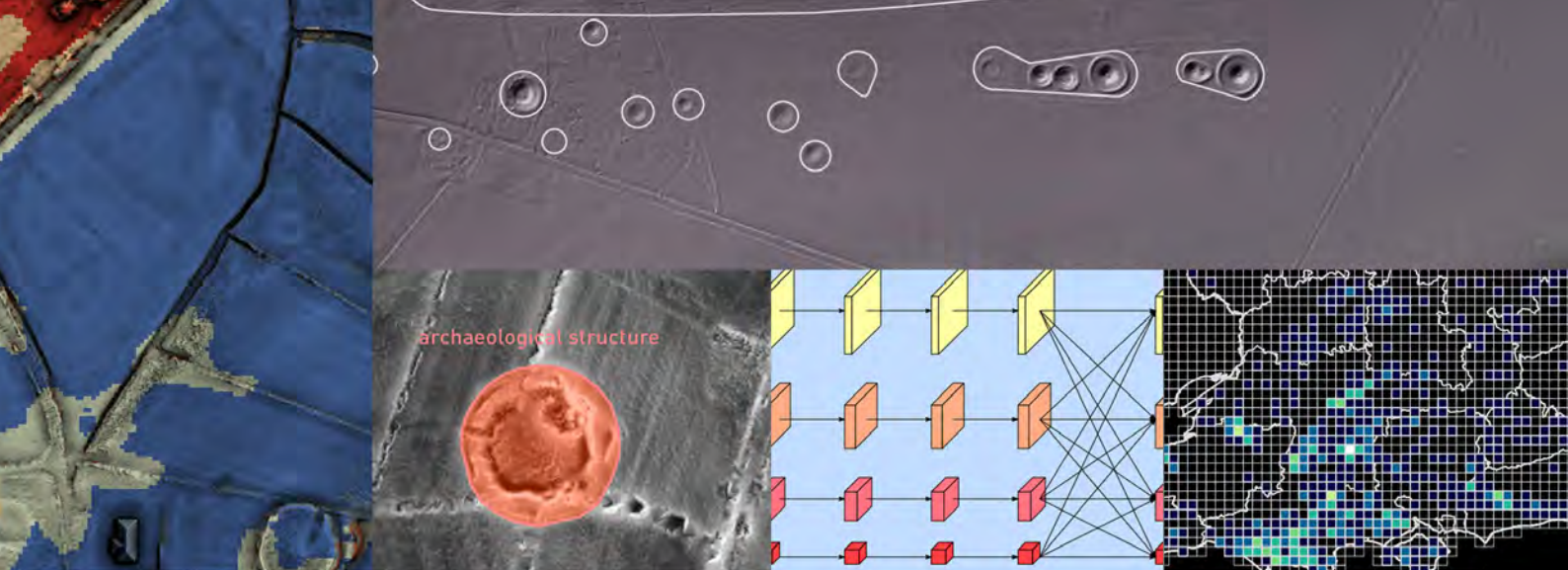




ALS/MLM Project Research Review & Recommendations Report

Version 1.3

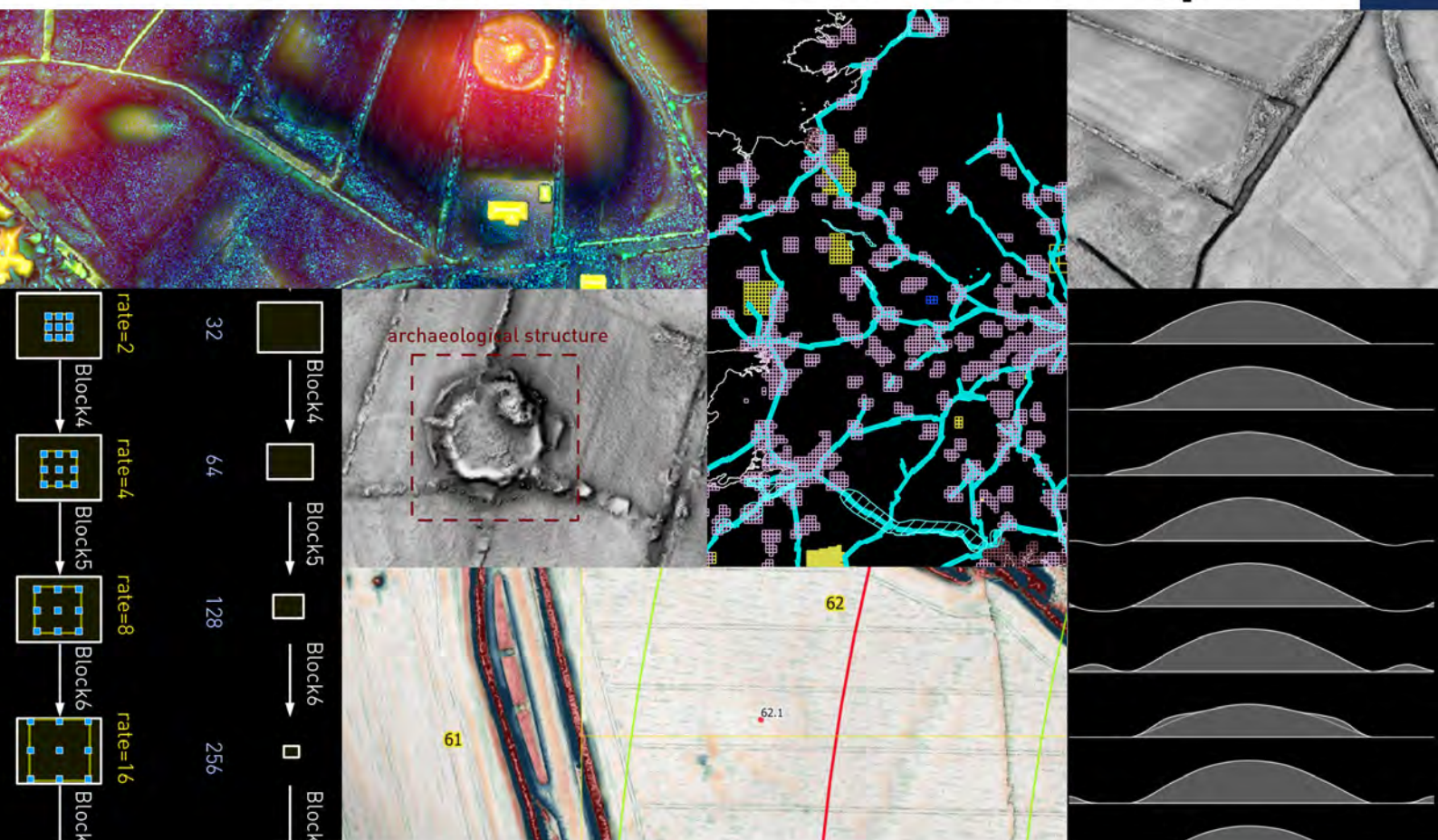


Table of Contents

1. Introduction.....	16
2. Pre-processing.....	19
2.1. DEM Specification	20
2.2. Visualisations	23
2.3. Learning dataset creation	29
2.4. Background (negative) examples.....	31
2.5. Data augmentation.....	32
3. Detection	32
3.1. Pre-training (and transfer learning).....	33
3.2. Frameworks	33
3.3. Object detection methods.....	34
3.3.1. Challenges and issues with using machine learning in archaeology	35
3.3.2. Selected candidate machine learning methods	39
4. Post-processing.....	47
4.1. Location-based ranking.....	47
4.2. Corroboration with other remote sensing data.....	47
4.3. Feedback loop.....	48
5. Metrics.....	48
5.1. Classification-based evaluation measures.....	49
5.1.1. Precision, Recall and F1 score.....	49
5.1.2. Accuracy	50
5.1.3. Area under the receiver operating characteristic curve (AUROC)	51
5.1.4. Matthews Correlation Coefficient (MCC)	52
5.2. Segmentation-based evaluation measures.....	52
6. Implications of using ML	52
7. Section I References	54
8. Statutory Definitions	59
9. Classification of Monuments.....	60
9.1. Why do we classify monuments according to particular ‘types’?.....	61
9.2. How are monuments classified.....	62
10. The use of ALS for archaeological survey	70
10.1. Limitations and Bias	72
11. Applied use of lidar in Ireland on the study and protection of archaeology.....	78
11.1. Initial use of ALS in Ireland	78
11.2. High-resolution surveys: Tara and beyond	79
11.3. ALS Data captured for non-heritage purposes.....	81
11.4. ALS for infrastructure in Ireland.....	82
11.5. Future uses of ALS within Irish archaeology	84

12. Methodology for monument classification selection.....	86
12.1. Stage 1 Filter	87
12.2. Stage 2 Filter	88
12.3. Stage 3 Filter	88
12.4. Stage 4 Filter - Shortlist	88
12.5. Stage 5 Filter – Final Selection	95
13. Supplementary ALS data	95
13.1. Evaluation of UK sites and monument data.....	96
13.2. Evaluation of UK ALS data	98
14. General observations	100
15. Conclusions and Recommendations.....	103
16. Section II References.....	105
17. Appendix i: Additional machine learning resources	110
18. Appendix ii: Monument classifications or features that may be considered to be ‘Enclosures’	112
19. Appendix iii: Classifications of Barrows as contained in the current Archaeological Survey of Ireland’s Class List Definitions	115
20. Appendix iv: Classifications of Ecclesiastical enclosures, Enclosures, Large Enclosures and Ringforts as contained in the current Archaeological Survey of Ireland’s Class List Definitions	116

Table of Figures

Figure 1: The original 2 m resolution DTM as provided by TII (a). Rasterised data to 1m (b) and 0.5 m resolution (c), and a 0.5 m resolution rasterised classified data (d). The original point cloud is not classified and the elevation models therefore include all of the above ground objects (b, c). Most of the vegetation is removed with a proper classification and the DEM shows bare earth with buildings (d).	19
Figure 2: Different algorithm settings for ground classification. 'Aggressive' algorithm settings for removal of all vegetation and large buildings (a), removal of some vegetation and smaller houses (b), classification with 'mild' settings, mindful of archaeology (c), combined (a) and (c) approach (d). Rasterised (gridded) with IDW from Noise Mapping ALS data.	21
Figure 3: Two approaches to rasterise the DEM were tested. Interpolation with Triangulated Irregular Network (TIN) (a) and Inverse Distance Weighted (IDW) (b). Reprocessed Noise Mapping ALS data.....	22
Figure 4: A combination of TIN- and IDW-interpolated rasters (a), based on the confidence map (b) (Štular, Eichert, and Lozić 2021). Reprocessed Noise Mapping ALS data.	22
Figure 5: Classification with 'mild' settings (a) and (c), and a combined approach (b) and (d). Rasterised with TIN from Reprocessed NM20 ALS data (a) and (b), and with IDW (c) and (d). Although the combined approach offers some advantages over simple classification mindful of archaeology and seems better suited for lower density datasets such as the Noise Mapping data, it emphasises some data irregularities such as flight line misalignments, as can be seen in the highlighted noise in the eastern-most meadow. This is not so evident in data rasterised with IDW.	23
Figure 6: Visualisations commonly reported in archaeological assessment reports provided by TII. Relief shading (a), multidirectional hillshading (b), sky-view factor (c), simple local relief model (SLRM) (d), multidirectional hillshade overlaid by SLRM (e), SVF overlaid by SLRM (f). 25	
Figure 7: Visualisations commonly reported in papers dealing with archaeological exploration with deep learning CNN models. Slope (a), local dominance (b), positive openness (c), negative openness (d), visualisation of archaeological topography combining settings for flat and general terrain (e), red relief image map (f), multiscale topographic position (g), multiscale topographic position enhanced version 2 (h). SLRM is often cited as well (Figure 6d).....	26
Figure 8: Image analysis using deep learning Convolutional Neural Networks for an archaeological site (a curvilinear enclosure). Image classification: a class or label associated with the image (a); object detection: a labelled bounding box that locates the object of interest within the image (b), and object segmentation: a labelled footprint that locates and delineates the object of interest within the image (c). Reprocessed TII Noise Mapping data. The bounding box in (a) and the mask in (b) are hand drawn.	29
Figure 9: a) R-CNN architecture: Object detection system overview. The system takes an input image, extracts around 2000 bottom-up region proposals, computes features for each proposal using a large convolutional neural network (CNN), and then classifies each region using class-specific linear SVMs. b) Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network.	39

Figure 10: YOLO Architecture. The detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. The convolutional layers are pre-trained on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.	40
Figure 11: YOLOR multi-purpose NN architectures. (a) distinct models for distinct tasks; (b) shared backbone, different heads for different tasks; and (c) our proposed unified network: one representation with explicit knowledge and implicit knowledge for serving multiple tasks.	41
Figure 12: DeepLabV3 architecture for image segmentation illustrating the different architectures to capture multi-scale context.	42
Figure 13: DeepLabV3 architecture for image segmentation illustrating cascaded modules without (a) and with (b) atrous convolution.	42
Figure 14: An example of a high-resolution network (HRNET). Only the main body is illustrated, and the stem (two stride-2 3×3 convolutions) is not included. There are four stages. The 1st stage consists of high-resolution convolutions. The 2nd (3rd, 4th) stage repeats two-resolution (three-resolution, four-resolution) blocks.	43
Figure 15: Segmentation transformer: The encoder self-attention unit in the grey box could be a local version and is optional and is useful for boosting the image features. The decoder self-attention unit serves as the role of interacting the category queries and.	44
Figure 16: (a) The proposed Swin Transformer builds hierarchical feature maps by merging image patches (shown in grey) in deeper layers and has linear computation complexity to input image size due to computation of self-attention only within each local window (shown in red). It can thus serve as a general-purpose backbone for both image classification and dense recognition tasks. (Middle image) In contrast, previous vision Transformers [20] produce feature maps of a single low resolution and have quadratic computation complexity to input image size due to computation of self-attention globally. (b) In order to better scale model capacity and window resolution, several adaptations are made on the original Swin Transformer architecture (V1): 1) a post-norm to replace the previous pre-norm configuration; 2) a scaled cosine attention to replace the original dot product attention; 3) a log-spaced continuous relative position bias approach to replace the previous parameterized approach. Adaptions 1) and 2) make the model easier to be scaled up in capacity. Adaption 3) makes the model more effectively transferred across window resolutions. The adapted architecture is named Swin Transformer V2.	45
Figure 17: Figure 15. (a) The architecture of a Swin Transformer (Swin-T); (b) two successive Swin Transformer Blocks (for the notation see (Liu et al., 2021b)). W-MSA and SW-MSA are multi-head self-attention modules with regular and shifted windowing configurations, respectively.	45
Figure 18: Overview of SpotNet: the input image first passes through a double-stacked hourglass network; the segmentation head then produces an attention map that multiplies the final feature map of the backbone network; the final centre keypoint heatmap is then produced as well as the size and coordinate offset regressions for each object.	46
Figure 19: Illustration of Intersection-over-union (IoU) as a similarity measure for object detection on images. Image taken from Wikipedia (https://en.wikipedia.org/wiki/Jaccard_index).	52

Figure 20: Bar chart displaying the total number of sites recorded within the SMR grouped by monument classification for the 10 most frequently recorded classes. Percentage values on the top of the bar chart indicate the percentage of total sites each monument classification represents. Calculation excludes the classification of "Redundant record" (data extracted from ASI on 07/01/2022).	64
Figure 21: Example morphologies of barrows as evidenced at Tara (after Newman, 1997).	68
Figure 22: The effects of progressive denudation on selected monument types including mounds, enclosures and similar barrow forms, demonstrating how this may affect classification: a) denudation only by weather; b) early stage of attrition through activities such as ploughing; c) advanced stage of attrition; d) upstanding elements levelled; e) subsoil reduction and truncation of sub-surface features. Notations: M: Mound barrow, upstanding earthwork when first recorded. E: Enclosure barrow, upstanding earthwork when first recorded. L: Very low mound when first recorded (generally less than 0.35m maximum height). R: Ring-ditch or circular patch when first recorded. U: Unclassified: mound (UM), very low mound (UL), enclosure (UE), or unknown (U); unclassified sites are those not accepted as barrows on current evidence, grades 4 and 5 above; they do not contribute to analysis. (after Needham and Anelay, 2021).	69
Figure 23: Profiles of mound and enclosure barrow variants. Notations: M: Mound barrow. E: Enclosure barrow. b: encircling bank; b*: embanked slope; d: encircling ditch; e: external bank and ditch; f: a flat, near-level annular zone - or berm - encircling the foot of the mound and enclosed by a ditch and/or bank; g: interior ground surface; s: skirt; s*: step. (after Needham and Anelay, 2021).	70
Figure 24: A Johari Window model of historic monument recognition	76
Figure 25: ALS derived DSM relief shaded model of the Brú na Bóinne landscape, enlargement around Newgrange passage tomb (Meath County Council/The Heritage Council/The Discovery Programme).	79
Figure 26: Sky view factor visualisation of FliMAP derived ALS data of the whole Skellig Michael with enlargements at the monastery and south peak (DoHLGH/The Discovery Programme). The use of high-resolution ALS surveys (commonly 20-40 points per square metre and occasionally more) in heritage research has become increasingly commonplace, and the commissioning of new datasets is now a frequent component of site management practice.	80
Figure 27: Example of a potential low-profile enclosure (62.1) identified as part of the N11-N25 Oilgate-Rosslare Harbour scheme archaeological assessment of ALS data. Visualisation scheme used LRM over SVF (Davis 2021b).	83
Figure 28: Screen image of the GSI Open Topographic Viewer. Inset highlights three separate organisational ALS datasets (OPW NASC/NPWS/TII) across the same region of County Offaly. (https://dcenr.maps.arcgis.com/apps/webappviewer/index.html?id=b7c4b0e763964070ad685)	85
Figure 29: Flowchart illustrating the staged processes involved in the analysis and proposal of the sites and monument classification scheme to use within the project.	87
Figure 30: Map illustrating the survey extents of the relevant ALS datasets which could be potentially used within the project.	91
Figure 31: Illustration of the overestimation of the number of monument contained within the ALS indexes provided by the GSI.	92

Figure 32: Image displaying part of the ALS derived relief shaded model for GB grid tile SU14SW located in County Wiltshire. Each of the enclosing polygons represents the scheduled area which contains at least one example of a barrow site. For this tile there are 89 scheduled barrow polygons	97
Figure 33: Map illustration of the density of scheduled barrows within The Environment Agency's National LIDAR Programme. Each grid cell corresponds to individual ALS tiles (5km tiles aligned to the ordnance survey grid).	99
Figure 34: Map illustration of the density of scheduled barrows within Scotland. The gey area corresponds to the ALS survey extent captured through Phase I to Phase V of the Scottish Lidar acquisition 2009 – 2021. Each grid cell corresponds to individual ALS tiles (5km tiles aligned to the ordnance survey grid).	100

List of acronyms

AI	<i>Artificial Intelligence</i>
ALS	<i>Airborne laser scanning / Airborne laser system / Airborne LiDAR system</i>
ANN	<i>Artificial neural network</i>
AP	<i>Average precision</i>
ASPP	<i>Atrous spatial pyramid pooling</i>
ASCII	<i>American Standard Code for Information Interchange</i>
ASI	<i>Archaeological Survey of Ireland</i>
AUC	<i>Area under the curve</i>
AUROC	<i>Area under the receiver operating characteristic curve</i>
CNN	<i>Convolutional neural network</i>
DoHLGH	<i>The Department of Housing, Local Government and Heritage</i>
DCHG	<i>The Department of Culture, Heritage and the Gaeltacht</i>
DEM	<i>Digital elevation model</i>
DFM	<i>Digital feature model</i>
DMV	<i>Deserted medieval village</i>
DP	<i>Discovery Programme</i>
DSM	<i>Digital surface model</i>
DTM	<i>Digital terrain model</i>
e2MSTP	<i>Enhanced Multiscale Topographic Position version 2</i>
FN	<i>False negative</i>
FP	<i>False positive</i>
FPR	<i>False positive rate</i>
GIS	<i>Geographic information system</i>
GPR	<i>Ground penetrating radar</i>
GPU	<i>Graphics processing unit</i>
GSI	<i>Geological Survey Ireland</i>
HERs	<i>Historic Environment Records</i>
HS	<i>Hillshade</i>
IDW	<i>Inverse distance weighting</i>
IoU	<i>Intersection over union</i>
LBR	<i>Location-based ranking</i>
LD	<i>Local dominance</i>

LiDAR	<i>Light detection and ranging</i>
LR	<i>Learning rate</i>
mAP	<i>Mean average precision</i>
MCC	<i>Matthews correlation coefficient</i>
ML	<i>Machine learning</i>
MLA	<i>Machine Learning Algorithm</i>
MLM	<i>Machine learning model(s)</i>
MSA	<i>Multiple sequence alignments</i>
MSTP	<i>Multiscale topographic position</i>
OBIA	<i>Object-based image analysis paradigm</i>
OPW	<i>Office of Public Works</i>
OSI	<i>Ordnance Survey Ireland</i>
OvO	<i>One vs one</i>
OvR	<i>One vs rest</i>
R-CNN	<i>Region-based Convolutional Neural Network</i>
RF	<i>Random forest</i>
RGB	<i>Red-Green-Blue</i>
ROC	<i>receiver operating characteristic</i>
RPN	<i>Region proposal network</i>
RRIM	<i>Red relief image map</i>
SLRM	<i>Simple local relief model</i>
SMR	<i>Sites and Monuments Record</i>
SNR	<i>Signal-to-noise ratio</i>
SVF	<i>Sky-view factor</i>
SVM	<i>Support vector machine</i>
SW-MSA	<i>Shifted window multiple sequence alignments</i>
TII	<i>Transport Infrastructure Ireland</i>
TIN	<i>Triangulated irregular network</i>
TN	<i>True negative</i>
TP	<i>True positive</i>
TPR	<i>True positive rate</i>
UAV	<i>Unmanned Aerial Vehicle</i>
VAT	<i>Visualisation for Archaeological Topography</i>

WHS	<i>World Heritage Site</i>
W-MSA	<i>Window multiple sequence alignments</i>
YOLO	<i>You Only Look Once</i>

Executive Summary & Report

Recommendations

The TII ALS Machine Learning Research Project aims to develop a user-friendly software tool capable of automatically detecting archaeological features from LiDAR data using machine learning, specifically deep learning algorithms. This will include the development of a 'trained' machine learning model that can automatically detect three classes of earthwork monuments from LiDAR data. The software tool will initially be designed for TII staff and to archaeological consultants appointed to TII projects and subsequently be made more widely available as free and opensource software to provide opportunities for future collaboration and development. This report is the first stage of this project and aims to explore the several critical aspects which will inform the decisions and future direction of project activities.

The report is divided into two sections:

Section I: Technical literature review – A comprehensive technical literature review of the current state of the art of automatic detection from ALS data (and other remote sensing data) using deep MLM. The review focuses on the current recommendations and limitations regarding detection of archaeological features, and also include other potentially useful methods and strategies. Specific areas of focus include:

- The pre-processing of ALS data – visualisation methodology when applied to MLM, creation of the training dataset and data augmentation
- Object detection methods and MLM frameworks
- Post processing MLM results against additional datasets
- Metrics and MLM evaluations strategies

Section II: Irish monument detection & suitability review – An in-depth exploration and evaluation of the potential application of automated detection processes on Irish monument forms and shortlisting of suitable earthwork monument types for the training process. Subjects discussed within this section include:

- Discussion on the development and concept of monument classification and its relationship to the physical representation in the landscape
- A review of the use of ALS within archaeological survey in Ireland and future developments
- Explanation of the methodology utilised to inform the selection of monument classifications to be included in this research project
- A proposition to supplement training data with additional monuments outside of Ireland

Technical Recommendations

Following on from the technical literature review, the following technical recommendations have been proposed:

General project guidance recommendation

- The primary interest of the MLM is the detection of archaeological remains across the entire dataset as discrete spatial objects. The whole process has to create information that is systematic and well-documented (i.e., with thorough metadata).

Pre-processing recommendation

- Recommendation: We propose to reprocess (classify, grid, visualise) the entire Noise Mapping dataset. Following a case-by-case basis of review of additional ALS datasets we will decide whether to reprocess them, weighing the benefits this might bring against the computational and time resources required.

DEM Specification recommendations

- Data must be handled more carefully for automatic detection in comparison to human (visual) inspection which can tolerate systemic irregularities in the data. For example, border edge transitions need to be seamless, and small areas without data need to be interpolated.
- Based on the characteristics of the ALS data provided by TII and obtained from other providers, we propose to create DFMs at 0.5 m resolution.
- Since much of the Irish landscape is gently sloping and TIN is the preferred method of data rasterization by ALS data providers, we recommend using this method for this project.

Visualisation recommendations

- Although the idea of using a DTM as a direct input to the ML process is worth pursuing, we believe that further studies are needed to prove its effectiveness against using advanced visualisation techniques. We suggest testing this within the project before a list of proposed inputs is finalised.
- Further testing is needed to determine which visualisation(s) produce the best results with Irish topography and selected archaeological types. We will also conduct tests with a new, customised visualisation that highlights local topographic variations, provides multiscale information, and is directionally invariant. Such a visualisation will be suitable for a wide range of data augmentation methods.

Learning dataset creation recommendations

- The sample dataset should contain only archaeological structures whose topographic characteristics are identifiable on the DTM. Sites that are considered above ground structures with no influence on the bare-earth topography, such as standing stones, or sites that are associated related only to small-object deposits should be excluded.
- The decision on the exact procedure to construct the training, validation and test sets will be made after the review of these reports by the whole project team and the TII and detailed in the Final Technical Report.

Background (negative) examples recommendation

- The CNN testing should be carried out which include negative examples.

Data augmentation recommendation

- We propose conducting tests that include a wide range of data augmentation techniques including at least rotation, horizontal and vertical flipping, scaling, blurring and sharpening, brightness and contrast alterations, translation, truncation, and noise injection.

Object detection pretraining recommendation

- Pre-training requires further exploration especially as archaeological learning sets are small in comparison to other training sets.

Candidate machine learning methods recommendations

- The work needs to be highly interdisciplinary involving experts from archaeology, remote sensing, earth sciences and machine learning
- The data needs to be prepared carefully, and special effort should be put on removing data biases
- The usage of data augmentation is strongly recommended
- There is no clear top performing model, rather a set of models performing well across different datasets, hence benchmarking several methods on the project data is a must
- The project proposes that the following models are evaluated and employed R-CNN, YOLO, DeepLabV3, HRNET, Swin Transformers, and SpotNET

Post processing recommendations

- Where possible historic land-use assessment and historic landscape characterisation datasets will be evaluated as an indicator for different zone ranks.
- The project will test the effectiveness of applying the Feedback loop through the initial labelling of archaeological objects on the performance of the model.

Metrics recommendation

- The project will test and assess several accuracy and evaluation strategies including AUROC, MCC and IoU during the development of the MLM.

Archaeological Recommendations

Following on from the Irish monument detection & suitability review, the following archaeological recommendations have been proposed:

Representation of monuments and sites in ALS data

- It is proposed that quality control of the dataset will be carried out before the training samples are created

Augmentation of training datasets

- If required, the training datasets created for barrows will be augmented through the inclusion of selected sites within England and Scotland and their corresponding ALS datasets following quality control measures. Full specifications and details of all the LAS datasets utilised within the project will be provided within the Final Technical Report

Sites and monument classification selection

- The tender specification notes that the classes of monuments to be included in the model will be agreed between TII and the Project Team. Based on the above review and evaluation of the potential application of automated detection processes on Irish monument forms and shortlisting of suitable earthwork monument types for the training process, the following monument classifications are recommended: Ringforts, Enclosures and Barrows

Section I: Technical literature review

1. Introduction

We start the report with the suggestion by Cowley et al. (2020) that ‘a systemic consideration of the implications of changing technology and data is sometimes desirable, rather than gradual assimilation of developments into existing practice. In particular, the issues being addressed relate to the challenges and opportunities of proliferating remote sensed data and associated digital workflows. These include the strategic assessment of threat, consideration of fitness for purpose of different datasets relative to landscape characteristics, the documentation of processes and sources of information (paradata & metadata), the suitability of data structures, and the mechanisms for automating site detection and data creation.’

This report summarises recent developments in the field of archaeological object recognition from airborne laser scanning data (ALS) using machine learning, and reviews state-of-the-art convolutional neural network (CNN) methods used in remote sensing in general. We first provide insight into why artificial intelligence should be used with proper understanding, and then proceed with considerations of required performance and completeness. The text is largely based on the literature reviewed and is supplemented by comments, recommendations, and insights.

In the context of this project, we divide the machine learning process into three major technical steps, each with several important sub-steps. The major ones are:

1. *Pre-processing*, which includes procedures and recommendations for creating the digital elevation model (DEM) (point cloud classification, gridding...), computing visualisations, creating the learning set, and data augmentation.
2. *Detection*, which deals with pre-training and transfer learning, frameworks and CNN methods.
3. *Post-processing*, which includes methods to improve detection rates and remove false positive detections. Some of these methods include location-based ranking, corroboration with other remote sensing data, and a feedback loop.

In computer vision, the term ‘feature’ refers to the properties of an image, while an ‘object’ refers to real-world entities (Traviglia, Cowley, and Lambers 2016, 14). In this report, therefore, the term ‘objects’ is used to refer to archaeological features, although some stray ‘features’ may still appear.

Artificial intelligence should be applied with proper understanding

Machine learning approaches require computational skills that are not routinely taught in archaeological training. Where archaeologists have taken the plunge, the complex alignment of sources, technical capacities and research questions required can produce disappointing outcomes. As a result, there is some understandable scepticism towards their practical utility (Casana 2014). There is a need to understand the performance of neural networks, including factors such as how they are trained, how they learn, how transferable systems are from one area to another, and how and why outputs from one network may differ from another. For example, the influence of pre-training of neural networks on down-the-line outputs is currently unclear, or how the character of inputs (i.e., which visualisation, what processing parameters, etc.) may influence the results (Cowley et al. 2020). This is supported by Trier et al. (2018, 9) who urge for exploration of the complex reasons for the variability in performance of deep neural networks. These reasons are varied, but they cite the (in)suitability of a particular deep

neural network structure for the pattern recognition problem at hand and the insufficient number of archaeological training examples as two important ones. It should also be emphasised that while it is only right to call for a better understanding of how and why artificial intelligence applications work, this also means that the practices and interpretations on which archaeologists identify and classify archaeological monuments needs to be made clearer. Black boxes in identification and classification processes are undesirable, whether they are expressed computationally or by the archaeologist. Synergistic research into the performance of neural networks and the way archaeologists identify and classify remains appears to be a useful shared concern.

On required performance and completeness

Through the efficient detecting and mapping of the presence and location of archaeological objects, machine learning facilitates the investigating of trends within the distribution and interrelationship of these objects in the landscape, the emerging of large-scale patterns between different types of objects and the structuring of the landscape in the past. Understanding this spatial relation between archaeological objects and their surroundings, that is, the landscape, lies at the core of landscape archaeology (Verhoeven 2017).

One of the main questions in the design of an object detection tool is what level of performance is acceptable? The required performance depends on:

1. the (envisioned) users,
2. the intended task of the tool,
3. the incorporation of the tool (and its results) within the wider archaeological research framework.

The question of completeness deals with how accurately the results reflect the extent of the identified archaeological objects.

Information gathered with machine learning can only be considered an approximation of the true number, size, form and location of the objects of interest, and as early as 2012, Cowley advocated for '*a key conceptual shift from a widely held fixation in archaeology on individual identifications being correct to overall patterns being descriptive*' (Cowley 2012).

The detection performance is related to the ability of a visualisation to enrich the original data representation with interpretable information. However, the image data is not the only information that influence human-based interpretation. An expert-based interpretation would include perceptions of the information included in the image, but also additional external information not contained in the data itself (e.g., geoarchaeological knowledge of the area or skills and knowledge based on experience). The computer-vision approach applied to an image cannot encompass fully the exhaustive aspects that influence archaeological interpretation. Therefore, the approaches should not aim at replacing expert-knowledge or imposing a single visualisation for archaeological interpretation of ALS-derived terrain models, but rather aim at proposing a significant support tool for archaeological analysis. With a high capacity of data processing, a consistent response against similar data and an interpretation bias (even if not entirely absent) that can be measured, the deep CNN approach provides a complementary tool for the identification and characterization of archaeological structures from ALS-derived relief model (Guyot, Lennon, and Hubert-Moy 2021).

If the percentage of false positive and negative detections produced through the application of CNN were similar to that of human experts, a machine learning procedure could save a large

amount of time, effort and money. For their study of ancient Mayan settlement, Canuto et al. (2018) state that human experts failed to discover approximately fifteen per cent of the objects in the research area, but do not provide information on the rates of their false positive and negative examples. Verschoof and Lambers (2021, 13) also note that when a part of their study area was re-examined, based on the results of the automated detection, even more potential archaeological objects were discovered, overlooked during both the manual analysis and the automated detection. Bundzel et al. (2020) therefore judge we have reason to believe that we are approaching the goal of automated detection at least in some landscapes and with particular archaeological objects. It is usually the very small objects that pose problems for machine learning, but these are also the most subjective in man-made interpretations because they can easily be confused with natural elements such as uprooted trees or large anthills. Medium-sized structures are therefore the most objective factor of identification from a practical point of view.

Levels of performance and completeness must be high when automated detection is used independently as the sole source of information, especially when the objects of interest are landscape patterns such as field systems or roads. Alternatively, if the tool is used to aid in limited manual inspection of the data or is used in conjunction with a manual analysis to improve results, a certain degree of exclusion can be afforded (Verschoof-van der Vaart and Lambers 2021, 13). In developing a tool for TII, this means that both performance and completeness must be high, but in contrast, it is more important to determine the true location of previously unknown (unrecorded) archaeological remains rather than to correctly classify them. We must therefore accept that rare classes of monument types could have the potential to be confused with more numerous classes. However, uncertainty between classes of archaeological structure may be regarded as a minor issue, when the identification of a broad range of archaeological site types is the main objective. We are more interested in the rates of false positive and false negative identifications of objects of archaeological interest. Cowley et al. (2020, 110) summarise on the subject – *what is not known cannot be managed or protected, or inform our understanding of the past*. Therefore, it is better to accept a higher number of false positive detections than have a large number of true negative detections (missed true objects).

The results of automated detection can be regarded as showing highlighted areas of interest – containing potential archaeological objects that require additional validation and verification (potentially in the field). When used in this way, the results are very comparable, in type and value, to archaeological predictive maps. The level of competence and especially completeness of an automated detection tool does not have to be extremely high as the results are merely one of multiple consulted data sources, that form the basis for subsequent fieldwork (Verschoof-van der Vaart and Lambers 2021, 13–14).

Recommendation: The primary interest of the MLM is the detection of archaeological remains across the entire dataset as discrete spatial objects. The whole process has to create information that is systematic and well-documented (i.e., with thorough metadata).

2. Pre-processing

Pre-processing of the ALS data in the creation of suitable surface models comprises an important part of the production process. As such, it requires a responsible approach and well thought out methods.

The ALS datasets and associated survey control reports provided by TII and others have so far revealed very diverse data quality (acquisition parameters, processing state) and formats.

At the time of preparation of this report ZRC SAZU had not received the majority of the ALS datasets and has therefore thus far focused on the TII Noise Mapping data from 2010–11, and the newer NM20 and Virginia bypass data as typical representatives of lower and higher scanning density data.

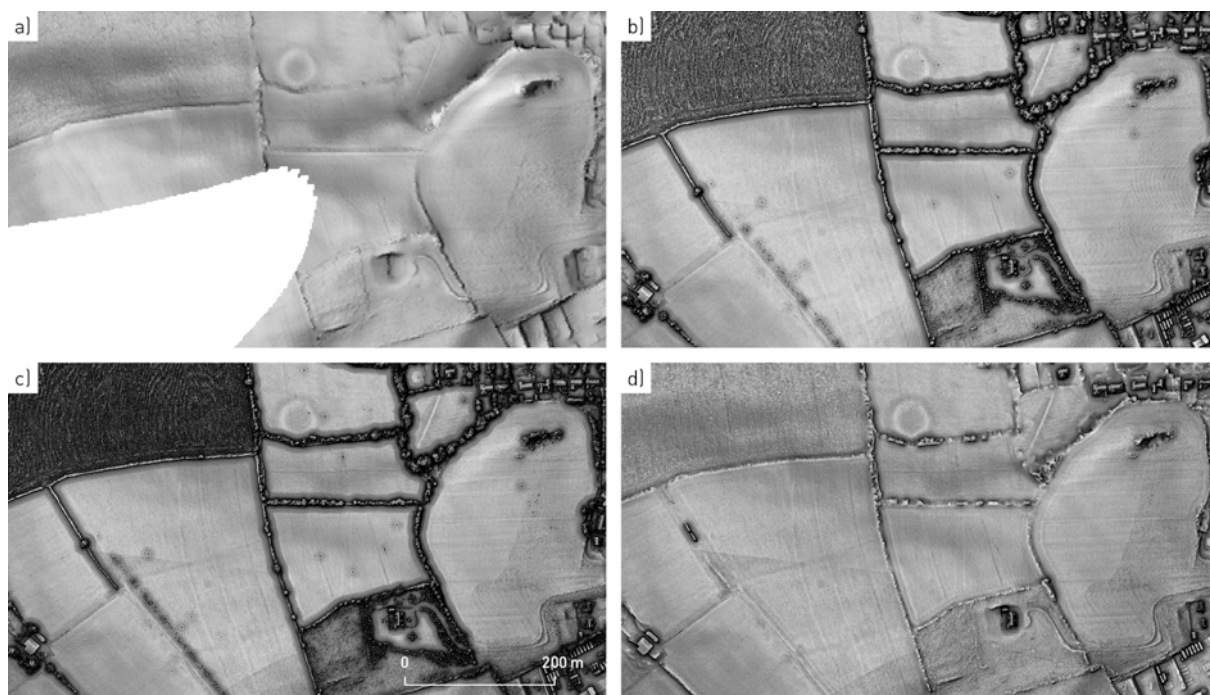


Figure 1: The original 2 m resolution DTM as provided by TII (a). Rasterised data to 1m (b) and 0.5 m resolution (c), and a 0.5 m resolution rasterised classified data (d). The original point cloud is not classified, and the elevation models therefore include all of the above ground objects (b, c). Most of the vegetation is removed with a proper classification and the DEM shows bare earth with buildings (d).

The ALS Noise Mapping ASCII files were rasterised into GeoTIFFs to make them smaller and easier to process. For the same reasons, we also converted point cloud data from ASCII to LAZ, (a dedicated compressed ALS point cloud format). This effectively reduced the amount of data by 95% without any subsequent loss of information. Following this, the scanning density was reviewed to evaluate if the production of higher resolution elevation models would be feasible and identified four things:

1. Scanning density is in places high enough to produce elevation models of up to 0.5 m resolution (Figure 1b, c, d).
2. ALS point clouds do not store classification information, therefore only digital surface models can be produced without classifying whole point clouds first (Figure 1b, c),
3. Point clouds cover wider areas than provided in the gridded files (DTMs) (Figure 1). The data strips are still quite narrow, about 1.5 km across the flight line, which limits the range of visualisations we can use as input to machine learning models. The reason for

this is that most visualisation techniques compute with a certain radius, so medium- and large-scale computations would suffer greatly from the edge effect. Geographical coverage of data stored in point clouds is also larger than in provided gridded data for the NM20 and Virginia bypass data.

- Data does not contain projection information; we assume it is the TM75 / Irish National Grid (EPSG:29903).

Recommendation: We propose to reprocess (classify, grid, visualise) the entire Noise Mapping dataset. Following a case-by-case basis of review of additional ALS datasets we will decide whether to reprocess them, weighing the benefits this might bring against the computational and time resources required.

2.1. DEM Specification

When processing ALS data for archaeological investigations using CNN methods, some fundamental questions arise. Some of the basic questions are:

1. What should be included in the DEM? Should we use a digital terrain model (bare earth) (DTM), a DTM with buildings, or make use of a digital feature model (DFM), which is an archaeology-specific DEM that also includes terrain and buildings but is created in a specialised, computationally intensive way?
2. Should filtering and rasterization (gridding) be done specifically for archaeology? This ensures that most of the archaeology is visible and intact, but on the other hand raises the question of whether everyone can afford both the time and costs required to (re)process data in this way. Point clouds, software, hardware, knowledge, and time should all be available. Such data is never normally delivered to the customer by the ALS data acquisition company; however, this could inform the specifications of any future TII ALS commissions.

Recommendation: Data must be handled more carefully for automatic detection, in comparison to human (visual) inspection which can tolerate systemic irregularities in the data. For example, border edge transitions need to be seamless, and small areas without data need to be interpolated.

Štular and Lozić (2021) recommend a mixture of averaging and segmentation techniques to process the data with a specific archaeological case in mind. The point cloud processing is based on averaging enhanced by sequential processing. First, the buildings are classified with settings optimized for building detection (Figure 2a). Then, the entire point cloud without buildings is reprocessed with settings optimized for ground point detection, which favours the retention of archaeological features over vegetation removal (Figure 2c). An additional step, targeting low-density data, is to classify all unclassified points that are within ± 0.2 m of ground points as ground points (Figure 2d). We refer to this process as classification with 'archaeological settings' or a 'combined approach'. The result is a point cloud with points classified into the following classes:

1. ground (2),
2. low vegetation (between 0.5 and 2.0 m) (3),
3. medium vegetation (between 2.0 and 5.0 m) (4),
4. high vegetation (above 5.0 m) (5),

5. building (6),
6. low point (noise) (7),
7. model key point (if present in the original classification (8), and
8. water (if present in the original classification) (9).

The process requires considerable processing power, as we need to detect ground points, determine their height above the ground, and reclassify them in successive steps, twice. The rasterisation of this point cloud is known as a Digital Feature Model (DFM) (Figure 2d).

If we compare the results of the combined approach with classification using 'mild' settings (Figures 2d, 5b and d), with classification with 'mild' settings (Figures 2c, 5a and c), which itself is the third 'component' of the combined approach, we can see the following:

1. The combined approach removes more of the vegetation and walls.
2. The combined approach classifies points as buildings. In 'mild' settings, buildings are not classified separately because most are included in the ground (2) class. This means that fewer buildings are present in the raster and their shape is not well preserved. With the combined approach, multiple DEMs can be created, both separately for buildings, vegetation, and ground, and for their combinations. The height above ground is also stored in the point clouds for each point.
3. In the combined approach, noise is not removed if the flight lines are not properly aligned as there can easily be 10-15 cm (or more) difference in elevation between the flight lines. This is clearly seen in the meadows on the east side of the areas in Figure 5. This is more of an issue with high density data like NM20, but not so much with lower density data like the Noise Mapping data.

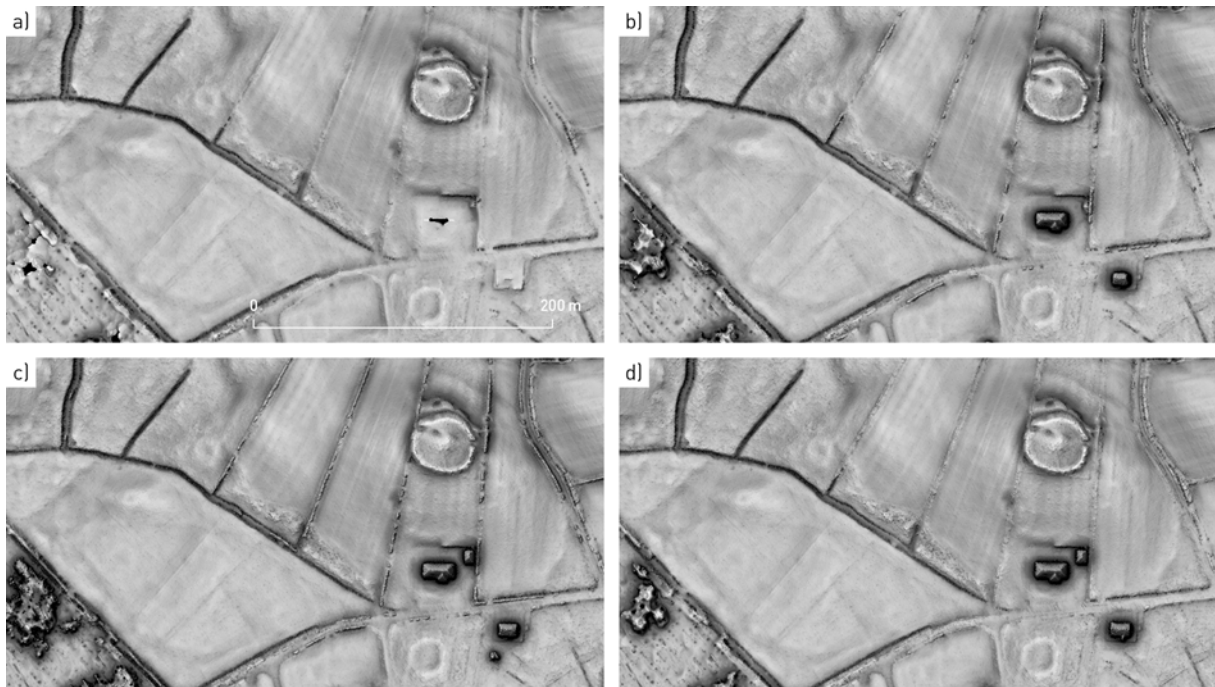


Figure 2: Different algorithm settings for ground classification. 'Aggressive' algorithm settings for removal of all vegetation and large buildings (a), removal of some vegetation and smaller houses (b), classification with 'mild' settings, mindful of archaeology (c), combined (a) and (c) approach (d). Rasterised (gridded) with IDW from Noise Mapping ALS data.

Recommendation: Based on the characteristics of the ALS data provided by TII and obtained from other providers, we propose to create DFMs at 0.5 m resolution.

It has been shown that a model trained by the proposed resolution generalises well to new sites and that higher resolution ALS data improves classification accuracy (Gallwey et al. 2019). These tests also show that although the model was trained at a resolution of 0.5 m, it can be applied at a different resolution without requiring additional training.

The interpolation method used to create a DTM from ALS point cloud data has a direct impact on the resulting terrain surface representation. While there is also a need to investigate how well machine learning models perform and transfer to DTMs created using different interpolation methods, that is outside the scope of this project. Judging from the tests that have been conducted for the visual representation of archaeological features (e.g. Kokalj and Hesse 2017; Štular and Lozić 2020) two methods stand out for various reasons. Gridding with triangulated irregular network (TIN) and inverse distance weighted (IDW) (Figures 3 and 5). Štular and Lozić (2021) recommend combining the two techniques based on a confidence map that is created by incorporating information on slope, ground point density, and low vegetation density (Figure 4). However, the procedure is time consuming and requires point cloud data with classified ground and low vegetation points.

Recommendation: Since much of the Irish landscape is gently sloping and TIN is the preferred method of data rasterization by ALS data providers, we recommend using this method for this project.

Data gridded using TIN are also better suited for some other studies, such as solar energy potential modelling, because the house roofs are more clearly defined and have less noise.

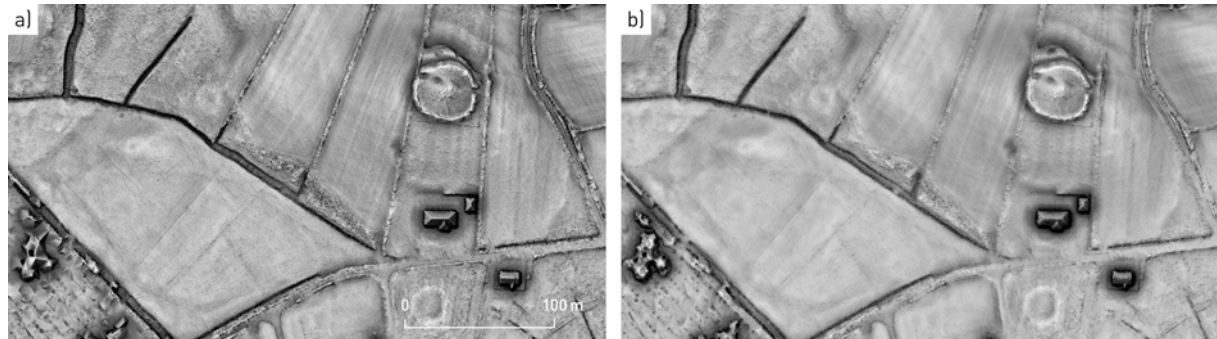


Figure 3: Two approaches to rasterise the DEM were tested. Interpolation with Triangulated Irregular Network (TIN) (a) and Inverse Distance Weighted (IDW) (b). Reprocessed Noise Mapping ALS data.

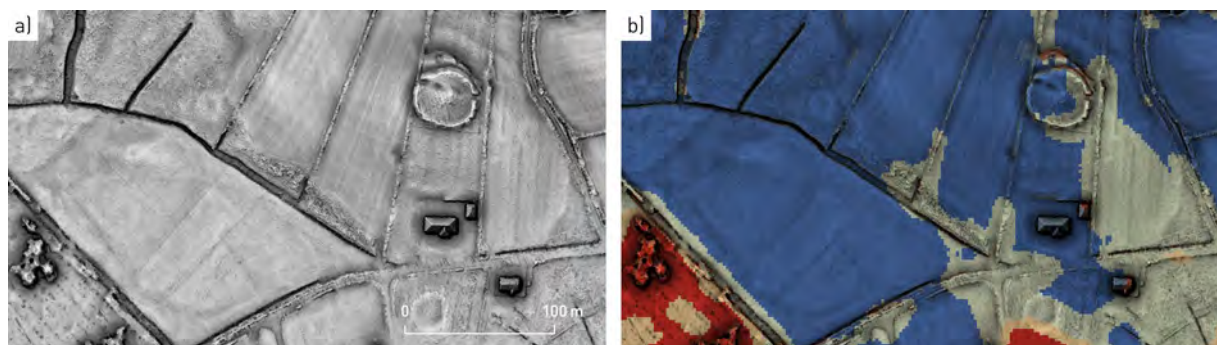


Figure 4: A combination of TIN- and IDW-interpolated rasters (a), based on the confidence map (b) (Štular, Eichert, and Lozić 2021). Reprocessed Noise Mapping ALS data.

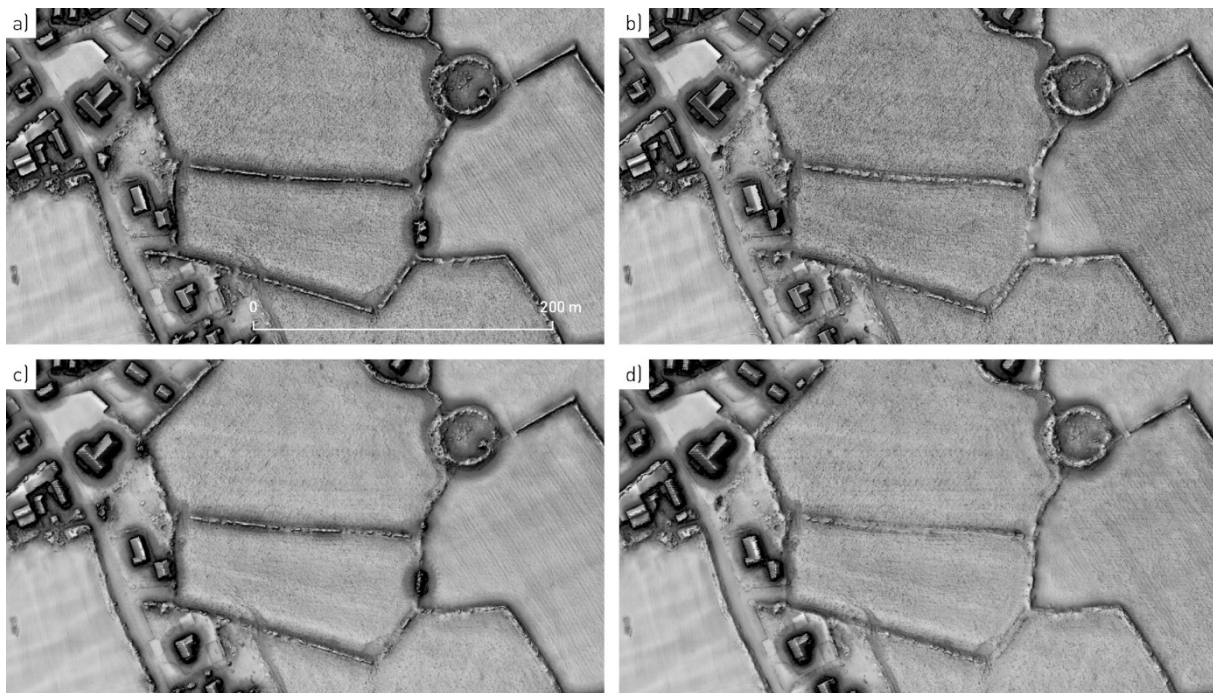


Figure 5: Classification with 'mild' settings (a) and (c), and a combined approach (b) and (d). Rasterised with TIN from Reprocessed NM20 ALS data (a) and (b), and with IDW (c) and (d). Although the combined approach offers some advantages over simple classification mindful of archaeology and seems better suited for lower density datasets such as the Noise Mapping data, it emphasises some data irregularities such as flight line misalignments, as can be seen in the highlighted noise in the eastern-most meadow. This is not so evident in data rasterised with IDW.

2.2. Visualisations

The archaeological assessment reports provided by TII describe various visualisation techniques previously used to visually inspect ALS data. Some used the default settings of the Relief Visualisation Toolbox (RVT) (Kokalj et al. 2019) while others fine-tuned the settings to further enhance the visibility of objects of interest. Typical adjustments include lowering the sun elevation angle to calculate hillshading, custom stretching of the histogram, inversion of the colour scale, and vertical exaggeration. Combinations of visualisations are often used (e.g., of simplified local relief model (SLRM), multidirectional hillshade, sky-view factor (Figure 6e and f)).

The most commonly used visualisations are described below. Unless otherwise noted, the RVT default settings are used in the calculations.

1. *Relief shading* (also known as hillshading or shaded relief) provides the most 'natural', i.e., intuitively readable, visual impression of all techniques. It is easy to compute and straightforward to interpret even by non-experts without training and is a wonderful exploratory technique (Figure 6a). Direct illumination restricts the visualisation in dark shades and brightly lit areas, where no or very little detail can be perceived. A single light beam also fails to reveal linear structures that lie along it, which can be problematic in some applications, especially in archaeology, or can completely change (inverse) the appearance of the landscape (i.e., false topographic perception). Typical settings used: Sun elevation angle set to 10°.
2. *Hillshading from multiple directions* (Multi-HS). Note that many reports claim of using 16-direction hillshade. It is unclear whether this is true because only three directions are shown in a GIS environment and switching between layers (directions) is

cumbersome. By experience, only the 8-bit version is often used, and this is computed with three predetermined directions in the RGB bands. This might in fact even be beneficial because hillshadings from nearby angles are highly correlated and the predetermined directions (azimuth 315° for red, 22.5° for green, and 90° for the blue band) reduce data volume and are optimised for best visual understanding. The convexities and concavities, as well as orientations, are clearly and unambiguously depicted (Figure 6b). Typical settings used: Sun elevation angle set to 20°, vertical exaggeration 3, although experience within Irish landscapes suggests a vertical exaggeration of approximately 1.6 provides the optimal visualisation outputs.

3. *Sky-view factor (SVF)* is a geophysical parameter that measures the portion of the sky visible from a certain point (Zakšek, Oštir, and Kokalj 2011). Locally flat terrain, ridges, and earthworks (e.g., building walls, cultivation ridges, burial mounds) which receive more illumination are highlighted and appear in light to white colours, while depressions (e.g., trenches, moats, plough furrows, mining pits) are dark because they receive less illumination. It is ineffectual in the case of low, subtle earthworks as the variation in how much sky is visible is negligible (Figure 6c). Typical settings used: histogram stretch to minimum/maximum between 0.94–0.99.
4. *Simple Local Relief Model (SLRM)* removes the large-scale morphological elements (hills, valleys...) from data so only small-scale features remain (e.g., archaeology). In the simplified process a simple mean filter computes the trend and a trend removed model is produced directly by subtracting the filtered model from the original (Figure 6d).

Additional visualisations, used in machine learning literature include:

5. *Slope severity* (gradient) is the first derivative of a DEM and is aspect independent. It represents the maximum rate of change between each cell and its neighbours. If presented in an inverted grayscale (steep slopes are darker), slope severity retains a very plastic representation of morphology (Figure 7a). However, additional information is needed to distinguish between positive/convex (e.g., banks) and negative/concave (e.g., ditches) features because slopes of the same gradient (regardless of rising or falling) are presented with the same colour.
6. *Local dominance (LD)* shows how dominant an observer standing above a certain pixel would be for a local surrounding area (Kokalj and Hesse 2017). Dominance is higher for points on local protrusions as well as on slopes, and lower for points in local depressions (Figure 7b). It is well suited for very subtle positive relief features such as former field boundaries or heavily eroded burial mounds, but also delivers very good results for topographic depressions such as dolines, mining traces, and hollow ways.
7. *Topographic openness* is a proxy for diffuse relief illumination (Yokoyama, Shlrasawa, and Pike 2002). Because it considers the whole sphere for calculation, the result is an image devoid of general topography—a kind of trend-removed image. It has the same valuable properties for visualisation as sky-view factor with the added benefit of more pronounced local terrain features, although the visual impression of the general topography is lost. Topographic openness can be positive (Figure 7c) or negative (Figure 7d) and their combination can also be used to enhance visibility of small, local topographic variations.

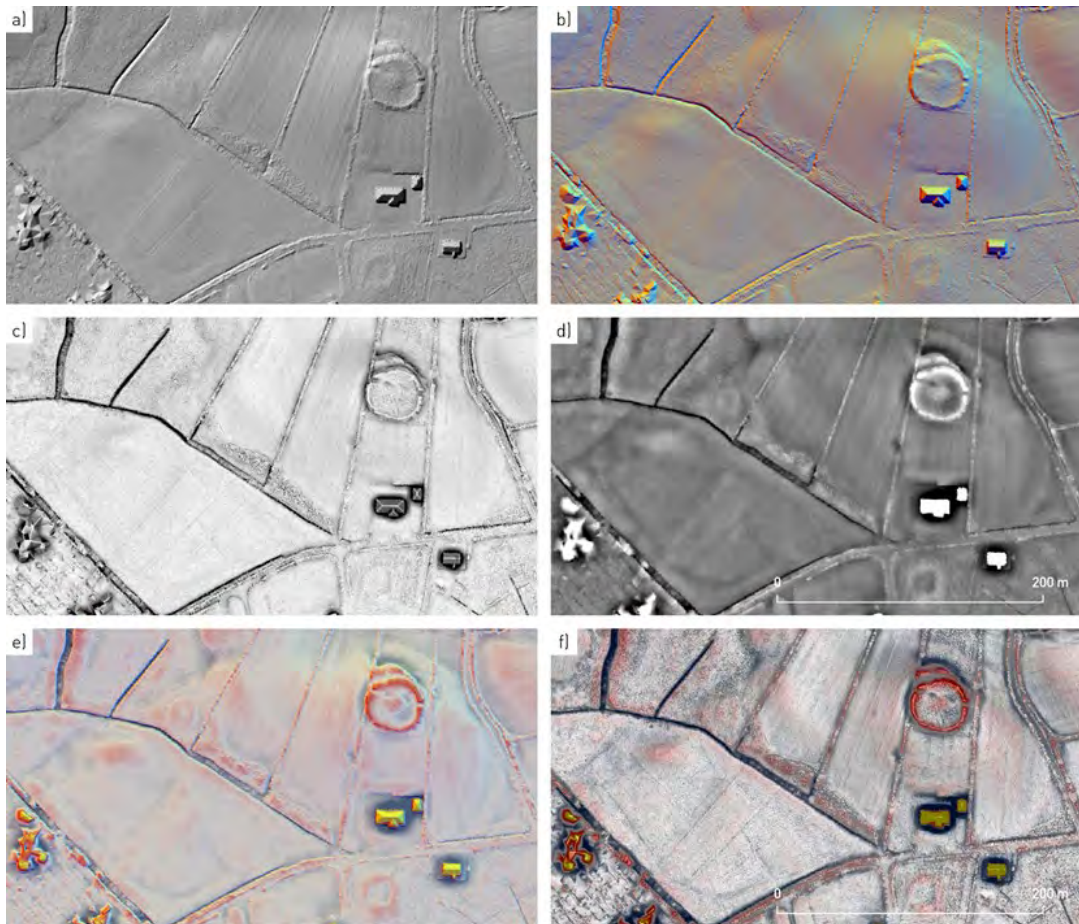


Figure 6: Visualisations commonly reported in archaeological assessment reports provided by TII. Relief shading (a), multidirectional hillshading (b), sky-view factor (c), simple local relief model (SLRM) (d), multidirectional hillshade overlaid by SLRM (e), SVF overlaid by SLRM (f).

8. *Visualisation for Archaeological Topography (VAT)* is based on a fusion of four information layers into a single image (hillshade, slope, positive openness, and sky-view factor). VAT can be used to enhance visibility of features of varied scale, height, orientation, and form that sit on terrain ranging from extremely flat to very steep (Figure 7f) (Kokalj and Somrak 2019; Somrak, Džeroski, and Kokalj 2020).
9. *Red Relief Image Map (RRIM)* is often used for mapping mass movement. It overlays a slope gradient image, coloured in white to red tones, with the 'ridge and valley index' computed from positive and negative openness in a grayscale colour map (Figure 7f) (Chiba, Kaneta, and Suzuki 2008; Asia Air Survey 2019).
10. *Multiscale Relief Model (MSRM)* was developed to extract microtopographic information at various scales employing micro-, meso- and large-scale DEMs (Orengo and Petrie 2018; Berganzo-Besga et al. 2021). Multiple low pass filters with different kernel radii are applied to the original DEM. The algorithm subtracts consecutive low pass filters between them and calculates the mean of these subtracted reliefs.
11. *Multiscale Topographic Position (MSTP)* is generated from a hyperscale datacube of the topographic metric deviation from mean elevation and reduced to three dimensions by extracting the absolute maximum value from micro, meso, and macro scale ranges (Figure 7g) (Guyot et al. 2021).

12. *Enhanced Multiscale Topographic Position version 2 (e^2MSTP)* is based on fusion of multiscale topographic position, red relief image map, and simplified relief model (Figure 7h). It is slope invariant and the best technique for CNN tested by Guyot et al. (2021).

Other promising detrending and orientation invariant techniques have been proposed (Toumazet, Simon, and Mayoral 2021) and appear promising as input to machine learning, but have not yet been tested.

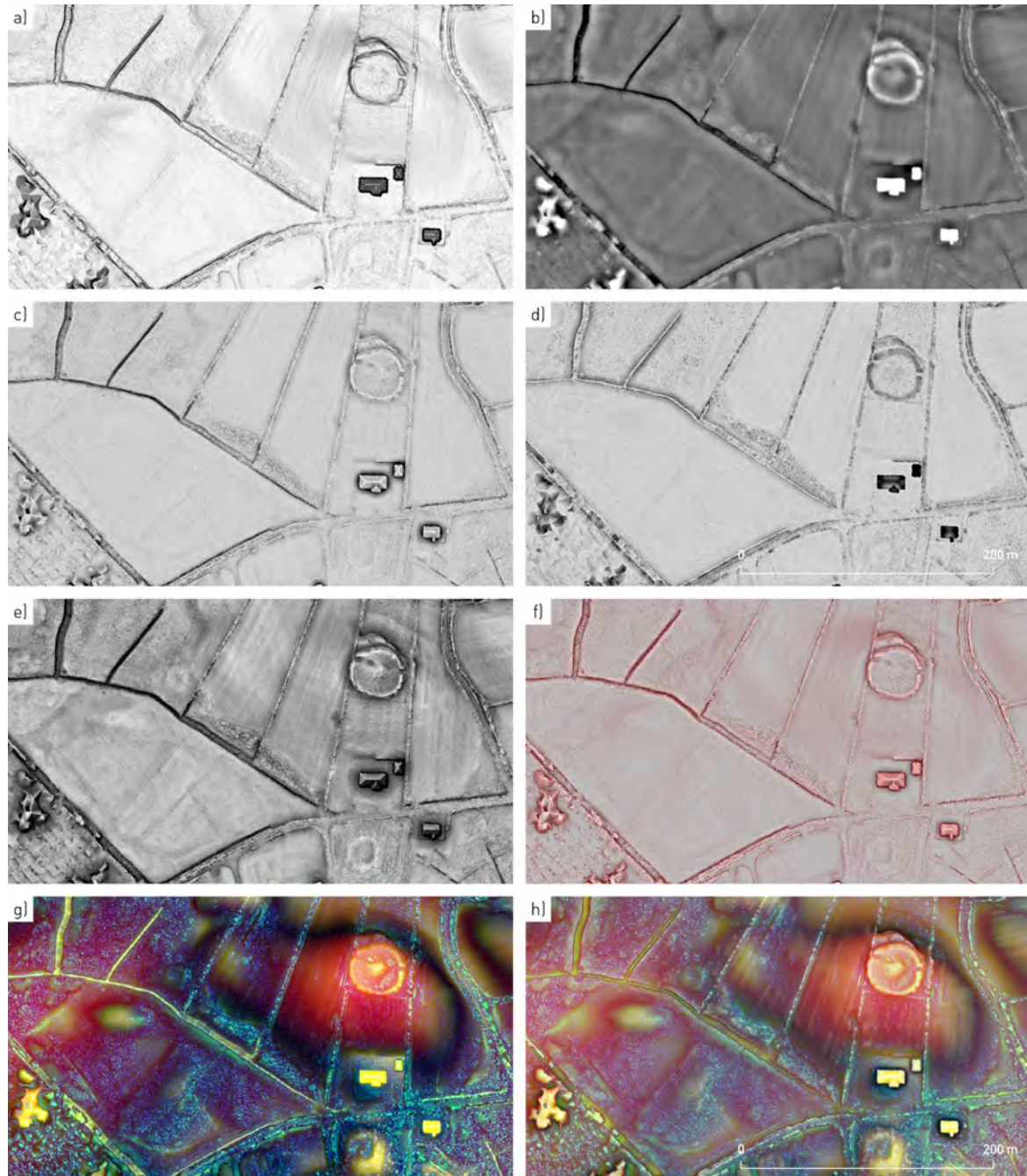


Figure 7: Visualisations commonly reported in papers dealing with archaeological exploration with deep learning CNN models. Slope (a), local dominance (b), positive openness (c), negative openness (d), visualisation of archaeological topography combining settings for flat and general terrain (e), red relief image map (f), multiscale topographic position (g), multiscale topographic position enhanced version 2 (h). SLRM is often cited as well (Figure 6d).

DEM versus visualised ALS data

Verschoof et al. (2021, 9) observe that the better performance of their model in detecting hollow roads on DTM data compared to visualised data is due to information loss as proposed by Shannon's channel coding theorem (Shannon 1949). This theorem from information theory states that any processing of information from one representation to another reduces (e.g., due to rounding errors) but never improves the information content. We believe that the conclusion of Verschoof et al. that the DTM contains more informative, less noisy data compared to the visualised data may be true, but this does not necessarily mean that the CNN is able to make this information useful. As demonstrated in several studies (see e.g. Shorten and Khoshgoftaar 2019; Chen et al. 2020), data augmentation techniques, such as histogram stretching, generally improve the performance of a ML model. Verschoof et al. also use histogram stretching to enhance their DTM data (as do Bundzel et al. (2020) and Gallwey et al. (2019)), but do not do so for hillshading, the visualisation technique with which they compared against. In addition, ALS data are inherently noisy due to processing artefacts (e.g., resulting from misaligned flight strips or acquisition errors), vegetation cover, and decay of archaeological structures. Adding noise is also one of the data augmentation techniques commonly used with remote sensing data (Stivaktakis, Tsagkatakis, and Tsakalides 2019; Abdelhack 2020). Finally, hillshading is one of the visualisation techniques that really does suffer from information loss or at least has serious drawbacks compared to other visualisations (see e.g. Kokalj and Hesse 2017, 16–17). Gallwey et al. (2019) also report better performance of 'raw 8-bit DSM height values' instead of SLRM, positive or negative openness in their test, but the DSM height values were indeed linearly stretched for all tiles individually to increase contrast. The authors suspect that while the visualisations are effective for human interpretation of archaeological data and also useful for more traditional machine learning techniques, it is not desirable to artificially alter the data representation prior to input because deep CNNs learn their own feature representations during training. This is in direct contrast to the findings of Guyot et al. (2021) who found that visualisations that are best suited for visual interpretation also work best as input to their CNN model. We believe that Gallwey et al.'s finding that their results were biased toward DSM is true because their CNN was pre-trained on 8-bit DSM height values. They conducted experiments to build a model from scratch with different visualisations. However, no meaningful results were obtained with any visualisation, probably due to the limited size of the training data. In order to comprehensively test which LiDAR visualisation is best suited for CNNs in the future, a robust CNN trained from scratch with multiple differently visualised representations of the same data would be required; however, such a model has not yet been made publicly available by any known source.

Recommendation: Although the idea of using a DTM as a direct input to the ML process is worth pursuing, we believe that further studies are needed to prove its effectiveness against using advanced visualisation techniques. We suggest testing this within the project before a list of proposed inputs is finalised.

Assessment of visualisations for CNN

Somrak and Kokalj (2020) tested several visualisations for classification of buildings, platforms and aguadas. They found that the models gave the best results with multiple visualizations in separate channels (that is, without combining them first), and when a smaller buffer around the objects was given (2 px rather than 15 px, except for visualizations that are computed with a wider convolutional filter (LD, flat VAT)). In their assessment of visualisations for visual (human) inspection, Guyot et al. (2021) report that the combination of multiscale representation and local morphological information (e^2 MSTP) allowed for better interpretation

of structured terrain; for example, terrain with pits, narrow ditches or embankments. Detrending techniques (local dominance and SLRM) were highly informative for small-scale terrain variations; however, their limits were apparent for the following geoarchaeological configurations:

1. small archaeological structures (smaller than the radius of analysis) within highly textured terrain (usually in undergrowth vegetation) and,
2. large archaeological structures (larger than the radius of analysis) considered to be the natural trend of the terrain and thus not enhanced.

Terrain openness and its variants (positive and negative openness, SVF, RRIM) appeared to be effective for small archaeological structures, especially for small mounds (<10 m radius) with a central pit; however, the lack of overall contextual information (such as multiscale topographic position) reduced the ability to interpret these structures. However, these visualisations were not visually informative for continuous and subtle variations that occur at a larger scale with the presence of levelled tumulus. This limitation was also observed for standard visualisations, such as slope, hillshading, and, to a lesser extent, PCA of hillshadings, as well as for the combined visualisation of archaeological topography (VAT and its 3-band variation VAT-HS channels), because it was designed to improve the recognition of small topographical features (Kokalj and Somrak 2019). While they found VAT informative for the visual description of small structures or local morphological characteristics, the results they obtained showed that it was not adapted for the visualisation of larger subtle topographic variations or for specific topographic positions. The performance (AP@IoU0.5) they achieved with machine learning ranged from a minimum average of 24% (analytical hillshading) to a maximum average of 65% (e²MSTP), indicating that the choice of visualisation has a fundamental impact on performance. Only the multiscale approaches, MSTP (58%) and e²MSTP (65%), obtained performances that exceeded 50%.

The best-performing visualisation was the e²MSTP generated using multiscale topographic analysis (MSTP) combined (via a blending technique) with a morphologically explicit image (RRIM) and a slope-invariant relief detrending technique (SLRM). This visualisation was created based on empirical knowledge and iterative selection of parameters that enhance the perception of ALS-derived terrain data for archaeological prospection. The computer-vision approach shows the suitability of such visualisations for identifying archaeological structures on ALS-derived terrain models in the study area. In particular, it highlights the utility of using a multiscale approach that provides contextual topographic position information and is more robust for varying structure size. It also confirms the advantage of combining complementary visualisations to address identified limitations of single visualisations (e.g. blending morphological and detrended information with multiscale information) (Guyot, Lennon, and Hubert-Moy 2021). It should be noted, however, that the morphology of the archaeological structures in the test was mostly upstanding and very subtle, and the authors themselves warn that further testing is needed. A similar approach applied to other contexts would not necessarily yield the same performance from a deep CNN perspective. For example, Suh et al. (2021) report that VAT performed best of the visualisations tested in a large-scale study. All new tests require new training and evaluation, which may not result in the same ranking of visualisations, especially if structures or landforms differ. However, Guyot et al. (2021) expect that ranking of visualisations should not change. If one visualisation works better for visual inspection than another, then it should also work better for automatic detection with CNN. Kazimi et al. (2020) report that the best performing visualisations for detection of bomb craters,

charcoal kilns, barrows and mining holes were local dominance and SLRM, with unprocessed DTM not far behind. They also report that the true positive rate for all the data types are similar, but there are fewer false positives in the case of SLRM and LD than with other visualisations. They also urge for further experiments to study the effect of different DTM derivatives for detection of other object categories and those in different regions, different data sources, and data resolutions.

Recommendation: Further testing is needed to determine which visualisation(s) produce the best results with Irish topography and selected archaeological types. We will also conduct tests with a new, customised visualisation that highlights local topographic variations, provides multiscale information, and is directionally invariant. Such a visualisation will be suitable for a wide range of data augmentation methods.

2.3. Learning dataset creation

In deep CNN, classification predicts the class of the whole image (patch, tile, chip) being considered (Figure 8a), object detection predicts the presence and location (surrounded by a bounding box) of an object in the image (Figure 8b), while instance segmentation adds a contour (also called a 'mask' or a 'segment') to the detected object (Figure 8c).

The nature of the training set largely influences the CNN method and its output. For example, the balancing of rapid coverage and detail of information means that archaeological (heritage management) records such as the Irish SMR often contain locations of objects indicated only by single point data, without specific details of at least the size of the recorded object. This makes it difficult to construct appropriately sized bounding boxes (Figure 8b) and more so masks for object detection and semantic segmentation (Figure 8c), introducing additional uncertainty in the machine learning model.

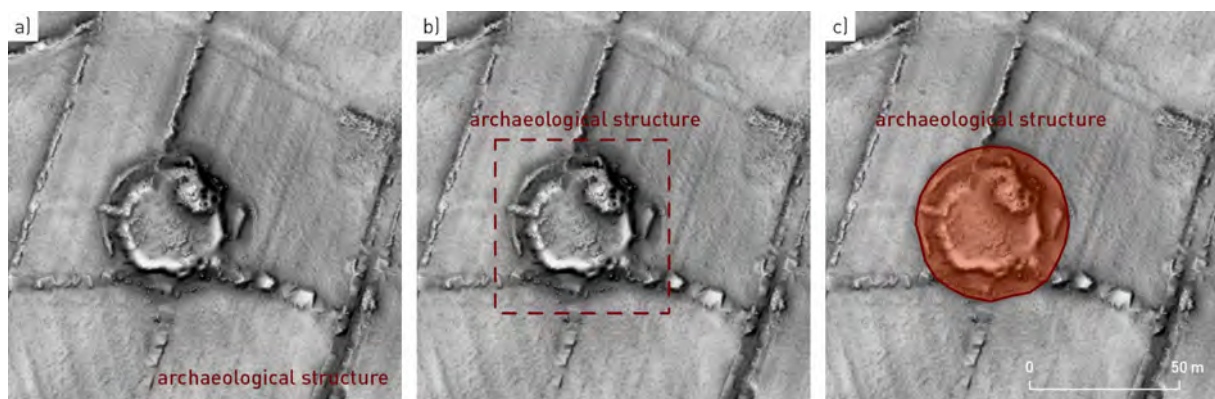


Figure 8: Image analysis using deep learning Convolutional Neural Networks for an archaeological site (a curvilinear enclosure). Image classification: a class or label associated with the image (a); object detection: a labelled bounding box that locates the object of interest within the image (b), and object segmentation: a labelled footprint that locates and delineates the object of interest within the image (c). Reprocessed TII Noise Mapping data. The bounding box in (a) and the mask in (b) are hand drawn.

The data examples (a dataset) used for machine learning are divided (split) into three (sub)(data)sets: the training (learning) set, the validation set, and the test set. The terms test and validation dataset are used interchangeably in the literature and their meaning can only be inferred from context. Both datasets are samples of data retained and excluded from training the CNN model. We propose to use the definitions provided by Ripley (1996, 354) in his seminal work:

1. *Training set*: A set of examples used to fit the model.
2. *Validation set*: The sample of data used to provide an unbiased evaluation of a model fit on the training set while tuning model hyperparameters. The evaluation becomes more biased as skill on the validation dataset is incorporated into the model configuration.
3. *Test set*: A set of examples used only to provide an unbiased evaluation of a final model.

Each experimental configuration must be checked to ensure that no contamination occurs between validation, test and training datasets. The full geographical separation of the training and testing areas enables easier construction of the training set (Simidjievski et al. 2021) and 1:1 division gives more confidence into the generality of results (Bundzel et al. 2020).

Guyot et al. (2021) notes that the availability of archaeological reference datasets remains a key issue. The collection, publication and maintenance of labelled archaeological data is not straightforward. Available inventories are gradually addressing this issue, but the limits of large-scale archaeological references persist: many sites remain to be discovered, and for many of them, the nature and spatial extent of the archaeological structures could, paradoxically, be confirmed only by destructive archaeological excavation. Manual analysis is biased towards the expectations, experience, knowledge and observational abilities of the interpreter(s), with the risk of missing or dismissing objects (Halliday 2013). The 'ground truthing', which serves as an essential base for all supervised remote-sensing classification or detection approaches, would remain wishful thinking in the archaeological prospection domain. Thus, there is a need for development of new paradigms based on fuzzy or partial reference datasets.

Bundzel et al. (2020, 18) provide some suggestions for archaeological labelling of lidar data that is to be used in machine learning:

1. If the labelling was done without ground-truthing, ideally, several experts should evaluate each object and indicate the confidence in their decision. The confidence level is a valuable information when calculating CNN loss.
2. If presence of an object was verified by ground-truthing, this should be indicated.
3. When labelling objects of ancient construction activity, the area covering the present-day features belonging to the object (its eroded material) should also be indicated.
4. When labelling objects such as looting trenches, agricultural fields, etc., the entire area of the objects' features should be indicated (for example the trench and the debris hill).

These recommendations are also reflected in suggestions by Cowley (Cowley 2016; Cowley et al. 2021) and Banaszek et al. (2018).

Recommendation: The sample dataset (i.e. training, validation and test datasets combined) should contain only archaeological structures whose topographic characteristics are identifiable on the DTM. Sites that are considered above ground structures with no influence on the bare-earth topography, such as standing stones, or sites that are associated related only to small-object deposits should be excluded.

The number and characteristics of image patches selected for the training and test datasets have an important influence on the machine learning model and therefore its selection for the particular task. It is generally known that the number of training samples required to successfully train deep CNNs has to be large, but just how large depends on the exact method, architecture, and configuration of the particular CNN. Mask R-CNN-like instance segmentation

involves a large and complex network that requires an abundance of data to be trained on, while U-net can generally perform well with less training samples. What is more concerning is that studies (e.g. Trier, Cowley, and Waldeland 2018; Verschoof-van der Vaart et al. 2020; Guyot et al. 2021) report the sensitivity of a model to the images selected for the training and test datasets. The performance metric can vary greatly for the same number of training images, depending on which particular images were used to train the model. Some of the variability is related to the metrics used to evaluate detection and segmentation performances, but the main sources of variability seem to be the images selected for model evaluation (the complexity of the test dataset) and training (whether the training dataset is representative and comprehensive) (Soroush et al. 2020). This is very much related to the issue of (very) small archaeological training sets.

Recommendation: The decision on the exact procedure to construct the training, validation and test sets will be made after the review of these reports by the whole project team and the TII and detailed in the Final technical Report.

2.4. Background (negative) examples

Trier et al. (2018, 10) observe that further consideration of the character of learning sets in specific implementations is crucial. The identification of objects of confusion appears to be an important consideration in improving outputs. Adding such false positives to the learning set in confusion classes has the added benefit of taking a more holistic approach to the topography of the landscape and creating more explicit understandings of how the archaeological elements of the landscape are defined and identified. The main purpose of adding confusion classes is to assign non-archaeological structures to non-archaeological classes, thus reducing the number of false positive identifications of the archaeological structures being sought. However, there is also a risk that true archaeological structures may, in some cases, be mistaken as one of the confusion classes. Adding confusion classes (providing the model with negative/background examples during training) is also supported by Somrak et al. (2020). Guyot et al. (2021) report this should be addressed because in their study, which used a single 'archaeological' class, consisting mostly of barrows, expected prediction errors (e.g. a roundabout) were accompanied by errors for objects sharing few or no similarities with the reference dataset (e.g. a horse training area, a large ditch). One such study by Verschoof et al. (2020) that dealt primarily with similar objects (barrows) reports no improvement in performance when the CNN was also trained with negative examples.

Trier et al. (2018) argue that the number of background examples should be large for three reasons:

1. 'Background' is the most frequent situation in the landscape.
2. There are many different natural terrain structures.
3. A large range of background examples may reduce the number of false positives.

However, if the number of background examples is too large, the training phase becomes very slow, precluding experimentation with various parameter settings by running several training phases and comparing the results.

Recommendation: The CNN testing should be carried out which include negative examples.

2.5. Data augmentation

There are many data augmentation techniques, which can be used, but we have to be careful using them, because we have to be mindful of the content and nature of images. For example, colour transformations are not desirable as they can completely change the meaning, and neither are rotations or flipping for directional visualisations, such as hillshading. Guyot et al. (2021, 5) used limited affine transformations:

1. scaling: 80 to 120% of the original image size,
2. translation: -20% to 20% of the original image position,
3. rotation: -25° to 25° of the original image orientation.

These transformations were defined within limited ranges of scaling, translation and rotation to avoid unrealistic versions of the used visualisations. The augmentation process was applied 50% of the time to ensure that the deep CNN received both augmented and non-augmented versions of the training dataset. Maxwell et al. (2020) implemented random augmentations of the original image chips (square class samples from the source image) including rotations at 90°, 180°, and 270°, horizontal and vertical flip, brightness and contrast alterations, and blurring and sharpening. They too attempted to avoid extreme augmentations of the original data. Bonhage et al. (2021, 5) report that pre-processing steps (contrast enhancement and data augmentation) increased the mean average precision (mAP) of the model by approximately 4% on average.

Recommendation: We propose conducting tests that include a wide range of data augmentation techniques including at least rotation, horizontal and vertical flipping, scaling, blurring and sharpening, brightness and contrast alterations, translation, truncation, and noise injection.

3. Detection

Multiscale approaches may be required if the archaeological structures are of different scales, requiring DTM of various resolutions (e.g. 0.25 m and 2.0 m), and/or neural networks with different input sizes (e.g. 96 × 96 pixels and 512 × 512 pixels) (Trier, Cowley, and Waldeland 2018, 20). Objects with lower certainty levels have to be labelled in order to increase the recall rate. The uncertainties, bias, and error during the labelling process will affect the model performance.

High metrics can also be achieved for models that are trained for a small task with no outside application. When models are trained for a complex task across a big region, better performance is achieved if the certainty measures are relaxed to increase recall, and quick post-processing steps are integrated to increase precision (Soroush et al. 2020).

Measures to prevent a model from overfitting (where the statistical model fits exactly against its training data thus preventing accurate matching against unseen data) the training set should be implemented (Hinton et al. 2012). For this purpose, a dropout layer was inserted towards the end part of the network. In the dropout layer, neurons and their connections are randomly dropped out (50% is a common practice). This regularization method prevents the network from memorization of training samples and instead encourages the network to learn more general representations. Generalization capabilities are thus improved—leading to better

model performance on unseen data—resulting in the final model being more robust (Srivastava et al. 2014).

Maxwell et al. (2020) circumvent poor predictions near the edge of image chips by making predictions on image chips twice the size of the chips in the train set, i.e. 1024 px × 1024 px instead of 512 px × 512 px, with a stride of 256, allowing for substantial overlap (each pixel is evaluated 8 times). In this way it is less likely that features are missed and more likely that they appear near the centre of at least one chip. The results are cropped so that only the centre 50% of each chip is used. Within this project where the DFM model resolution is 0.5m the image chip dimensions will be 512 px x 512 px which will capture objects which are approximately less than 125m in diameter.

3.1. Pre-training (and transfer learning)

Pires de Lima and Marfurt (2020) present a systematic review of transfer learning and their results show that transfer learning is a powerful tool for remote sensing. Furthermore, they show that models pre-trained on more generic natural images outperform those pre-trained on domain-specific datasets. This is not supported by many other studies who wish that ideally, in the future, large tagged training sets of archaeological images can be used for pre-training, rather than hoping that common processes of detection learned on generic photographs from databases such as MS COCO (Lin et al. 2014) and ImageNet (Deng et al. 2009) may be applicable to specifically archaeological images (Trier, Cowley, and Waldeland 2018; Maxwell, Pourmohammadi, and Poyner 2020).

Gallwey et al. (2019) successfully used a fixed network (DeepMoon), pre-trained on very similar data (DTM) and objects, but report bias and lower performance with other input success (visualisations).

Training from scratch by initializing the model with random weights is only reported as successful for the U-Net CNN (Küçükdemirci et al. 2021; Seitsonen and Ikäheimo 2021; Soroush et al. 2020; Suh et al. 2021), while it is unsuccessful for more complex configurations such as the Mask R-CNN because there are not enough training examples available (Bundzel et al. 2020; Maxwell, Pourmohammadi, and Poyner 2020).

Recommendation: Pre-training requires further exploration especially as archaeological learning sets are small in comparison to other training sets.

3.2. Frameworks

Garcia-Molsosa et al. (2021) give several reasons why they used an existing online machine learning platform instead of developing their own open access detectors:

1. Automatic pre-processing steps that include data augmentation.
2. The training of the models uses cloud-based distributed computing, which greatly speeds the process.
3. Proprietary postprocessing techniques separate the per-pixel classification results into separate objects.
4. The training and testing of detectors can be achieved very quickly without the need to gather large amounts of annotated images.

5. The research aim was to test the potential of deep learning for the detection of multiple objects, and therefore, a fast and efficient method allowed them to experiment until an adequate detector for each object was achieved.
6. The symbology and representation of archaeological objects can vary greatly between different datasets, meaning other researchers will have to train their own algorithm that fits the specific object and the data they are using.

But they warn that the methods to train and run the models cannot be shared given the use of a proprietary cloud computing platform, which is in contrast to the demand for open science which may limit its application within this project.

3.3. Object detection methods

In the literature, the recognition of relatively simple features, such as various mounds and pits, predominates, while more complex, large-scale landscape patterns, such as historic field systems and roads, are rarely used in archaeology (Traviglia and Torsello 2017), with recent studies trying to balance this (Verschoof-van der Vaart and Landauer 2021; Verschoof-van der Vaart and Lambers 2021; Somrak, Džeroski, and Kokalj 2020). The nature (size, linearity...) of the archaeological objects influences the decision on which type of machine learning method to choose and what size the input tiles should be. Kazimi et al. (2018) report that in general, as the tile size grows, there is a decrease in prediction accuracy for linear structures like tracks and streams, but an increase in prediction accuracy for areal structures (lakes) for all tested models. They tested classification on patch sizes from $8 \times 8 \text{ m}^2$ to $100 \times 100 \text{ m}^2$. Similarly, Verschoof and Landauer (2021) feed the CNN with small, individual sections of hollow roads, as opposed to whole sections used as input images. This is because the variation between small sections is minimal and has several advantages:

1. The use of sections (and therefore multiple samples per single hollow road) makes it much more cost-effective to create a sufficient training dataset for deep learning algorithms; and the training of the models uses cloud-based distributed computing, which greatly speeds the process.
2. by using this method, a CNN is only used for image classification, a relatively simple task, instead of more complicated tasks such as segmentation.

Another benefit of using classification over segmentation is that some objects are difficult to define on a pixel level because of their heterogeneous nature, especially if their traces have been 'transformed' by various natural and anthropogenic processes overtime. Therefore, using a CNN for classification will produce better detection results at a lower cost/effort.

The number of iterations per epoch (the number of times a learning algorithm sees all the training examples). has to be dynamically adjusted to the number of training images available at the beginning of each experiment (assuming a batch size of 1 and 1 image per GPU). This configuration ensures that the deep CNN observes each image (or its augmented version) only once per epoch.

3.3.1. Challenges and issues with using machine learning in archaeology

The need for using ML in archaeology is ever-growing with the rapid developments in image analysis techniques – the development of different methods from ML has exploded and penetrated all areas of life: science, technology, multimedia and business. Considering the wealth of data collected in archaeology, it seems that using these data to learn ML models is a match-made-in-heaven.

In order to be able to exploit the power of ML, one needs to first prepare the data to be ready for ML exploitation. This means that the data selection and data preparation stages have been performed and the outcome is a dataset representing the problem under investigation: data were cleaned of processing errors, selection of study area and unit of analysis, selection of the main question addressed with the data (e.g., localization of specific objects at a given study site), removing data biases, definition of evaluation strategy (training, validation, and testing parts of the data), and selection of quality criteria for the models (i.e., how will the quality of learned models be judged – evaluation measures).

The process of digitalisation of archaeology is fairly recent. This was mainly executed through the object-based image analysis paradigm (OBIA). Reviews on the subject are provided in (Davis, 2018), (Davis, Sanger and Lipo, 2018), (Magnini and Bettineschi, 2019) and (Olivier and Verschoof-van der Vaart, 2021). The first and foremost question is dealing with the need to progress the digitalisation of archaeology with automatic analysis methods. The answer is very much related with the need to speed-up the examination of the large amounts of gathered data and process data captured within much larger study areas as compared to traditional archaeological practices. Moreover, this is also related to the provision of improved reproducibility when carrying out image analysis processing. Davis (2018) provides a very comprehensive table of the developments in the adoption of OBIA in archaeology, which, for completeness, we include here as Table 1.

Several challenges and practical issues for advancing the agenda of digitalisation (as well as machine learning) in archaeology have been identified. These include:

- User-friendliness of the software solutions both commercial and free, open-source software. Having the user experience in mind when designing software solutions for specific domains is crucial for adoption. This is even more relevant for software which includes ML. In this project, we will strive to minimize the adoption barrier by enhancing the user experience.
- Segmentation, a process which is highly dependent on the domain expert/user, that is controlled by the scale parameter and is dependent on the: data source, the aim of the study, the geographical context, the surrounding contrast and the internal heterogeneity. Generally speaking, over-segmentation should be preferred over under-segmentation as this is necessary to grant the necessary detail to traces of features which appear secondary in the given context (e.g., evanescent crop-marks within a network of agricultural channels and rural houses).
- Incompleteness of the residual record, i.e., quantification of the post-depositional process, which alter the original characteristics of the different sites to a different degree. This leads to the problem of high variability within the same class of objects of interest, hence, making it more difficult for the machine learning models to correctly recognize and classify them.

- Archaeological sites are often associated within a complex palimpsest of structures from different time periods coming from both natural and, or anthropological origin. This makes these sites more difficult to recognize and classify not just by machine learning methods but also by human experts.
- The barrier of meaning (Casini et al. 2021) states that while the traits leading to decision and identification of archaeological sites used by domain experts (typically learned through years of experiences) are not necessarily the same traits predominantly employed by the ML methods to make a decision. More to the point, rarely do ML methods provide direct explanations as to why a given prediction is made (with the rare exceptions, e.g., decision trees). In the case of deep learning, there are separate methods that study the behaviour of the models and their decision-making process. More generally speaking, the area of explainable artificial intelligence studies this in more detail.
- The various forms of bias in the collected data might hinder the performance of the models (Casini et al., 2021). These biases include: 1) "lumping an assemblage into previously determined categories, meaning that rare and unusual objects can easily be missed by being grouped with a more common type"; 2) "models cannot fully incorporate the variability of the features being classified. ML analyses are susceptible to missing the "forest for the trees" because the data used to train the models are often stripped of contextual information (especially in the case of images) or operate on a limited set of prechosen variables that may not include sufficient information to distinguish between important (that is, archaeologically relevant) classes."; and 3) inclusion of non-objective user preferences during data collection (e.g., focusing on a specific subtype of archaeological objects of interest and consider it as a representative for the larger classes, selection of a study area) which must be considered considering the focus of this project being only on a limited number of monument classifications.
- Limited amount of validated data. The deep learning approaches typically used for image analysis are data hungry machines that need large quantities of labelled data in order to learn good predictive models. In archaeology, obtaining large, labelled curated datasets is very expensive and might require field studies in remote areas (since not everything can be assessed using the available imagery/data). There are several approaches that could be used to alleviate this. Firstly, data augmentation methods could be used. Second, transfer learning methods could be used. These methods learn predictive models using data from one domain and then transfer the model (and fine-tune it) to the domain of interest. Here, the closer the domains are the more successful the transfer is. Third, methods for self-supervised/unsupervised learning of models do not need (large amounts) labelled data to learn good predictive models. However, this comes at the price that they do not have the information about the archaeological objects of interest and will look for any types of objects encountered in the imagery data. Based on the labelled data availability, we shall select the most appropriate method to use here (starting from data augmentation, through transfer learning to self-supervised learning).
- The size of the objects of interest is much smaller than the available image. In standard image analysis, typically the objects of interest take up the whole image, in archaeology this is not always the case.

- There is a shortage of publicly available, AI/ML-ready archaeological data. This hinders the customization of ML methods for archaeology since the available data for use is limited to smaller groups of archaeologists, and ML experts do not have easy access to such data. This affects our work here in that there are no clear established guidelines on which methods should be focused on first.

We performed a literature overview to clarify the potential candidate ML methods to use for the data at hand. The focus here is on methods for object detection and image segmentation for archaeological imagery, hence, we consider two main sources of information:

- Discover the mysteries of the Maya competition 2021 (at ECML PKDD 2021). Detailed description of the competition, data, and resources are available for download at https://biasvariancelabs.github.io/maya_challenge/. The main objective of the competition was to “explore the potential of the Sentinel satellite data, in combination with the available lidar data, for integrated image segmentation in order to locate and identify “lost” ancient Maya settlements (aguadas, buildings and platforms), hidden under the thick forest canopy.” The data pertained to Mayan sites in the Chactun area in Mexico, and the goal was to identify and locate three types of objects: platforms, aguadas and buildings. Based on the competition entries, we identified the best performing methods to include variants of DeepLabV3 and U-Net with ResNet100 encoders and/or HRNETs.
- Benchmarks of methods for semantic segmentation available at <https://paperswithcode.com/task/semantic-segmentation>. Here, we see the performance of a variety of methods for semantic segmentation (described in over 2400 papers with code related to this topic) across 56 benchmarks. By browsing the benchmarks, we can see the current front-runners in this area which will be the prime candidates for evaluation on the data from this project (R-CNN, YOLO, DeepLabV3, HRNET, Swin Transformers, and SpotNET).

Time period	Advances	Limitations
2000–2010	- Segmentation and mathematical algorithms (Bescoby, 2006) - Edge detection for roadway identification (De Laet et al., 2007) - Morphometric variables used for classification include: - Shape (De Laet et al., 2007; Menze et al., 2006) - Compactness and smoothness (De Laet et al., 2007) - Colour/pixel value (Jahjah et al., 2007) - Neighbourhood analysis (De Laet et al., 2007, 2008, 2009) - Elevation (Menze et al., 2006) - Pattern recognition (template matching) is implemented as an automatic detection method (Trier et al., 2009)	- Most studies use two-dimensional (2D) aerial and satellite imagery, not three-dimensional (3D) topographic datasets such as LiDAR (Menze et al., 2006 is the exception). This limits the variables that can be used to identify features. - Low number of morphometric criteria were inefficient at capturing morphological diversity of certain feature types - High number of false-positive and false-negative results (e.g. De Laet et al., 2007, 2008, 2009; Menze et al., 2006, 2007; Trier et al., 2009)
2010–2015	- Use of LiDAR and 3D datasets becomes prevalent - Morphometric variables used for classification include: - Elevation, slope, and curvature (Schneider et al., 2015; Verhagen & Drăguț, 2012) - Nearest neighbour analysis, size, shape, and circularity (Harrower et al., 2013; Scheutter et al., 2013) - Hillshade and topographic position index (Schneider et al., 2015) - Principle component analysis (Chen, Comer, Priebe, Sussman, & Tilton, 2013) - Orientation and topographic contours (D'Orazio et al., 2012; Figorito & Tarantino, 2014) - Eccentricity (Figorito & Tarantino, 2014) - Volume (Menze & Ur, 2012) - Pattern recognition (template matching) continues to be utilized (Kvamme, 2013; Schneider et al., 2015; Trier et al., 2015; Trier & Pilø, 2012) - Increased accuracy achieved by: - Use of statistical classifiers (Chen et al., 2013; Trier et al., 2015) - Using multiple datasets to cross-reference automated results from different sources (Trier et al., 2015) - Using multitemporal datasets to account for possible seasonal lapses in visibility in remotely sensed imagery (Menze & Ur, 2012)	- Spatial resolution of remote sensing data (> 1 m) sometimes prevent accurate detection of small deposits (Verhagen & Drăguț, 2012) - Many of these methods have significant issues with false-positive and false-negative results (e.g. Kvamme, 2013; Harrower et al., 2013; Scheutter et al., 2013; Schneider et al., 2015; Trier & Pilø, 2012; Trier et al., 2015) - Some methods cannot detect features (or present false positives) that are located in close proximity to certain types of topographic anomalies (e.g. D'Orazio et al., 2012; Trier et al., 2015; Trier & Pilø, 2012)
2015–2018	- Co-opting of hydrological depression analyses for mound detection (Davis et al., In Press; Freeland et al., 2016) - Morphological variables used for classification include: - Circularity (Davis et al., 2018; Freeland et al., 2016; Witharana et al., 2018) - Rectangularity (Zingman et al., 2016) - Area (Davis et al., 2018; Magnini et al., 2017; Witharana et al., 2018) - Length and width (Magnini et al., 2017; Toumazet, Vautier, Roussel, & Dousteysier, 2017) - Size (Cerrillo-Cuenca, 2017; Davis et al., 2018; Zingman et al., 2016) - Curvature (Cerrillo-Cuenca, 2017) - Edge detection (Traviglia & Torsello, 2017; Witharana et al., 2018; Zingman et al., 2016) - Elevation (Davis et al., 2018; Guyot et al., 2018) - Multiscalar analysis with multiple datasets are incorporated to cross-validate results at small-to-large scales (Guyot et al., 2018; Witharana et al., 2018) - Pattern analysis continues to be used (Davis et al., 2018; Trier et al., 2015; Wang et al., 2017)	- False-positive and false-negative results (Freeland et al., 2016; Schneider et al., 2015; Trier et al., 2015; Witharana et al., 2018) - Lack of temporal control during feature detection (Traviglia & Torsello, 2017) - Small sample sizes for pattern recognition (Davis et al., 2018; Wang et al., 2017; Zingman et al., 2016)

Table 1: Outline of developments of object-based image analysis (OBIA) as a method in archaeology from 2000 to 2018 (taken from (Davis, 2018)).

The literature overview revealed the following findings.

- The applications of deep learning for archaeological imagery data are very recent.
- There is no strong computationally experimental work to evaluate a wide range of methods.

- The performance of the methods is very much dataset dependent meaning that a given method might excel for dataset A and have mediocre performance for dataset B. Prompting the need to benchmark more methods before selecting the best method.
- The performance of the methods can approach human recognition levels. For example, the best performing methods on the Maya competition have IoU over 0.83, which is remarkable. However, it is worth noting that the data used here was carefully prepared for AI/ML exploitation in the sense that a good amount of labelled data was made available.
- Parametrization of the methods, i.e., performing a wide parameter sensitivity analysis and ablation studies (removing some “feature” of the model or algorithm and seeing how that affects performance), will be needed to assure robust model performance.

3.3.2. Selected candidate machine learning methods

Based on the review, we selected several machine learning methods to benchmark and investigate for potential use here: R-CNN, YOLO, DeepLabV3, HRNET, Swin Transformers, and SpotNET. Summary of the selected methods is given in Appendix i.

R-CNN model family (Regions with CNN). R-CNN combines two insights (Girshick et al., 2013): applying CNNs to regions proposed in a bottom-up manner to localize and segment objects and use supervised pre-training for another task (when labelled data is scarce) followed by domain-specific fine tuning. Fast and Faster R-CNN further improve the performance of R-CNN: the former jointly optimizes classification and bounding box regression, while the latter learns an additional sub-network for generating region proposals (Ren et al., 2015). Implementations of these methods are available at: <https://github.com/rbgirshick/rcnn> and <https://github.com/facebookresearch/detectron>.

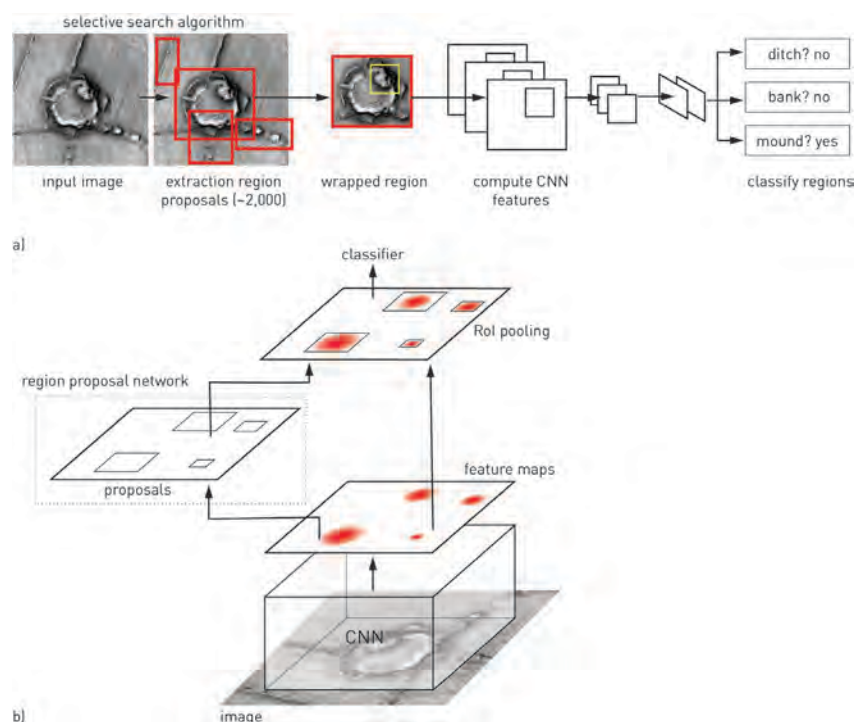


Figure 9: a) R-CNN architecture: Object detection system overview. The system takes an input image, extracts around 2000 bottom-up region proposals, computes features for each proposal using a large convolutional neural network (CNN), and then classifies each region using class-specific linear SVMs. b) Faster R-CNN is a single, unified network for object detection. The RPN module serves as the ‘attention’ of this unified network.

YOLO model family (You Only Look Once). YOLO (Redmon et al., 2015) models treat the task of object detection as a regression task where spatially separated bounding boxes and the associated class probabilities are predicted directly from full images in one evaluation. Since the whole detection pipeline is a single network, it can be optimized end-to-end directly on detection performance. This unified architecture is extremely fast. Figure 10 shows the YOLO architecture. Implementation of this method is available at <https://github.com/AlexeyAB/darknet>.

YOLOR (Wang, Yeh and Hong-Yuan Mark Liao, 2021) is a recent, successful extension of the YOLO framework. It is a unified network to encode implicit knowledge and explicit knowledge together, just like the human brain can learn knowledge from normal learning as well as subconsciousness learning. The unified network can generate a unified representation to simultaneously serve various tasks. Thus, kernel space alignment, prediction refinement, and multi-task learning in a convolutional neural network can be performed. The results demonstrate that when implicit knowledge is introduced into the neural network, it benefits the performance of all tasks. Furthermore, the analysis of the implicit representation learnt from the proposed unified network shows great capability on catching the physical meaning of different tasks. Figure 11 shows the YOLOR architecture. Implementation of this method is available at <https://github.com/WongKinYiu/yolor>.

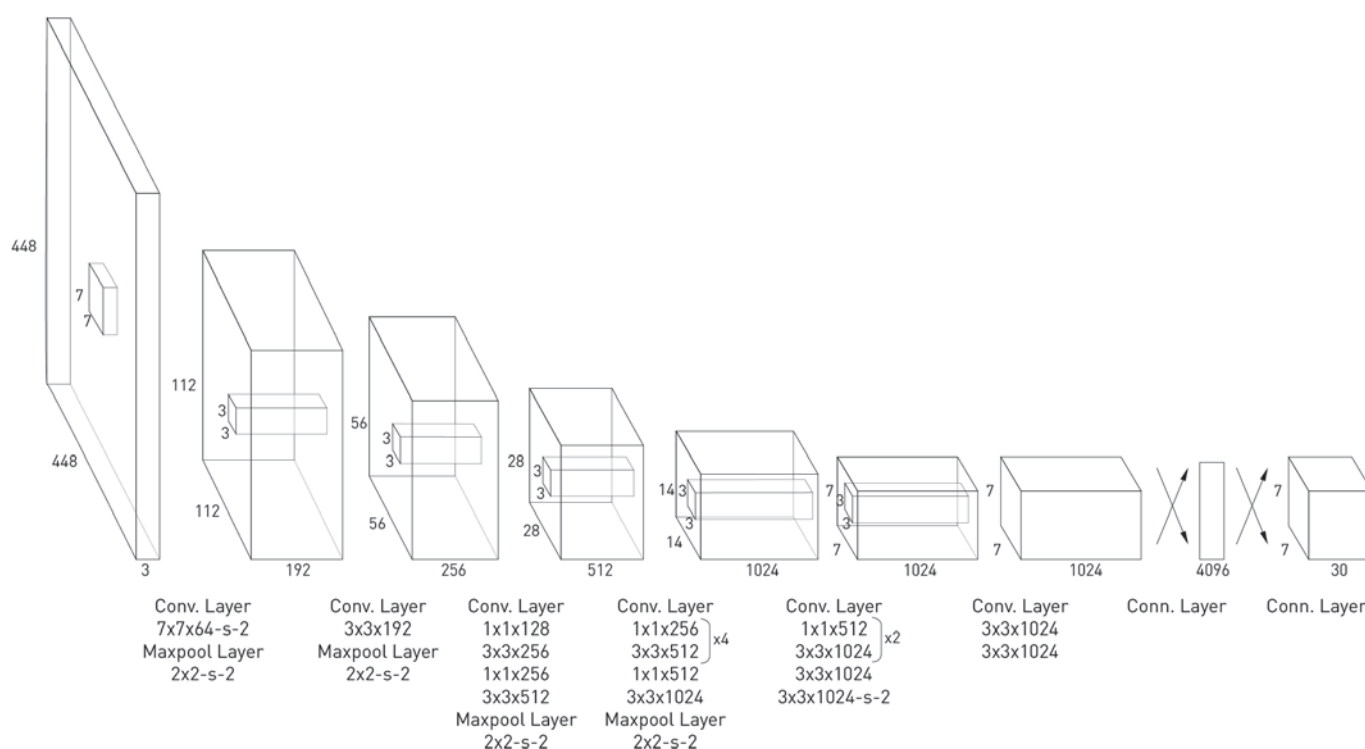


Figure 10: YOLO Architecture. The detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. The convolutional layers are pre-trained on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

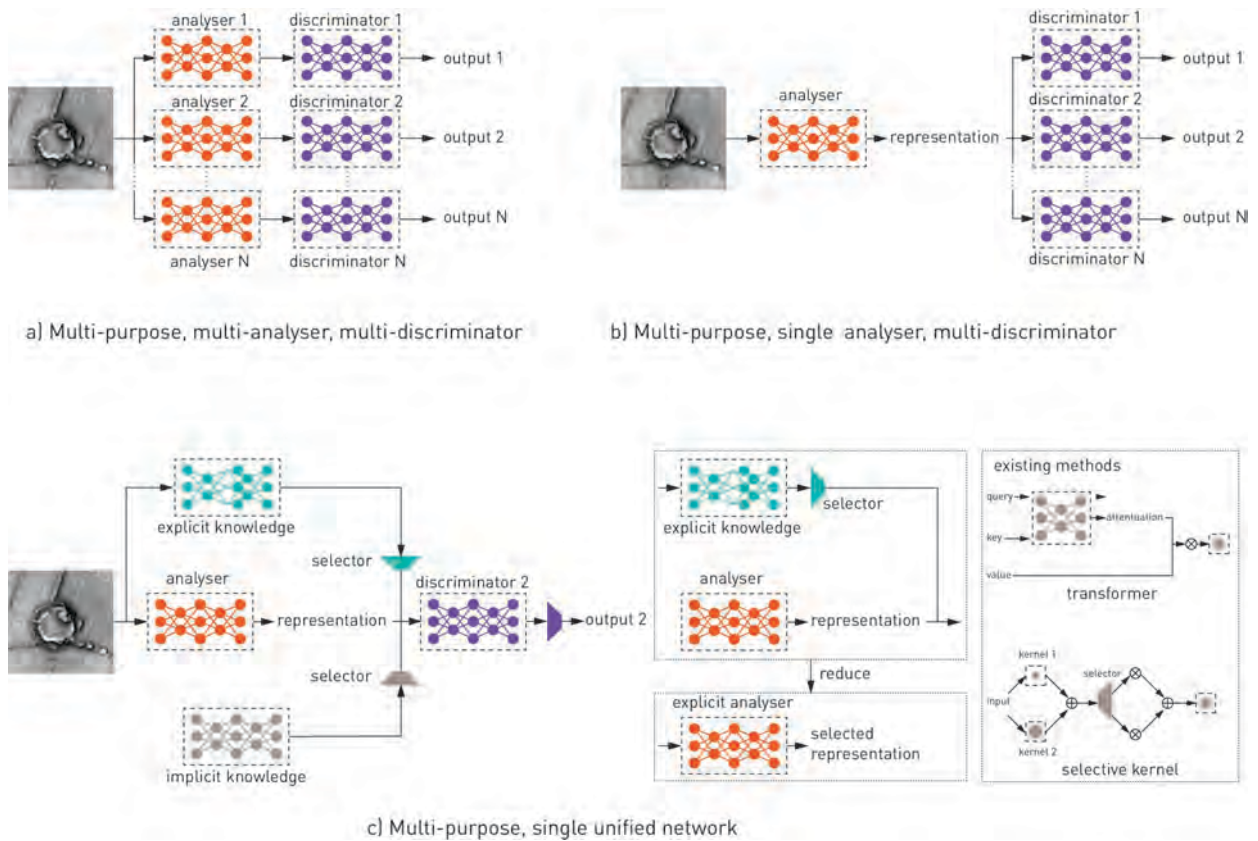


Figure 11: YOLOR multi-purpose NN architectures. (a) distinct models for distinct tasks; (b) shared backbone, different heads for different tasks; and (c) our proposed unified network: one representation with explicit knowledge and implicit knowledge for serving multiple tasks.

DeepLabV3. DeepLabV3 is a semantic segmentation architecture from the DeepLab family of methods. DeepLabV1 uses atrous (dilated) convolution to explicitly control the resolution at which feature responses are computed within Deep Convolutional Neural Networks. DeepLabv2 uses atrous spatial pyramid pooling (ASPP) to robustly segment objects at multiple scales with filters at multiple sampling rates and effective fields-of-views.

DeepLabv3 improves upon DeepLabv2 with several modifications. To handle the problem of segmenting objects at multiple scales, modules are designed which employ atrous convolution in cascade or in parallel to capture multi-scale context by adopting multiple atrous rates. Furthermore, the ASPP module from DeepLabv2 is augmented with image-level features encoding global context and further boost performance. More specifically, global average pooling is applied on the last feature map of the model, feed the resulting image-level features to a 1×1 convolution with 256 filters (and batch normalization), and then bilinearly upsample the feature to the desired spatial dimension. In the end, the improved ASPP consists of (a) one 1×1 convolution and three 3×3 convolutions with rates = (6, 12, 18) when output stride = 16 (all with 256 filters and batch normalization), and (b) the image-level features.

The architecture of DeepLabV3 is illustrated in Figure 12 and Figure 13 . Implementations of the different DeepLabV3 varieties is available at <https://github.com/tensorflow/models/tree/master/research/deeplab>.

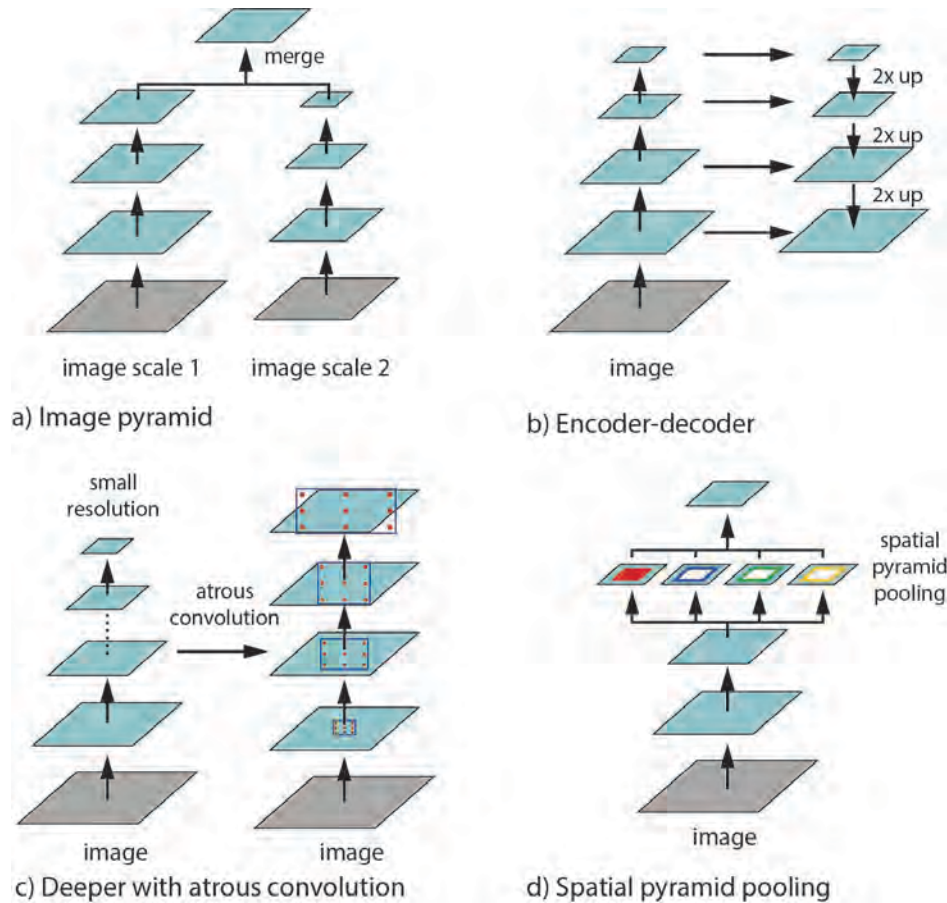


Figure 12: DeepLabV3 architecture for image segmentation illustrating the different architectures to capture multi-scale context.

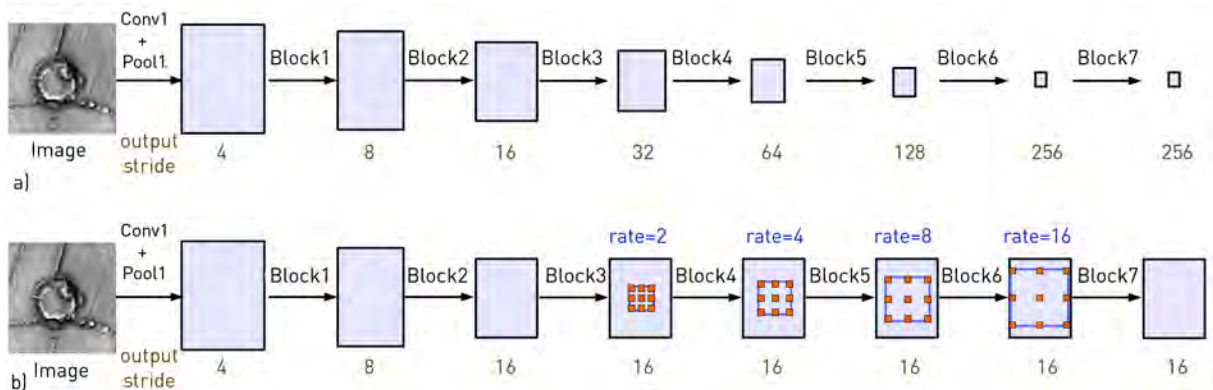


Figure 13: DeepLabV3 architecture for image segmentation illustrating cascaded modules without (a) and with (b) atrous convolution.

HRNET, or High-Resolution Net, is a general purpose convolutional neural network for tasks like semantic segmentation, object detection and image classification (Wang et al., 2019). It is able to maintain high resolution representations through the whole process. It starts from a high-resolution convolution stream, gradually add high-to-low resolution convolution streams one by one and connect the multi-resolution streams in parallel. The resulting network consists of several (e.g., 4) stages and the n th stage contains n streams corresponding to n resolutions. Repeated multi-resolution fusions by exchanging the information across the parallel streams

are conducted over and over. Figure 14 illustrates the HRNET architecture. Implementation of this method is available at <https://github.com/HRNet/>.

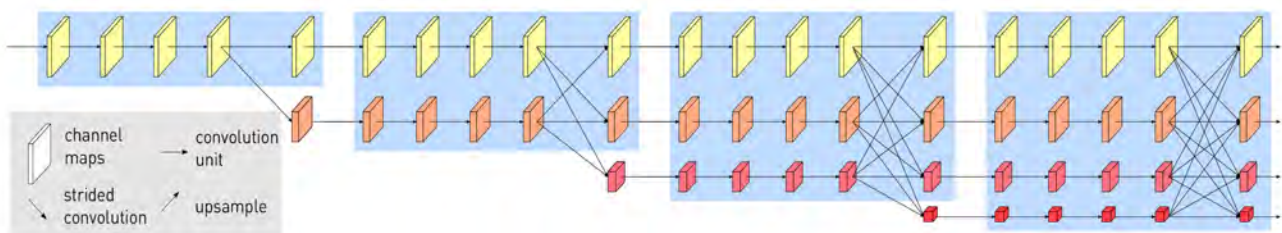


Figure 14: An example of a high-resolution network (HRNET). Only the main body is illustrated, and the stem (two stride-2 3×3 convolutions) is not included. There are four stages. The 1st stage consists of high-resolution convolutions. The 2nd (3rd, 4th) stage repeats two-resolution (three-resolution, four-resolution) blocks.

HRNET has been used in a recent work on Segmentation transformers, showing state-of-the-art performance on a variety of tasks and datasets (Yuan, Chen and Wang, 2019). There, the semantic segmentation problem is addressed by focusing on the context aggregation strategy. The main motivation is that the label of a pixel is the category of the object that the pixel belongs to. The proposed approach is simple yet effective and is based on object-contextual representations, characterizing a pixel by exploiting the representation of the corresponding object class. First, object regions are learned under the supervision of ground-truth segmentation. Second, the object region representation is computed by aggregating the representations of the pixels lying in the object region. Last, the relation between each pixel and each object region is calculated and augment the representation of each pixel with the object-contextual representation which is a weighted aggregation of all the object region representations according to their relations with the pixel. Figure 15 illustrates this approach. Implementation of this method is available at <https://github.com/open-mmlab/mmdetection> and <https://github.com/HRNet/HRNet-Semantic-Segmentation>.

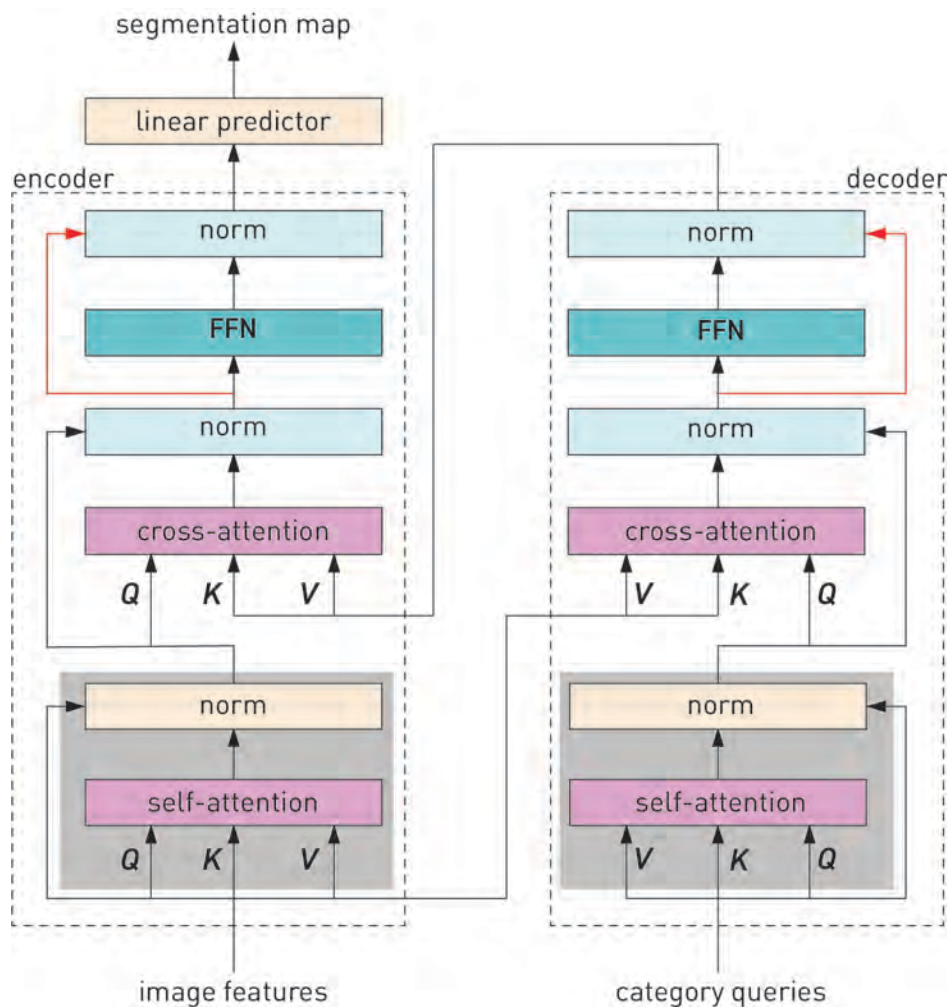


Figure 15: Segmentation transformer: The encoder self-attention unit in the grey box could be a local version and is optional and is useful for boosting the image features. The decoder self-attention unit serves as the role of interacting the category queries and

Swin Transformers. The Swin Transformer is a type of Vision Transformer (Liu et al., 2021b). It builds hierarchical feature maps by merging image patches (shown in grey, Figure 16) in deeper layers and has linear computation complexity to input image size due to computation of self-attention only within each local window (shown in red, Figure 16). It can thus serve as a general-purpose backbone for both image classification and dense recognition tasks. In contrast, previous vision Transformers produce feature maps of a single low resolution and have quadratic computation complexity to input image size due to computation of self-attention globally. Challenges in adapting Transformer from language to vision arise from differences between the two domains, such as large variations in the scale of visual entities and the high resolution of pixels in images compared to words in text. To address these differences, we propose a hierarchical Transformer whose representation is computed with Shifted WINdows. The shifted windowing scheme brings greater efficiency by limiting self-attention computation to non-overlapping local windows while also allowing for cross-window connection. Figure 16 illustrates the Swin transformers architecture. Liu et al. (2021a) present SwinTransformersV2 for scaling up capacity and resolution (see Figure 15, right image) – these yield state-of-the-art results in different object recognition tasks. Implementation of this method is available at <https://github.com/microsoft/Swin-Transformer>.

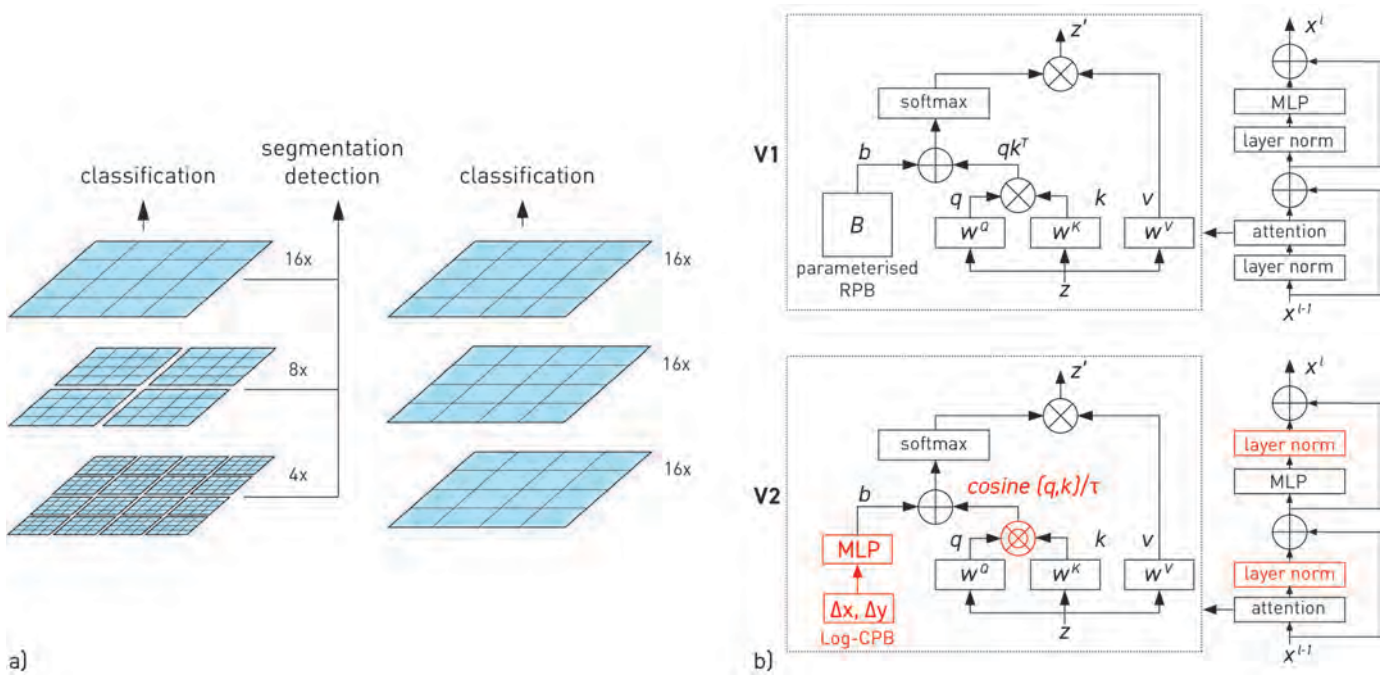


Figure 16: (a) The proposed Swin Transformer builds hierarchical feature maps by merging image patches (shown in grey) in deeper layers and has linear computation complexity to input image size due to computation of self-attention only within each local window (shown in red). It can thus serve as a general-purpose back-bone for both image classification and dense recognition tasks. (Middle image) In contrast, previous vision Transformers [20] produce feature maps of a single low resolution and have quadratic computation complexity to input image size due to computation of self-attention globally. (b) In order to better scale model capacity and window resolution, several adaptations are made on the original Swin Transformer architecture (V1): 1) a post-norm to replace the previous pre-norm configuration; 2) a scaled cosine attention to replace the original dot product attention; 3) a log-spaced continuous relative position bias approach to replace the previous parameterized approach. Adaptions 1) and 2) make the model easier to be scaled up in capacity. Adaption 3) makes the model more effectively transferred across window resolutions. The adapted architecture is named Swin Transformer V2.

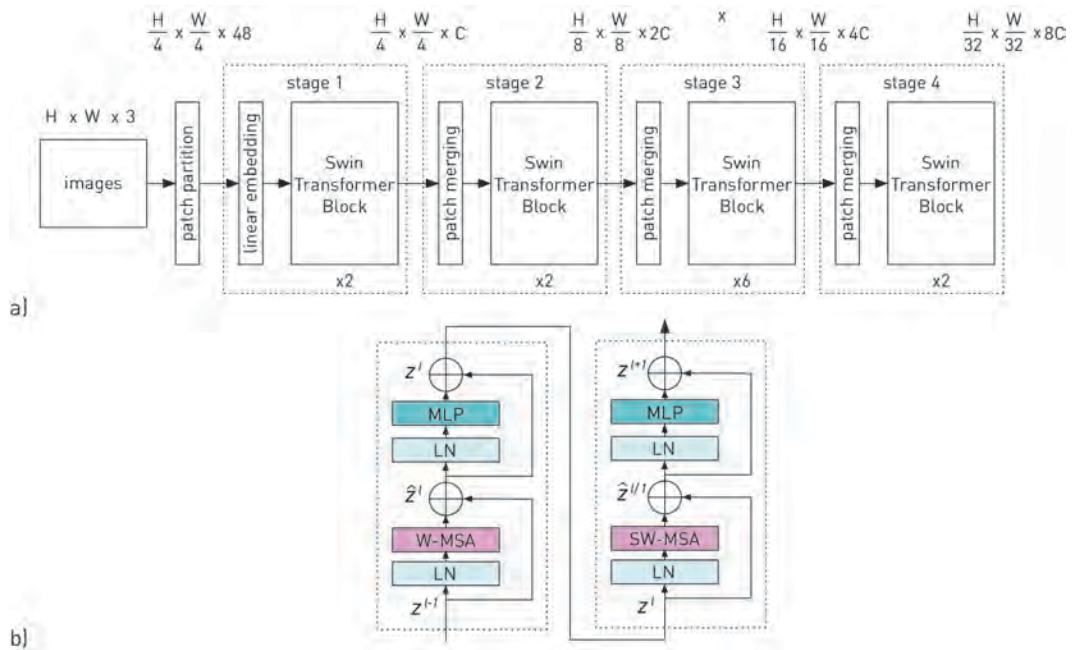


Figure 17: Figure 15. (a) The architecture of a Swin Transformer (Swin-T); (b) two successive Swin Transformer Blocks (for the notation see (Liu et al., 2021b)). W-MSA and SW-MSA are multi-head self-attention modules with regular and shifted windowing configurations, respectively.

SpotNET. Humans are very good at focusing their visual attention toward relevant areas when they search for different types of objects. For instance, when we search for cars, we will look at the streets, not at the top of buildings. The motivation behind SpotNET is to train a network to do the same via a multi-task learning approach (Perreault et al., 2020). To train visual attention, foreground/background segmentation labels are produced in a semi-supervised way, using background subtraction or optical flow. Using these labels, an object detection model is trained to produce foreground/background segmentation maps as well as bounding boxes while sharing most model parameters. Those segmentation maps are then used inside the network as a self-attention mechanism to weight the feature map used to produce the bounding boxes, decreasing the signal of non-relevant areas. Figure 16 illustrates the SpotNet architecture. Implementation of this methods is available at <https://github.com/hu64/SpotNet>.

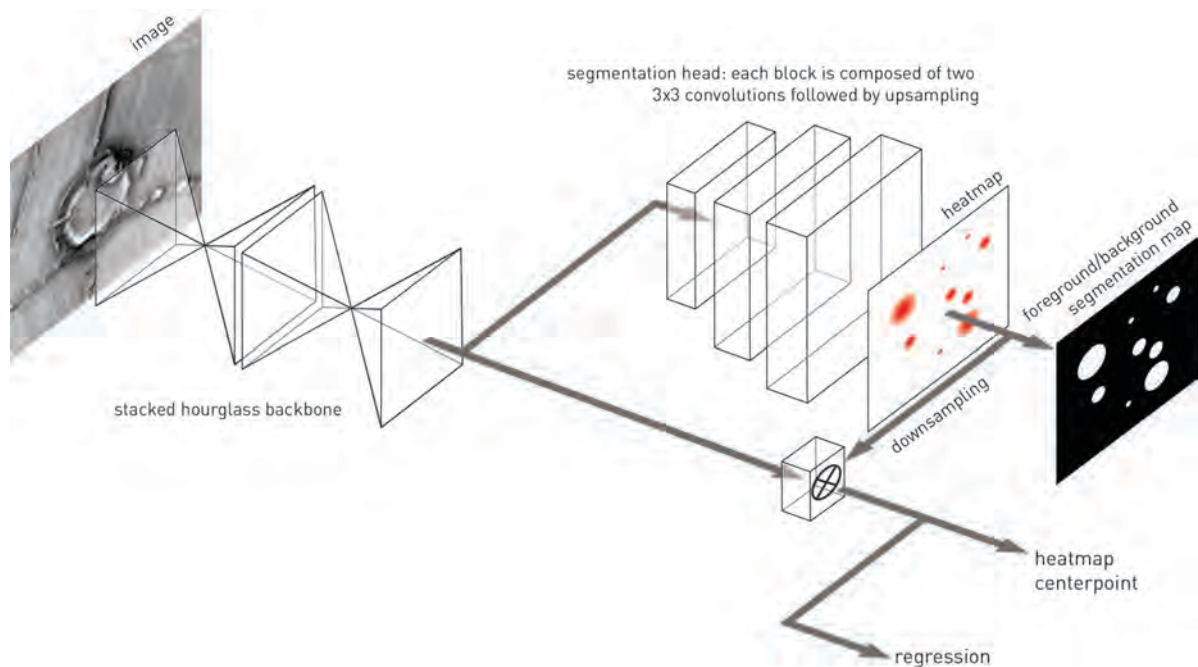


Figure 18: Overview of SpotNet: the input image first passes through a double-stacked hourglass network; the segmentation head then produces an attention map that multiplies the final feature map of the backbone network; the final centre keypoint heatmap is then produced as well as the size and coordinate offset regressions for each object.

Recommendations: Based on the overview, the following recommendations regarding the use of machine learning in archaeology emerge:

1. the work needs to be highly interdisciplinary involving experts from archaeology, remote sensing, earth sciences and machine learning;
2. the data needs to be prepared carefully, and special effort should be put on removing data biases;
3. usage of data augmentation is strongly recommended;
4. there is no clear top performing model, rather a set of models performing good across different datasets, hence benchmarking several methods on the project data is a must
5. the project proposes that the following models are evaluated and employed R-CNN, YOLO, DeepLabV3, HRNET, Swin Transformers, and SpotNET

4. Post-processing

Post-processing can improve the results (performance of the CNN) by reducing the false positive detections and provides more readable, understandable and usable results, e.g., georeferenced vector maps equipped with detection confidence.

4.1. Location-based ranking

Location-Based Ranking (LBR) can be implemented, to reduce false positives and thereby improve performance (Verschoof-van der Vaart et al. 2020; Verschoof-van der Vaart and Landauer 2021). Detections in high-ranking zones (rank 1) are more likely to be true, while detections in low-ranking zones (rank 2) have a much higher likelihood of being false positives. LBR can be used to reduce the number of false positives by ignoring detections in low-ranking zones. To determine the principal landscape characteristics involved, a broad-brush landscape characterization is performed. This evaluation identifies processes that are the most detrimental to the preservation and visibility of archaeological features and causing the most false positives. Areas where such processes occur may include:

1. (post)medieval agricultural areas,
2. modern agricultural fields,
3. disturbed areas (quarries, sod cutting, etc.) and
4. to a different extent, urbanized or built-up areas and modern paths and roads and areas of additional landscape development (e.g., golf courses).

The process is reflected in the work carried out by Garcia-Molsosa et al. (2021) who also divided their results into a high-probability and a low-probability categories, although these are separated by a size threshold. They reason that a probabilistic framework is particularly well suited for large-scale approaches where cross-referencing of information obtained through different methods and sources can be used to weight possible sites. Previously created landscape characterisation datasets such as historic land-use assessment / historic landscape characterisation (HLC) (Lambrick, Hind, and Wain 2013) could be utilised to assist in the creation of a LBR however, HLC coverage within Ireland is not universal and datasets were created using a range of approaches (Lambrick, Hind, and Wain 2013)

Recommendation: Where possible historic land-use assessment and historic landscape characterisation datasets will be evaluated as an indicator for different zone ranks.

4.2. Corroboration with other remote sensing data

The results of CNN machine learning models can be corroborated with probability analyses based on other remote sensing data. Davis et al. (2021) used multitemporal Synthetic Aperture Radar (SAR) data from Sentinel-1 and multispectral data from Sentinel-2 to conduct a random forest (RF) probability analysis of shell ring locations in their study region. SAR details soil properties and multispectral data highlights additional information about moisture content and vegetation. The method works best in areas with limited vegetation cover. They evaluated geophysical separability between confirmed shell rings and adjacent areas without shell rings in Sentinel-1 and Sentinel-2 data using a 10 m buffer. While the process was successfully applied, it was done so with limiting assumptions and by greatly reducing the studied areas. There are studies that support the inclusion of satellite imagery in the deep learning model

(e.g. Berganzo-Besga et al. 2021; Orengo et al. 2020), but others have been less successful, primarily because their study area was dense forest (Simidjievski et al. 2021).

Incorporating free satellite data, either directly into the deep learning process or into post-processing, seems tempting but is outside the scope of this project. It has, however, potential for further study.

4.3. Feedback loop

Verschoof et al. (2021) implemented a feedback loop between the archaeological interpreter and the classification algorithm to enhance the quality of the (training) dataset and the classification model. The objects within the dataset were initially rapidly labelled. Subsequently, the algorithm was trained and tested on this data. The incorrect detections (false positives) were manually compared to the ALS data to determine their cause. This resulted into multiple unlabelled objects, also called 'new positives', which were missed during the initial tagging. The training dataset was updated with these new positives and used to retrain the algorithm. This process was repeated two times until no new positives were detected. This approach allows for the rapid labelling of the initial training dataset, reducing the time needed for pre-processing. Furthermore, it results in the improved performance and a thorough training dataset (Verschoof-van der Vaart and Landauer 2021). Two iterations of the feedback loop improved the performance (MCC) of the model to detect hollow roads by circa 20 % (a huge increase).

Recommendation: The project will test the effectiveness of applying the Feedback loop through the initial labelling of archaeological objects on the performance of the model.

5. Metrics

Evaluation of the performance of the different machine learning methods is not a straightforward task. It encompasses two major parts: evaluation strategy and evaluation measures. The selection of the evaluation strategy depends on the problem definition and the available data. There are two widely used evaluation strategies: holdout set evaluation and cross-validation. In holdout set validation, a portion of the data (typically $\frac{1}{3}$, $\frac{1}{4}$ or $\frac{1}{5}$) is set aside and only used for evaluation of the performance of the learned models. The remainder of the data is then used to train the predictive models and, if needed, to perform parameter tuning. In cross-validation, the data is first divided into K-folds (very often 10 folds are used). Next, one of the folds is set aside for testing and the remaining K-1 folds (in this case 9) are used for learning and parameter tuning, if needed. Each of the K folds is then used as a test set in round-robin manner. The performance is then aggregated across the K folds.

In terms of selecting evaluation measures, we shall focus on two types of measures. The first type will measure whether we predicted correctly that there is or not an object of interest in the test image (classification-based measures) and the second type will measure the overlap of the predicted regions of the objects of interest with the ground truth annotations (segmentation-based measures). Note that it is difficult to compare the performance of the machine learning models because it is sensitive to the models' parameters, target features, input data (acquisition, processing, land cover, visualisation...), and the spatial scale for model prediction. Overall shapes of the archaeological objects in the automated and reference (i.e., manually digitized) datasets are often very similar, but the boundaries do not overlay exactly due to their fuzzy nature.

To assess an approach not only statistically but within an archaeological prospection scheme, a deep CNN model (the deployment model) can be trained using all possible reference data and applied to an independent set of images of a new study area. The model is then evaluated through human-interpretation and field survey.

Evaluation metrics respond differently to changes in study area and object characteristics. For example, a broad spatial scale for the model test area can increase omission (false negatives) and commission (false positives), resulting in a decrease of the F1-score. It is therefore difficult to compare the performance of the machine learning models in terms of any metric because most, if not all, the parameters that influence the metrics vary. Even if the models themselves and their parameters were fixed, target features, input data (acquisition, processing, land cover, visualisation...), and the spatial scale for model prediction are different.

The overall mAP hides local discrepancies that could be identified only through case-by-case visual analysis of model predictions. The metrics used for object detection and segmentation are often based on an overlap measurement (i.e., IoU) that is used as a threshold for determining a match or non-match. However, the complex relation between remotely sensed archaeological information and comprehensive archaeological information (e.g., excavation and field reports, archives) is not considered regardless of the threshold value (i.e., one or more values). The definition of reference data frequently raises issues in archaeological mapping, such as how remote sensing perceives the footprint of a known archaeological structure or diffuse footprints, such as large artificial mounds that have been eroding for thousands of years (Guyot et al. 2021).

In area-based assessment, producer's and user's accuracies are generally lower than the associated recall and precision values. This highlights the difficulty in mapping and assessing boundaries that are poorly defined or gradational in nature. When centre-weighting the areas to reduce the influence of boundaries (fuzzy accuracy method modified from Brandtberg et al. (2003)) an increase in user's and producer's accuracies, can be seen, which suggests that the objects were generally successfully mapped, though the boundaries may be disparate (Maxwell, Pourmohammadi, and Poyner 2020).

Similar concerns also arise for detecting undiscovered archaeological structures. A false detection by machine learning could become a true positive after in-situ verification. Therefore, a liberal strategy (rather than a conservative strategy) is required to define the detection thresholds (related to the confidence score and overlap measurement), which allows for a certain number of false negatives. Such a strategy can be justified to detect a maximum number of potential structures, as long as the prediction corresponds to a relevant response from the deep CNN considering the input examples it was trained on. Then, potential structures are interpreted based on human expertise.

These issues highlight that the current evaluation metrics, which originated from computer-vision and image-analysis domains, are only partially adapted to archaeological mapping. This could be considered in future studies such as by using fuzzy approaches.

5.1. Classification-based evaluation measures

5.1.1. Precision, Recall and F1 score

Precision is the ability of the classifier not to label as positive a sample that is negative, and **Recall** is the ability of the classifier to find all the positive samples. The **F-measure** ($F\beta$ and F1

measures) can be interpreted as a weighted harmonic mean of the precision and recall. A F_β measure reaches its best value at 1 and its worst score at 0. With $\beta=1$, F_β and F_1 are equivalent, and the recall and the precision are equally important.

In a binary classification task, the terms "positive" and "negative" refer to the classifier's prediction, and the terms "true" and "false" refer to whether that prediction corresponds to the external judgment (sometimes known as the "observation"). Given these definitions, we can formulate the following table known as confusion matrix:

Predicted class	Actual class	
	TP (true positive)	FP (false positive)
	FN (false negative)	TN (true negative)

Hence, we can formally define the notions of precision, recall and F-measure:

$$precision = \frac{TP}{TP + FP},$$

$$recall = \frac{TP}{TP + FN},$$

$$F_\beta = (1 + \beta^2) \frac{precision \times recall}{\beta^2 precision + recall}, \text{ and for } \beta = 1$$

$$F_1 = 2 \frac{precision \times recall}{precision + recall}.$$

In the multiclass classification task (as the one considered here), the notions of precision, recall, and F-measures can be applied to each class independently. There are a few ways to combine results across classes, specified by the average argument to the F1 score, precision score and recall score functions. We consider three averaging approaches micro, macro and weighted:

- **Micro:** Calculate metrics globally by counting the total true positives, false negatives and false positives.
- **Macro:** Calculate metrics for each label and find their unweighted mean. This does not take label imbalance into account.
- **Weighted:** Calculate metrics for each label, and find their average weighted by support (the number of true instances for each label). This alters 'macro' to account for label imbalance; it can result in an F-score that is not between precision and recall.

Note that if all labels are included, 'micro'-averaging in a multiclass setting will produce precision, recall and F1 score that are all identical to accuracy, while 'macro'-averaging of recall will produce the same results as balanced accuracy. Also note that 'weighted' averaging may produce an F-score that is not between precision and recall.

5.1.2. Accuracy

Accuracy is the most widely used evaluation measure for classification. Essentially, accuracy gives the fraction of correct predictions made by a predictive model.

If $\hat{y}_i$ is the predicted value of the i -th sample and y_i is the corresponding true value, then the fraction of correct predictions over n examples is defined as

$$accuracy(y, \hat{y}) = \frac{1}{n_{examples}} \sum_{i=0}^{n_{examples}-1} 1(\hat{y}_i - y_i)$$

where $1(x)$ is the indicator function.

Considering the notation from the previous subsection and the confusion matrix, accuracy can be also calculated as

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

5.1.3. Area under the receiver operating characteristic curve (AUROC)

A receiver operating characteristic (ROC) curve plots the performance of a (binary) classifier by varying the discrimination or decision threshold. More precisely, it plots the true positive rate (TPR) vs. false positive rate (FPR) at varying thresholds. TPR is the fraction of true positive examples out of all positive examples (aka *sensitivity*), and FPR is the fraction of false positive examples out of all negative examples (equal to $1 - specificity$).

Plotting of the ROC curve needs the true binary value and the target scores, which can either be probability estimates of the positive class, confidence values, or binary decisions. Moreover, by computing the area under the ROC curve, the curve information is summarized in one number.

For the task of binary classification, plotting of the ROC curve and the calculation of the AUROC value is done according to the explanation above. For the task of multi-class classification, this is no longer that straightforward. Here, two averaging strategies are currently supported: the *one-vs-one* algorithm computes the average of the pairwise ROC AUC scores, and the *one-vs-rest* algorithm computes the average of the ROC AUC scores for each class against all other classes. In both cases, the predicted labels are provided in an array with values from 0 to $n_{classes}$, and the scores correspond to the probability estimates that a sample belongs to a particular class. The OvO and OvR algorithms support weighting uniformly (*average='macro'*) and by prevalence (*average='weighted'*).

One-vs-one (OvO) Algorithm: Computes the average AUC of all possible pairwise combinations of classes and it is weighted uniformly:

$$\frac{1}{c(c-1)} \sum_{j=1}^c \sum_{k>j}^c (AUC(j|k) + AUC(k|j))$$

where c is the number of classes and $AUC(j|k)$ is the AUC with class j as the positive class and class k as the negative class. In general, $AUC(j|k) \neq AUC(k|j)$ in the multiclass case. This algorithm is used by setting the keyword argument *multiclass* to 'ovo' and *average* to 'macro'.

The multiclass AUC metric can be extended to be weighted by the prevalence:

$$\frac{1}{c(c-1)} \sum_{j=1}^c \sum_{k>j}^c p(j \cup k) (AUC(j|k) + AUC(k|j))$$

where c is the number of classes. This algorithm is used by setting the keyword argument multiclass to 'ovo' and average to 'weighted'. The 'weighted' option returns a prevalence-weighted average.

One-vs-rest (OvR) Algorithm: Computes the AUC of each class against the rest. The algorithm is functionally the same as the multilabel case. To enable this algorithm, set the keyword argument multiclass to 'OvR'. Like OvO, OvR supports two types of averaging: 'macro' and 'weighted' calculated in the same way.

5.1.4. Matthews Correlation Coefficient (MCC)

MCC is a measure of the correlation between the observed and predicted binary (two-class) classification, even if the classes are very imbalanced (Herfort et al. 2019). It takes into account the balance ratios of all four categories (TP, FN, TN, and FP) and is therefore a more reliable indicator of quality of the model over the other metrics used (Luque et al. 2019). Range [-1, 1].

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP \times FN) \times (TN + FP) \times (TN \times FN)}}$$

5.2. Segmentation-based evaluation measures

The most widely used evaluation measure for the tasks of object detection and image segmentation is Intersection-over-Union (IoU). Essentially, this is a Jaccard index score of the pixel representation of the images. More specifically, it is calculated as the ratio of the intersection of the mask from the ground truth and the prediction, and the union of the masks from the ground truth and prediction (as illustrated in Figure 19). An IoU score bigger than 0.5 is normally considered a "good" prediction.



Figure 19: Illustration of Intersection-over-union (IoU) as a similarity measure for object detection on images. Image taken from Wikipedia (https://en.wikipedia.org/wiki/Jaccard_index).

Recommendation: The project will test and assess several accuracy and evaluation strategies including AUROC, MCC and IoU during the development of the MLM.

6. Implications of using ML

A certain degree of involvement from an archaeological expert is and should remain necessary, in the least to interpret the automated detection results, as the goal of these methods is not to entirely replace the archaeological expert or 'automate archaeology' (Traviglia, Cowley, and Lambers 2016, 20). One thing that needs to be taken into account is that this development can

easily skew the known and protected archaeological sites towards localities with features visible in the DEMs, at the expense of, for example, prehistoric habitation sites with few or no above ground structures. The tools can also only detect objects similar to the predefined target class(es) while other objects are ignored (Lambers, Verschoof-van der Vaart, and Bourgeois 2019). Although automated detection can extend our knowledge on known archaeological classes, it may be unable to find potential new and/or unique types of archaeology. This process might unintentionally reinforce the dominance of the objects of interest in the archaeological record, by vastly multiplying the number of examples, while further marginalizing archaeological objects that are not detected (Nuninger et al. 2020), especially when automated detection is the sole source of information. It also remains to be seen what this means for heritage management, if the final frontiers devoid of archaeological sites will suddenly be teeming with potential features that need to be checked by archaeologists. Heritage management instances will need to find the resources and ways for inspecting these potential features and protecting the archaeological ones. One possible way forward could be offered by crowdsourcing and citizen science engagement where archaeology and history enthusiasts could be trained to detect potential sites from the ALS datasets, and then to inspect, document, and report them to the authorities, for instance using a mobile application (Seitsonen and Ikäheimo 2021). Within this project however, the efficient prospection of potential new archaeological sites through the application of MLM to ALS data should be considered to be only a single process within a much larger archaeological study which would include additional activities such as desk based research, analysis of aerial and satellite imagery, field survey and, for the preferred route at least, geophysical survey and if required, test excavation.

7. Section I References

- Abdelhack, Mohamed. 2020. 'An Open-Source Tool for Hyperspectral Image Augmentation in Tensorflow'. *ArXiv:2003.13502 [Cs, Eess]*, July. <http://arxiv.org/abs/2003.13502>.
- Asia Air Survey. 2019. 'Patents for RRIM'. 2019. <https://www.rrim.jp/en/patent/>.
- Banaszek, Łukasz, David C. Cowley, and Mike Middleton. 2018. 'Towards National Archaeological Mapping. Assessing Source Data and Methodology - A Case Study from Scotland'. *Geosciences* 8 (8): 272. <https://doi.org/10.3390/geosciences8080272>.
- Berganzo-Besga, Iban, Hector A. Orenge, Felipe Lumbreras, Miguel Carrero-Pazos, João Fonte, and Benito Vilas-Estévez. 2021. 'Hybrid MSRM-Based Deep Learning and Multitemporal Sentinel 2-Based Machine Learning Algorithm Detects Near 10k Archaeological Tumuli in North-Western Iberia'. *Remote Sensing* 13 (20): 4181. <https://doi.org/10.3390/rs13204181>.
- Bickler, S.H. (2021). Machine Learning Arrives in Archaeology. *Advances in Archaeological Practice*, 9(2), pp.186–191.
- Bonhage, Alexander, Mahmoud Eltaher, Thomas Raab, Michael Breuß, Alexandra Raab, and Anna Schneider. 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'. *Archaeological Prospection* 28 (2): 177–86. <https://doi.org/10.1002/arp.1806>.
- Brandtberg, Tomas, Timothy A Warner, Rick E Landenberger, and James B McGraw. 2003. 'Detection and Analysis of Individual Leaf-off Tree Crowns in Small Footprint, High Sampling Density Lidar Data from the Eastern Deciduous Forest in North America'. *Remote Sensing of Environment* 85 (3): 290–303. [https://doi.org/10.1016/S0034-4257\(03\)00008-7](https://doi.org/10.1016/S0034-4257(03)00008-7).
- Bundzel, Marek, Miroslav Jaščur, Milan Kováč, Tibor Lieskovský, Peter Sinčák, and Tomáš Tkáčik. 2020. 'Semantic Segmentation of Airborne LiDAR Data in Maya Archaeology'. *Remote Sensing* 12 (22): 3685. <https://doi.org/10.3390/rs12223685>.
- Canuto, Marcello A., Francisco Estrada-Belli, Thomas G. Garrison, Stephen D. Houston, Mary Jane Acuña, Milan Kováč, Damien Marken, et al. 2018. 'Ancient Lowland Maya Complexity as Revealed by Airborne Laser Scanning of Northern Guatemala'. *Science* 361 (6409): eaau0137. <https://doi.org/10.1126/science.aau0137>.
- Casana, Jesse. 2014. 'Regional-Scale Archaeological Remote Sensing in the Age of Big Data: Automated Site Discovery vs. Brute Force Methods'. *Advances in Archaeological Practice* 2 (3): 222–33. <https://doi.org/10.7183/2326-3768.2.3.222>.
- Casini, L., Roccetti, M., Delnevo, G., Marchetti, N. and Orrù, V. (2021). The Barrier of meaning in archaeological data science. *CoRR*, [online] abs/2102.06022. Available at: <https://arxiv.org/abs/2102.06022>.
- Chen, L.-C., Papandreou, G., Schroff, F. and Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *CoRR*, [online] abs/1706.05587. Available at: <http://arxiv.org/abs/1706.05587>.
- Chen, Ting, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. 'A Simple Framework for Contrastive Learning of Visual Representations'. *ArXiv:2002.05709 [Cs, Stat]*, June. <http://arxiv.org/abs/2002.05709>.
- Chiba, Tatsuro, Shin-ichi Kaneta, and Yusuke Suzuki. 2008. 'Red Relief Image Map: New Visualization Method for Three Dimension Data'. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37 (Part B2): 1071–76.
- Cowley, David C. 2012. 'In with the New, out with the Old? Auto-Extraction for Remote Sensing Archaeology'. In *SPIE Proceedings 8532*, edited by Charles R. Bostater, Stelios P. Mertikas, Xavier Neyt, Caroline Nichol, Dave Cowley, and Jean-Paul Bruyant, 853206. Edinburgh. <https://doi.org/10.1117/12.981758>.
- Cowley, David. 2016. 'What Do the Patterns Mean? Archaeological Distributions and Bias in Survey Data'. In *Digital Methods and Remote Sensing in Archaeology: Archaeology in the Age of Sensing*, edited by Maurizio Forte and Stefano Campana, 147–70. Quantitative Methods in the Humanities and Social Sciences. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40658-9_7.
- Cowley, David C., Łukasz Banaszek, George Geddes, Angela Gannon, Mike Middleton, and Kirsty Millican. 2020. 'Making LiGT Work of Large Area Survey? Developing Approaches to Rapid Archaeological Mapping and the Creation of Systematic National-Scaled Heritage Data'. *Journal of Computer Applications in Archaeology* 3 (1): 109–21. <https://doi.org/10.5334/jcaa.49>.
- Cowley, David C., George Geddes, Łukasz Banaszek, Kirsty Millican, and Iris Kramer. 2021. 'Expedite and Upscale Large Area Archaeological Survey in the Age of Big Data and Machine Learning' 112 (January): 15–17.
- Davis, D.S. (2018). Object-based image analysis: a review of developments and future directions of automated feature detection in landscape archaeology. *Archaeological Prospection*, 26(2), pp.155–163.
- Davis, D.S., Sanger, M.C. and Lipo, C.P. (2018). Automated mound detection using lidar and object-based image analysis in Beaufort County, South Carolina. *Southeastern Archaeology*, 38(1), pp.23–37.

- Davis, Dylan S., Gino Caspari, Carl P. Lipo, and Matthew C. Sanger. 2021. 'Deep Learning Reveals Extent of Archaic Native American Shell-Ring Building Practices'. *Journal of Archaeological Science* 132: 105433. <https://doi.org/10.1016/j.jas.2021.105433>.
- Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. 'ImageNet: A Large-Scale Hierarchical Image Database'. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–55. Miami, FL, USA: IEEE. <https://doi.org/10.1109/CVPR.2009.5206848>.
- Gallwey, Jane, Matthew Eyre, Matthew Tonkins, and John Coggan. 2019. 'Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial Heritage through Deep Transfer Learning'. *Remote Sensing* 11 (17): 1994. <https://doi.org/10.3390/rs11171994>.
- Garcia-Molsosa, Arnau, Hector A. Orengo, Dan Lawrence, Graham Philip, Kristen Hopper, and Cameron A. Petrie. 2021. 'Potential of Deep Learning Segmentation for the Extraction of Archaeological Features from Historical Map Series'. *Archaeological Prospection* 28 (2): 187–99. <https://doi.org/10.1002/arp.1807>.
- Girshick, R.B., Donahue, J., Darrell, T. and Malik, J. (2013). Rich feature hierarchies for accurate object detection and semantic segmentation. CoRR, [online] abs/1311.2524. Available at: <http://arxiv.org/abs/1311.2524>.
- Guyot, Alexandre, Marc Lennon, and Laurence Hubert-Moy. 2021. 'Objective Comparison of Relief Visualization Techniques with Deep CNN for Archaeology'. *Journal of Archaeological Science: Reports* 38: 103027. <https://doi.org/10.1016/j.jasrep.2021.103027>.
- Guyot, Alexandre, Marc Lennon, Thierry Lorho, and Laurence Hubert-Moy. 2021. 'Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach'. *Journal of Computer Applications in Archaeology* 4 (1): 1–19. <https://doi.org/10.5334/jcaa.64>.
- Halliday, Stratford. 2013. 'I Walked, I Saw, I Surveyed, but What Did I See?...and What Did I Survey?' In *Interpreting Archaeological Topography: Airborne Laser Scanning, 3D Data and Ground Observation*, edited by Rachel S. Opitz and David C. Cowley, 63–75. Oxford and Oakville, UK: Oxbow Books.
- Herfort, Benjamin, Hao Li, Sascha Fendrich, Sven Lautenbach, and Alexander Zipf. 2019. 'Mapping Human Settlements with Higher Accuracy and Less Volunteer Efforts by Combining Crowdsourcing and Deep Learning'. *Remote Sensing* 11 (15): 1799. <https://doi.org/10.3390/rs11151799>.
- Hinton, Geoffrey E., Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R. Salakhutdinov. 2012. 'Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors'. *ArXiv:1207.0580 [Cs]*, July. <http://arxiv.org/abs/1207.0580>.
- Kazimi, Bashir, Katharina Malek, Frank Thiemann, and Monika Sester. 2020. 'Effectiveness of DTM Derivatives for Object Detection Using Deep Learning'. In , 4. Vienna, Austria. <https://www.chnt.at/wp-content/uploads/Effectiveness-of-DTM-Derivatives-for-Object-Detection-Using-Deep-Learning.pdf>.
- Kazimi, Bashir, Frank Thiemann, Katharina Malek, Monika Sester, and Kourosh Khoshelham. 2018. *Deep Learning for Archaeological Object Detection in Airborne Laser Scanning Data*.
- Kokalj, Žiga, and Ralf Hesse. 2017. *Airborne Laser Scanning Raster Data Visualization: A Guide to Good Practice*. Prostor, Kraj, Čas 14. Ljubljana, Slovenia: Založba ZRC.
- Kokalj, Žiga, and Maja Somrak. 2019. 'Why Not a Single Image? Combining Visualizations to Facilitate Fieldwork and On-Screen Mapping'. *Remote Sensing* 11 (7): 747. <https://doi.org/10.3390/rs11070747>.
- Kokalj, Žiga, Klemen Zakšek, Krištof Oštir, Peter Pehani, Klemen Čotar, and Maja Somrak. 2019. *Relief Visualization Toolbox* (version 2.2.1). IDL. Ljubljana: ZRC SAZU. <https://iaps.zrc-sazu.si/en/rvt>.
- Küçükdemirci, Melda, Giacomo Landeschi, Nicolo Dell'Unto, and Mattias Ohlsson. 2021. 'Mapping Archeological Signs From Airborne Lidar Data Using Deep Neural Networks: Primary Results'. *ArcheoSciences* 451 (1): 291–93.
- Lambers, Karsten, Wouter Baernd Verschoof-van der Vaart, and Quentin P. J. Bourgeois. 2019. 'Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection'. *Remote Sensing* 11 (7): 794. <https://doi.org/10.3390/rs11070794>.
- Lambrick, George, Jill Hind, and Ianto Wain. 2013. *Historic Landscape Characterisation in Ireland: Best Practice Guidance*. The Heritage Council of Ireland Series. Dublin: The Heritage Council.
- Lin, Tsung-Yi, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2014. 'Microsoft COCO: Common Objects in Context', May. <https://arxiv.org/abs/1405.0312v3>.
- Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., Wei, F. and Guo, B. (2021a). Swin transformer V2: Scaling up capacity and resolution. CoRR, [online] abs/2111.09883. Available at: <https://arxiv.org/abs/2111.09883>.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. and Guo, B. (2021b). Swin transformer: Hierarchical vision transformer using shifted windows. CoRR, [online] abs/2103.14030. Available at: <https://arxiv.org/abs/2103.14030>.
- Lozić, Edisa, and Benjamin Štular. 2021. 'Documentation of Archaeology-Specific Workflow for Airborne LiDAR Data Processing'. *Geosciences* 11 (1): 26. <https://doi.org/10.3390/geosciences11010026>.
- Luque, Amalia, Alejandro Carrasco, Alejandro Martín, and Ana de las Heras. 2019. 'The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix'. *Pattern Recognition* 91 (July): 216–31. <https://doi.org/10.1016/j.patcog.2019.02.023>.

- Magnini, L. and Bettineschi, C. (2019). Theory and practice for an object-based approach in archaeological remote sensing. *Journal of Archaeological Science*, 107, pp.10–22.
- Maxwell, Aaron E., Pariya Pourmohammadi, and Joey D. Poyner. 2020. 'Mapping the Topographic Features of Mining-Related Valley Fills Using Mask R-CNN Deep Learning and Digital Elevation Data'. *Remote Sensing* 12 (3): 547. <https://doi.org/10.3390/rs12030547>.
- Nuninger, Laure, Philip Verhagen, Thérèse Libourel, Rachel Opitz, Xavier Rodier, Clément Laplaige, Catherine Fruchart, Samuel Leturcq, and Nathanael Levoguer. 2020. 'Linking Theories, Past Practices, and Archaeological Remains of Movement through Ontological Reasoning'. *Information* 11 (6): 338. <https://doi.org/10.3390/info11060338>.
- Olivier, M. and Verschoof-van der Vaart, W. (2021). Implementing State-of-the-Art Deep Learning Approaches for Archaeological Object Detection in Remotely-Sensed Data: The Results of Cross-Domain Collaboration. *Journal of Computer Applications in Archaeology*, 4(1), pp.274–289.
- Orengo, Hector A., Francesc C. Conesa, Arnau Garcia-Molsosa, Agustín Lobo, Adam S. Green, Marco Madella, and Cameron A. Petrie. 2020. 'Automated Detection of Archaeological Mounds Using Machine-Learning Classification of Multisensor and Multitemporal Satellite Data'. *Proceedings of the National Academy of Sciences* 117 (31): 18240–50. <https://doi.org/10.1073/pnas.2005583117>.
- Orengo, Hector A., and Cameron A. Petrie. 2018. 'Multi-Scale Relief Model (MSRM): A New Algorithm for the Visualization of Subtle Topographic Change of Variable Size in Digital Elevation Models'. *Earth Surface Processes and Landforms* 43 (May): 1361–69. <https://doi.org/10.1002/esp.4317>.
- Perreault, H., Bilodeau, G.-A., Saunier, N. and Maguelonne Héritier (2020). SpotNet: Self-attention multi-task network for object detection. CoRR, [online] abs/2002.05540. Available at: <https://arxiv.org/abs/2002.05540>.
- Pires de Lima, Rafael, and Kurt Marfurt. 2020. 'Convolutional Neural Network for Remote-Sensing Scene Classification: Transfer Learning Analysis'. *Remote Sensing* 12 (1): 86. <https://doi.org/10.3390/rs12010086>.
- Redmon, J., Santosh Kumar Divvala, Girshick, R.B. and Farhadi, A. (2015). You only look once: Unified, real-time object detection. CoRR, [online] abs/1506.02640. Available at: <http://arxiv.org/abs/1506.02640>.
- Ren, S., He, K., Girshick, R.B. and Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. CoRR, [online] abs/1506.01497. Available at: <http://arxiv.org/abs/1506.01497>.
- Ripley, Brian D. 1996. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511812651>.
- Seitsonen, Oula, and Janne Ikäheimo. 2021. 'Detecting Archaeological Features with Airborne Laser Scanning in the Alpine Tundra of Sápmi, Northern Finland'. *Remote Sensing* 13 (8): 1599. <https://doi.org/10.3390/rs13081599>.
- Shannon, C.E. 1949. 'Communication in the Presence of Noise'. *Proceedings of the IRE* 37 (1): 10–21. <https://doi.org/10.1109/JRPROC.1949.232969>.
- Shorten, Connor, and Taghi M. Khoshgoftaar. 2019. 'A Survey on Image Data Augmentation for Deep Learning'. *Journal of Big Data* 6 (1): 60. <https://doi.org/10.1186/s40537-019-0197-0>.
- Simidjievski, Nikola, Dragi Kocev, Ana Kostovska, Ivica Dimitrovski, Žiga Kokalj, Bertrand Le Saux, Andrej Draksler, et al. 2021. 'Discover the Mysteries of the Maya: ECML PKDD 2021 - Discovery Challenge'. Bias Variance Labs; ZRC SAZU; Jožef Stefan Institute; European Space Agency. https://biasvariancelabs.github.io/maya_challenge/about/.
- Somrak, Maja, Sašo Džeroski, and Žiga Kokalj. 2020. 'Learning to Classify Structures in ALS-Derived Visualizations of Ancient Maya Settlements with CNN'. *Remote Sensing* 12 (14): 2215. <https://doi.org/10.3390/rs12142215>.
- Soroush, Mehrnoush, Alireza Mehtash, Emad Khazraee, and Jason A. Ur. 2020. 'Deep Learning in Archaeological Remote Sensing: Automated Qanat Detection in the Kurdistan Region of Iraq'. *Remote Sensing* 12 (3): 500. <https://doi.org/10.3390/rs12030500>.
- Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. 'Dropout: A Simple Way to Prevent Neural Networks from Overfitting'. *Journal of Machine Learning Research* 15 (56): 1929–58.
- Stivaktakis, Radamanthys, Grigorios Tsagkatakis, and Panagiotis Tsakalides. 2019. 'Deep Learning for Multilabel Land Cover Scene Categorization Using Data Augmentation'. *IEEE Geoscience and Remote Sensing Letters* 16 (7): 1031–35. <https://doi.org/10.1109/LGRS.2019.2893306>.
- Štular, Benjamin, Stefan Eichert, and Edisa Lozić. 2021. 'Airborne LiDAR Point Cloud Processing for Archaeology. Pipeline and QGIS Toolbox'. *Remote Sensing* 13 (16): 3225. <https://doi.org/10.3390/rs13163225>.
- Štular, Benjamin, and Edisa Lozić. 2020. 'Comparison of Filters for Archaeology-Specific Ground Extraction from Airborne LiDAR Point Clouds'. *Remote Sensing* 12 (18): 3025. <https://doi.org/10.3390/rs12183025>.
- Suh, Ji Won, Eli Anderson, William Ouimet, Katharine M. Johnson, and Chandi Witharana. 2021. 'Mapping Relict Charcoal Hearths in New England Using Deep Convolutional Neural Networks and LiDAR Data'. *Remote Sensing* 13 (22): 4630. <https://doi.org/10.3390/rs13224630>.
- Toumazet, Jean-Pierre, François-Xavier Simon, and Alfredo Mayoral. 2021. 'Self-Adaptive Local Relief Enhancer (SAILORE): A New Filter to Improve Local Relief

- Model Performances According to Local Topography'. *Geomatics* 1 (4): 450–63. <https://doi.org/10.3390/geomatics1040026>.
- Traviglia, Arianna, David C. Cowley, and Karsten Lambers. 2016. 'Finding Common Ground: Human and Computer Vision in Archaeological Prospection'. *AARGnews - The Newsletter of the Aerial Archaeology Research Group*, no. 53: 11–24.
- Traviglia, Arianna, and Andrea Torsello. 2017. 'Landscape Pattern Detection in Archaeological Remote Sensing'. *Geosciences* 7 (4): 128. <https://doi.org/10.3390/geosciences7040128>.
- Trier, Øivind Due, David C. Cowley, and Anders Ueland Waldeland. 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'. *Archaeological Prospection* 26 (2): 165–75. <https://doi.org/10.1002/arp.1731>.
- Verhoeven, Geert J. 2017. 'Are We There Yet? A Review and Assessment of Archaeological Passive Airborne Optical Imaging Approaches in the Light of Landscape Archaeology'. *Geosciences* 7 (3): 86. <https://doi.org/10.3390/geosciences7030086>.
- Verschoof-van der Vaart, W. B., and Lambers, K.. 2021. 'Applying Automated Object Detection in Archaeological Practice: A Case Study from the Southern Netherlands'. *Archaeological Prospection*, 1–17. <https://doi.org/10.1002/arp.1833>.
- Verschoof-van der Vaart, Wouter Baernd, Karsten Lambers, Wojtek Kowalczyk, and Quentin P. J. Bourgeois. 2020. 'Combining Deep Learning and Location-Based Ranking for Large-Scale Archaeological Prospection of LiDAR Data from The Netherlands'. *ISPRS International Journal of Geo-Information* 9 (5): 293. <https://doi.org/10.3390/ijgi9050293>.
- Verschoof-van der Vaart, Wouter Baernd, and Juergen Landauer. 2021. 'Using CarcassonNet to Automatically Detect and Trace Hollow Roads in LiDAR Data from the Netherlands'. *Journal of Cultural Heritage* 47 (January): 143–54. <https://doi.org/10.1016/j.culher.2020.10.009>.
- Wang, C.-Y., Yeh, I-Hau. and Hong-Yuan Mark Liao (2021). You only learn one representation: Unified network for multiple tasks. CoRR, [online] abs/2105.04206. Available at: <https://arxiv.org/abs/2105.04206>.
- Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W. and Xiao, B. (2019). Deep high-resolution representation learning for visual recognition. CoRR, [online] abs/1908.07919. Available at: <http://arxiv.org/abs/1908.07919>.
- Wheatley, David, and Mark Gillings. 2002. *Spatial Technology and Archaeology: The Archaeological Applications of GIS*. London: CRC Press. <https://doi.org/10.1201/b12806>.
- Yokoyama, Ryuzo, Mlchlo Shlrasawa, and Richard J. Pike. 2002. 'Visualizing Topography by Openness: A New Application of Image Processing to Digital Elevation Models'. *Photogrammetric Engineering and Remote Sensing* 68: 251–66.
- Yuan, Y., Chen, X. and Wang, J. (2019). Object-contextual representations for semantic segmentation. CoRR, [online] abs/1909.11065. Available at: <http://arxiv.org/abs/1909.11065>.
- Zakšek, Klemen, Kristof Oštir, and Žiga Kokalj. 2011. 'Sky-View Factor as a Relief Visualization Technique'. *Remote Sensing* 3 (2): 398–415. <https://doi.org/10.3390/rs3020398>.

Section II: Irish monument detection & suitability review

8. Statutory Definitions

Following the Partition of Ireland in 1921, both jurisdictions on the island developed similar (though not identical) systems of monument identification, description and classification for inclusion in Sites and Monuments Records (SMR). While there are local variations in form and survival, most field monuments recorded in either Sites and Monuments Record may be found anywhere across the island of Ireland. There are a number of sites, mostly military-related sites from the second half of the twentieth century, that are only found within the jurisdiction of Northern Ireland. Many different kinds of archaeological sites found in Ireland are also found elsewhere in western and northern Europe, but there are also certain sites that are only, or mostly, found in Ireland, especially certain kinds of secular and ecclesiastical sites of the first millennium AD.

The definition of what constitutes an archaeological site or monument is very relevant to this project. Indeed, given the potential that ALS has for the identification of previously unrecorded sites, it is essential that one takes stock about what, precisely, one is seeking to identify. As outlined in the tendered proposal in respect of this project¹, a challenge in the use of a machine learning model (MLM) for the automatic detection of archaeological remains is its ability to identify and discriminate in the interpretation of features:

1. That recognise a set of features to be anthropogenic in origin and potentially of archaeological interest;
2. that ascribe initial interpretation to describe the likely form, function and complexity of the remains; and
3. the recognition that many sites do not readily conform to a perceived 'standard' form, with most sites being variations on a theme.

Arguably, these are challenges that already exist, and which people seek to overcome through the classification of sites into various groups or typologies. Through a process of learning and experience, one may begin to recognise the features that distinguish one feature in the landscape from another, to consider if it is likely to be the result of human interaction with the landscape, what form did that interaction take and what was the rationale for it, and (perhaps most importantly) what significance one may attach to any of those elements.

The definitions of what constitutes a monument or site of archaeological interest is provided for in law and is held in common by many states world-wide. The definitions do not, however, provide a name for each kind of archaeological site beyond the wider term on "monument". No distinction is made in Ireland, for example, between a burial site of the Neolithic period or one from the Iron Age; both would come under the definition above that refers to a burial "prehistoric or ancient". Similarly, a medieval cultivation ridge may be considered to be an "artificial structure", just as the major monuments of Newgrange or the Stone Forts of the Atlantic Coast may also each be considered to be an "artificial structure".

It is important to note that the statutory definition of a monument is not quite the same as the classifications of monuments compiled by archaeologists. As will be observed below, most classifications relate to composite structures that attempt to describe form, composition,

¹ As per Section 1.2.1, pp2-3 of DP tender submission, TII268 Services Contract Lot 5 2021

function or date. Those classifications move beyond simply identifying a feature of interest to the interpretation of the feature in question.

9. Classification of Monuments

Following the announcement of new legislation concerning 'ancient monuments' and archaeology on the island of Ireland after Partition in 1921, there was an important drive towards describing and explaining the nature of archaeological sites and providing information about the sites to the wider public. This can be observed in field guides and in the various county surveys and inventories that have been published since the *Preliminary Survey of Ancient Monuments of Northern Ireland* of 1940.

Ó Ríordáin's *Antiquities of the Irish Countryside*, first published in 1942, noted in its preface that it was intended "to give some answer to the questioning man in the street, who asks the archaeologist for a simple explanation of some monument which he has noticed in the countryside and which is, in all probability, a typical example of a class of antiquity widespread throughout Ireland" (Ó Ríordáin 1942). While Ó Ríordáin's publication was not directly intended for specialists in archaeology, the approach within it, to identify specific monument types and classifications, and to discuss these in the context of their form, composition, date, function, distribution, and excavated evidence, remains central to most field identifications undertaken in subsequent surveys and inventories compiled across the island of Ireland.

The subsequent development of lists or registers of monuments has occurred within the context of both structured surveys of geographical areas, surveys of certain monument types, and in-depth studies including excavations of particular sites. In these surveys one may observe something of a combination of the gathering of empirical evidence (the location, dimensions and condition of specific features at each location), interpretation of the sites (an estimation of their age, purpose and how they were formed) and attribution of significance (that is, why the individual site was of interest). The use of classifications enables one to 'make sense' of the data. Walsh (Walsh 1993, 101), writing in the context of classification for prehistoric monuments, notes that the process "*involves the identification of recurrent architectural forms or phenomena which are presumed to have been important to those communities who built and used them*", further noting "*consequently it is reasonable to infer that monuments of similar morphological form had some common function*". The process is equally applicable in the context of monuments from any other period. In essence, however, this may be read as 'find the similar morphologies, one finds a similar function'. Walsh counsels specifically against this, noting "we must beware also that when comparing monuments to classify them, we do not impose patterns and thereby infer a relationship where, in fact, none existed" (*ibid.*).

This caution is particularly relevant to the development of a machine learning model for monument recognition within pre-defined classifications. The identification of specific architectural forms or phenomena is one part of a process of site classification, arguably based upon a conceptual understanding of the recurrence of those forms or phenomena that has a particular meaning. In essence, one is attempting to establish how the composition or arrangement of such features may be described collectively as a kind of site and thus understood as an entity that is greater than simply the sum of its parts. Certain monuments with very different functions, dates or arrangement may conform to common morphological groupings which have no direct relationship in cultural terms, and thus no direct relationship

in terms of individual classifications. For example, certain large enclosures such as those associated with church sites, certain large ringforts and prehistoric ceremonial enclosures will share common elements which often vary as much from site to site within one classification as they do from one classification to another. Generalities, or more precisely concepts, about the classification may be offered, but a ploughed-down ceremonial enclosure may appear, superficially, to share the same kinds of morphology as a similarly altered large ringfort, hilltop enclosure or church site. Differentiating between each class may rely upon information that is not immediately visible in the field, and which has no tangible upstanding element, such as the occurrence of artefacts, historical sources (either about the site or about its earlier recorded morphology), or it may derive from tangible elements outside of the individual site, such as its topographical position or its physical and temporal relationship to other sites.

Thus, the classification may begin with the identification of recurrent architectural forms or phenomena, but it progresses through something of an intellectual and phenomenological interaction with the site, understanding its setting, its relationship to the wider landscape, and its relationship to other sites, based on the knowledge and experience of the individual carrying out the classification and the nature of the evidence used. Elements of that intellectual interaction may be conducted remotely from the site, but with caution; validation of the remains will normally be needed 'on the ground', by way of inspection or further survey including excavation.

9.1. Why do we classify monuments according to particular 'types'?

Ó Ríordáin's note in the preface to *Antiquities of the Irish Countryside* provides an insight into one of the simplest reasons why monuments are classified according to particular types: it enables an explanation of individual sites and monuments based on a wider consideration of the monument beyond just one or two examples of a 'typical' site. Such classification prompts the kind of interpretation (Walsh 1993, 101), whereby one may identify features that are presumed to have been important to whoever first constructed or used the monument, and to identify repetitions of those features elsewhere that slowly enables the reconstruction of the material world of the past and facilitates the discussion of why those features were important.

By classifying individual monuments, one may embark upon further studies of the individual sites, seeking places for comparison and analysis. For example, one may take the knowledge learned from one site that has been subject to detailed investigation (potentially including excavation) and apply it to another which has not been subject to that level of investigation in the past. Conversely, such study allows one to make inferences about monuments that are no longer extant but about which sufficient is known to make a meaningful comparison and identification of age, function or significance. Without such classifications academic and popular study of the past would be very difficult.

The classification process also enables the management of sites. Individual monuments are identified and once included in State records, become subject to a number of regulatory and policy controls, especially in the context of spatial planning. Tied into the process of classification is a process of ascribing significance to the remains, which in turn enables informed decisions to be made about how a site may be managed, preserved or indeed removed.

9.2. How are monuments classified

Most systems of classification tend to focus upon age, i.e., the period when the monument was first constructed, the materials used in the construction of the monument, and the final appearance of the monument, or the purpose for which the monument was originally constructed and used. These attributes are often used in the titles of individual categories (see below). At the present time there is a Class List of definitions of archaeological monument types, used in State records in Ireland. This Class List comprises some 480 individual monument definitions/classifications, with new monument classifications still being developed. While the initial drive towards the publication of county surveys halted in both Ireland and Northern Ireland by the early 1980s and 1990s, one may observe in publications such as *Recognising Irish Antiquities* (Archaeological Survey (Northern Ireland) 1980) and *Irish Field Monuments* (C. Manning 1991) that primary identification of field monuments has continued to be grouped according to period, composition, form and function. The current Class List of Definitions is listed alphabetically according to monument type.

This approach is an historical inheritance in archaeological research. It is something of a necessary and logical method of making sense of the wide range of sites and monuments that survive across the landscape, and key in the compilation of schedules, lists or registers of such sites. It takes the collection of discrete features found at a place and describes a site as a whole. The processes of classification and description have also evolved, whereby monuments are normally classified according to one or more of the following features or attributes:

1. Their physical form in plan (their shape, such as round, oval, square, rectangular, rectilinear, etc.).
2. Their physical form in profile or three dimensions (e.g., mound, enclosure, structure);
3. Their perceived original function (e.g., burial, grave, tomb, fort, fortification, mill);
4. The materials from which the monument is most obviously composed (e.g., earthwork);
5. A composite of their physical form and their perceived function (e.g., burial mound, ring fort, embanked enclosure);
6. An historical name by which the monument was known, frequently a Gaelic form that has endured and often now applied to certain types of monuments (*lios*, *crannog*, *caiseal*, etc.).
7. Individual classifications and some limitations and opportunities.

The extensive *Class List Definitions* used by the Archaeological Survey of Ireland (ASI) gives an indication of the complexity of site classification, but also clarity about the kinds of characteristics that may be considered to inform the interpretation of a set of features encountered in the field. It should be noted, however, that where a monument comprises multiple features, many of those discrete features have been given individual reference numbers – dots or polygons on maps – which means that individual complexes or complex monuments may have multiple references in the SMR. For example, an ecclesiastical site that comprises a church ruin, burial ground, a high cross and enclosing earthwork is likely to have at least four different references in the record, each of which essentially referring to the same place but emphasising one feature or another.

Within the current Class List there appears to be considerable effort given – where information is available to do so – to distinguish between features or monuments that have similar forms

or functions, but which have significant differences in terms, for examples, of their historical associations, their industrial outputs, temporal occurrence or, in some instances, size. For example, some 24 individual classifications are given for 'Religious Houses', based upon the classification of such foundations compiled by Gwynn and Hadcock (1970). In the Class List Definitions these sites are described as being of "the later medieval period", with a date range of twelfth to sixteenth centuries AD. Notwithstanding a debate about how the medieval period may be broken into early, high, or later periods (or how these 24 individual classifications could be sub-divided), the arrangement of this classification presents some challenges.

In many instances a conventual layout of buildings is normally associated with a religious order in the Middle Ages, and such conventual layouts follow generally accepted forms and principles. As Stalley (1987) demonstrated, however, even within a single religious order, there was considerable variation in the architecture and layout of its various buildings and estates, reflecting the wealth, influence, topography and materials used in the construction and maintenance of individual sites. Furthermore, in several instances the primary religious order using a site changed, commodifying the existing structures and using those rather than establishing a new set of buildings. Ultimately, after the dissolution of religious houses in Ireland, most of these foundations became ruins, often a source of stone for new buildings, or on occasion used in the sixteenth and seventeenth centuries as mansion-type dwellings. As such, each site has its own 'biography' and unique set of archaeological remains or survival. While the general arrangement of buildings may be compared between sites (as done, for example, by Stalley) the precise spatial arrangement of such buildings varies in all cases.

These religious houses are a useful case in point in terms of how inventories are compiled and some of the limitations that arise. In the period c.1130-1540 around 400 new religious houses were established in Ireland², with a number of pre-existing ecclesiastical foundations adopting the way of life of the major religious orders of the day. Allowing for some error in classification or interrogation of the records, one may suggest that the actual number of sites involved (including granges and out-farms) was perhaps in the order of 750 locations. As individual sites this would represent approximately 0.5% of all recorded archaeological sites and monuments in Ireland, covering a period of around 400 years, i.e., about 3.5% of the known human occupation of Ireland since the last Ice Age. However, the 24 different classifications used account for nearly 5% of all monument classifications. It is an instance where a lot of different kinds of a similar site type are known, and as such has a significant impact upon the number of individual classifications used. Conversely, the number of sites associated with the early medieval period, from c. AD 400 – 1100 represents at least one third of the total number of recorded sites (including some 45,000+ ringforts, early church sites, milling sites and areas of cultivation). However, settlement sites are broken into three primary classifications: ringfort – cashel, ringfort – rath, and ringfort – unclassified. A further classification – 'Enclosure' is often taken as a ringfort, but where the site has been largely destroyed or where few diagnostic elements have been observed that would enable a more precise definition as a ringfort. In this instance, therefore, a monument type (ringfort, including cashels, raths, enclosures and unclassified ringforts) which make up 0.8% of the classification categories account for some 30% of all recorded monuments, relating to about 6.5% of the known period of occupation of Ireland since the last Ice Age. Similar spikes of classification categories and site records may be observed throughout the record, with some 19 classifications relating to forms of grave

² <http://monastic.ie/religious-orders/>

markers dating from AD 400 and later (including tombs, sarcophagi, memorial stone and cenotaph) and a further 20 classifications relating to forms of crosses (high crosses, graveyard crosses, termon cross etc.). Where such sub-division into new classifications occur, one normally has a smaller total number of sites of a particular type overall, and issues about what may constitute a representative percentage of the monument type overall.

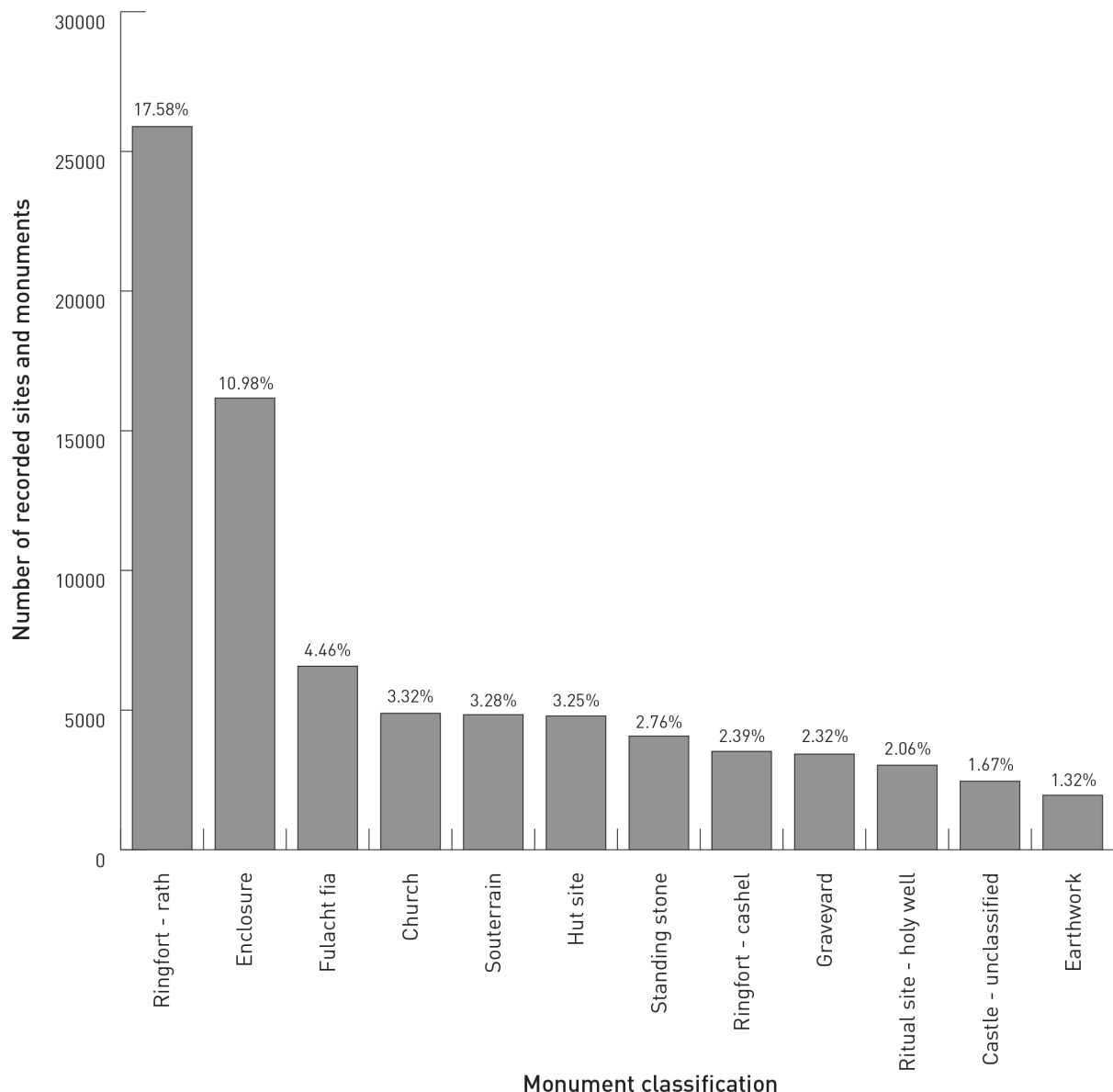


Figure 20: Bar chart displaying the total number of sites recorded within the SMR grouped by monument classification for the 10 most frequently recorded classes. Percentage values on the top of the bar chart indicate the percentage of total sites each monument classification represents. Calculation excludes the classification of "Redundant record" (data extracted from ASI on 07/01/2022).

As previously noted, there are features that are common to many different classifications of monuments, but their arrangement and function may differ considerably through time and location. Indeed, the monuments themselves (rather than just individual elements) may have been formed at very distant periods one from the other but share important overarching morphological characteristics. This is a fundamentally important issue for this project. The broad, or indeed the very specific classifications of monuments used by many archaeologists

today are inherently laden with meanings that frequently expand beyond the simple description of the remains. The inclusion (for classification) of attributes around function or date may be inferred from an examination of the physical remains at an individual site, but this in itself is part of an ongoing process of understanding individual and groups of sites.

For example, ringforts and ringworks, in addition to sounding similar, frequently look alike. Tadhg O’Keeffe (2000, 30) notes in his discussion of ringworks and ringforts that, when a monument is described as one or the other, one immediately associates with it not just morphology or date, but “function, culture and context”. Ringworks as a class of monument closely post-date the construction of ringforts, but are generally not considered to be an evolution of one into the other, and indeed while some of the everyday activities that occurred at these sites were largely the same (family life, tending livestock etc.) their broader function and especially how we understand these sites in a specific cultural context is fundamentally different, especially with regard to aspects of social and economic structure including personal or community identities of those who composed and used the sites. While there has been debate about the respective classification and distinction of ringforts and ringworks (see, for examples, Barry (1987, 45); Breen (2005, 43); Sweetman (Sweetman 1999, 4-16); O’Conor (1998, 18-20, 35-38)) our collective understanding of archaeological field monuments continues to develop over time. This site type, and others, is frequently examined and re-examined, with new ideas developing about their form and function (see for example Gardiner and O’Conor (2017)). While work continues to update ‘official’ records in light of new research, such updates rarely keep pace with academic debate about individual monuments or classifications. One should, therefore, consider the identifications included in resources such as the Record of Monuments and Places to be the best estimation of the monument’s classification at the time it was last surveyed.

There are practical considerations too, where features are shared across different classifications of monuments. Sometimes this may lead to simple misidentification, where one type is confused with another based on the understanding of its morphology at the time of survey. The debate around ringforts and ringworks is a case in point, but not an isolated one. For example, Barrett and Graham (1975, 40), noted that, in the analysis of potential ringforts in County Meath, at least some sites identified from aerial photography may in fact have been barrows (a prehistoric burial monument) rather than ringforts (an early historic domestic settlement site). Stout (1992, 58), in discussion of the accuracy of plans of monuments made in the course of Ordnance Survey field surveys in the nineteenth century, noted that in many instances precise classification was not always possible: one could not, on the basis of their mapped depictions, readily differentiate between ringforts and certain other circular earthworks. This observation is likely to hold true for the precise identification of monuments based on new ALS data. While ALS data will have the benefit of providing elevational data (and therefore not constrained to only two-dimensional depictions), similarities in monument form may well result in misidentification as one class of monument or another. Exploration of the elevational data requires focused attention, outside of the wider mapping exercise that is facilitated with ALS survey data. This has particular relevance in using the data for decision-making in the context of spatial development and land-use change; the topographical similarities may well mask significantly different sub-surface remains, or indeed mask the significance of the site in question.

Within the 480 monument definitions used by the Archaeological Survey of Ireland, one could develop multiple sub-classifications that describe aspects of a monument’s use, re-use,

abandonment and decay, as well as aspects of the cultural identity ascribed to it. Sometimes this is threaded into the description of a monument, such as noting a religious order that founded or maintained the site, or a major episode of re-use that re-modelled a site. In each case the additional information about the site helps build a clearer picture of what it was, of its significance, and the likely impacts of change on either its physical survival or its cultural value. ALS data is likely to provide rapid information to establish the presence, and extent, of upstanding features, especially earthwork features, and by extension inform an assessment of the likely below-ground remains that may be encountered.

It should also be understood that while monuments may be grouped under one classification or another, not every site is the same. There are significant variations in size, shape and composition of individual monuments. Megalithic 'tombs', for example, may be grouped into discrete typologies of court tomb, portal tomb, passage tomb or wedge tomb, but while individual monuments share very specific similarities, no two examples are perfectly the same (see de Valera and Ó Nualláin's *Survey of the Megalithic Tombs of Ireland* (1961)). Many classifications of one or another site that can be described as a kind of megalithic tomb may make use of additional historical information about a site which has been subsequently demolished, including written details, drawings or photographs, that allows one to classify a site that is no longer there.

This kind of insight is not possible in many instances, where the subsequent history of land use or decay has removed surface traces of the monument. Considering site taphonomy, be that at macro or micro levels, is important for the understanding of why the sites differ in appearance today. Monuments, as all objects, would have encapsulated entropy and life cycle in that monuments are built, used and decay, reaching a point of ruination at which decay slows and the monuments survive as 'fossils' in the modern landscape. Those processes can and do result in radically different outcomes for the historical survival of individual sites, to the point where a monument may be classified according to the survival of two or three key elements (such as traces of a court, passage, orientation of the structure and artefact associations) rather than the complete survival of the whole site. Furthermore, the monuments as first constructed may have had radically different appearances but based on the survival of key elements they have an attributed classification in the modern day.

As noted above, ALS surveys present a fundamentally different approach to investigating the landscape than that used for much of the twentieth century. Its development has also coincided with a wider appreciation of archaeological sites and monuments as features of wider landscapes in the past and in the present day. In the context of this project, the identification of features of interest is focused principally on the avoidance, or mitigation, of features that may be impacted upon by new development. The intention is to identify sites, but not to analyse them beyond identification and initial classification. Subsequent projects may enable a more detailed investigation of any sites identified, and along the way enable new insights into the archaeology of a particular area or region.

A challenge here is how one uses the technological opportunity presented by ALS to recognise new forms of historical monuments and archaeological sites. The development of a machine learning method will, eventually, enable the identification of more sites of a kind that are already known. It provides for the identification of previously undocumented monuments, and an output may be the identification of 'new' ringforts, barrows or church sites. There is an inherent danger that the approach would exclude the discovery of new monument types, not just the discovery of 'old' monument examples.

There are sites that have yet to be recognised, while other places of human activity never existed as a tangible 'site' in the first instance. With this in mind – a realisation that we (a) do not have a complete picture of the past, and (b) we may lack the ability, at least at present, to recognise many aspects of the past – seriation can become the end in itself, creating a catalogue with little or no idea about how the objects relate to the past or what other contexts they had beyond place of discovery (c) a 'site', or locus of activity, doesn't need to leave a strong physical trace and certainly not a monumental one. The simple gathering of data is not sufficient in trying to understand the past more fully, nor does it assist in successful long-term management. For example, there may be several tens of thousands of ringfort sites in Ireland. This in itself does not necessarily explain why they should be considered as important; this is only achieved through the detailed investigation of this monument type, with special consideration of its attributed meanings (see Darvill (1993, 7)).

Current systems of classification frequently lead to the compilation of lists of sites, and increasingly the provision of mapped data that indicates the presence of a feature of interest. A potential cause of concern is that this may lead, especially in the context of regulatory activities, to the creation of a 'paper landscape' (as described by Andrews (1975)). Today we are surrounded by an evolved physical landscape that has upstanding (and buried) elements from many different periods. Similarly in the past people were also dealing with an evolved landscape, one which they used and introduced further change. Increasingly, the recognition of different features in the landscape relating to different periods has led to archaeologists describing the landscape as a palimpsest, comparing the landscape to a document that has been written and over-written, to be deciphered and understood (Kostyrko and Kierszys (2019)).

ALS can be used to record and identify (or vice-versa) the shape of features on the ground surface. Without further analysis of the data, it cannot directly attribute function or interpretation to those features and will seldom 'see' the kinds of differences between individual monument classifications that enables the kinds of classifications noted above with regard to medieval religious houses or the retro-identification of sites that have since been destroyed.

There are, however, certain features of archaeological sites and monuments which are readily apparent as artificial structures in the wider landscape, and which can be readily identified, in terms of their form at least, from ALS surveys. These are:

- Enclosures - A ubiquitous monument descriptor, though a term that can encompass the majority of archaeological sites and monuments. An enclosure may be defined as an area of ground that is contained within a recognisable boundary, be that a structure, either artificial or natural, and which had a cultural function in the past. In many instances the enclosure is formed by a curvilinear or rectilinear feature such as a bank, ditch, routeway, fence or wall. An indicative list of features and monument types that may be considered to be a form of enclosure is provided at Appendix ii in this paper.
- Accretions/Conglomerations - A mass of material that may survive from a denuded solid structure, or the collapsed material arising from a former structure.. Certain enclosures may also be considered as accretions/conglomerations or comprise elements of accretions/conglomerations.
- Topographical features that ALS can help identify rapidly which are important in the identification of previously unrecorded sites and monuments. They potentially

indicate places that were attractive in the past for certain types of human activity, and where archaeological remains may be found.

Following examination of monument morphologies and published ALS data, one may observe that there are four principal monuments, often (but not exclusively) composed of earthworks that are most readily identifiable through ALS, though as noted above not every monument is the same:

1. Ringfort;
2. Barrows;
3. Ecclesiastical enclosures; and
4. Settlement clusters including deserted medieval villages (DMV)

In an Irish context, Newman (1997, 156-170), see Figure 21) has presented an overview of the most common barrow forms in Ireland. The current Class List Definitions provides short descriptions of eight classifications of barrows (see Appendix iii below). More recently, Needham and Anelay (2021) have provided an extensive overview of barrow morphologies. While the focus of their studies is outside of Ireland, their discussion of monument morphology, including the acceptance of certain other monument classifications as barrows, is insightful (*ibid.*, 438-470). As shown in Figure 22 and Figure 23 they present a series of schematic representations of the kinds of barrows that they encountered in the Rother Valley region, in the Weald, lands in the south-east of England. The schematic representation of 'typical' barrows contrasts with the surveyed and excavated examples (see for example *ibid.* Figure 5.21, page 102), where the general characteristics of the monument may be observed, but where those characteristics have been impacted upon by degradation of the site over the centuries.

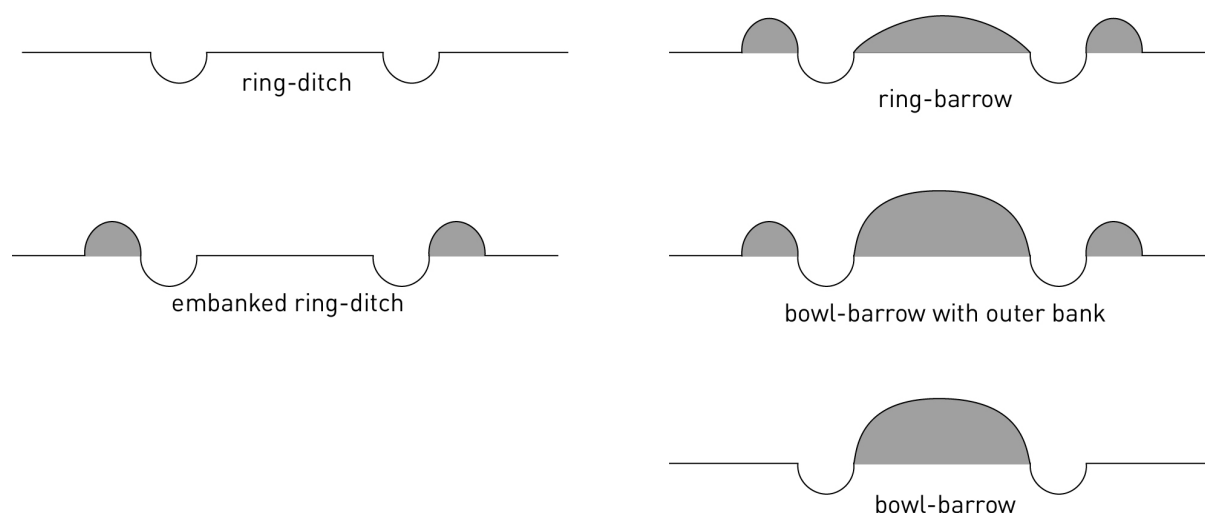


Figure 21: Example morphologies of barrows as evidenced at Tara (after Newman, 1997)

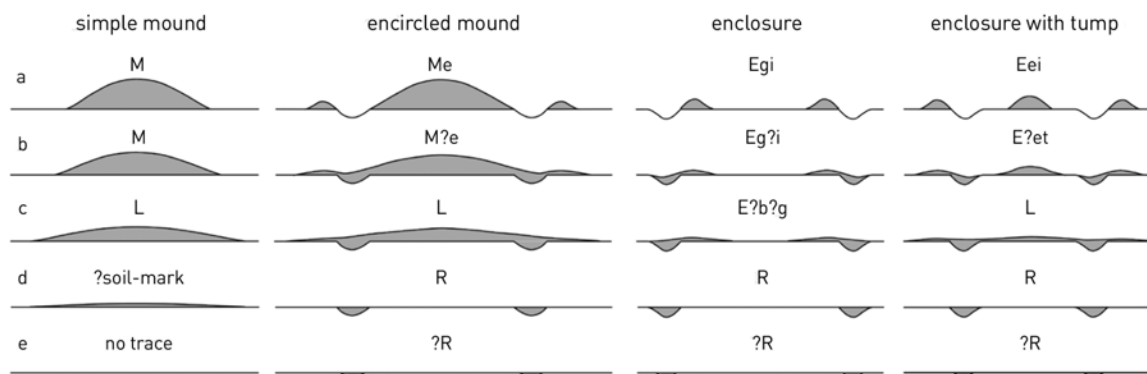


Figure 22: The effects of progressive denudation on selected monument types including mounds, enclosures and similar barrow forms, demonstrating how this may affect classification: a) denudation only by weather; b) early stage of attrition through activities such as ploughing; c) advanced stage of attrition; d) upstanding elements levelled; e) subsoil reduction and truncation of sub-surface features. Notations: M: Mound barrow, upstanding earthwork when first recorded. E: Enclosure barrow, upstanding earthwork when first recorded. L: Very low mound when first recorded (generally less than 0.35m maximum height). R: Ring-ditch or circular patch when first recorded. U: Unclassified: mound (UM), very low mound (UL), enclosure (UE), or unknown (U); unclassified sites are those not accepted as barrows on current evidence, grades 4 and 5 above; they do not contribute to analysis. (after Needham and Anelay, 2021).

Ringforts, ecclesiastical enclosures and sites associated with Early Medieval Settlement have received considerable attention in recent years as a result of dedicated research projects (see in particular O'Sullivan *et al.* (2014); Ó Carragáin (2021); Ó Carragáin and Turner (2017). The primary morphologies of these sites are discussed at length by O'Sullivan *et al.* (2014, 47-57, 145-148, 167-175), though the primary elements of site morphologies generally accord with the Class List definitions used by the Archaeological Survey of Ireland (see Appendix 3). Ongoing research (e.g. Harney (2020a); (2020b)) is presenting new information and understanding about the precise morphology of individual sites, their dates and construction histories, especially in relation to ecclesiastical enclosures.

It should be noted, however, that on occasions monuments of the kind described above are so substantial that the structural elements – be that a bank, mound or ruinous structure – may well be filtered out of processed digital models, since they may be confused with modern buildings, however this should be eliminated by adopting a DFM as the relevant surface from which to derive the training models. The analysis of both digital surface or elevation models and digital terrain models for the same survey areas is important, but may yield different kinds of results, especially where upstanding buildings are excluded from the models produced.

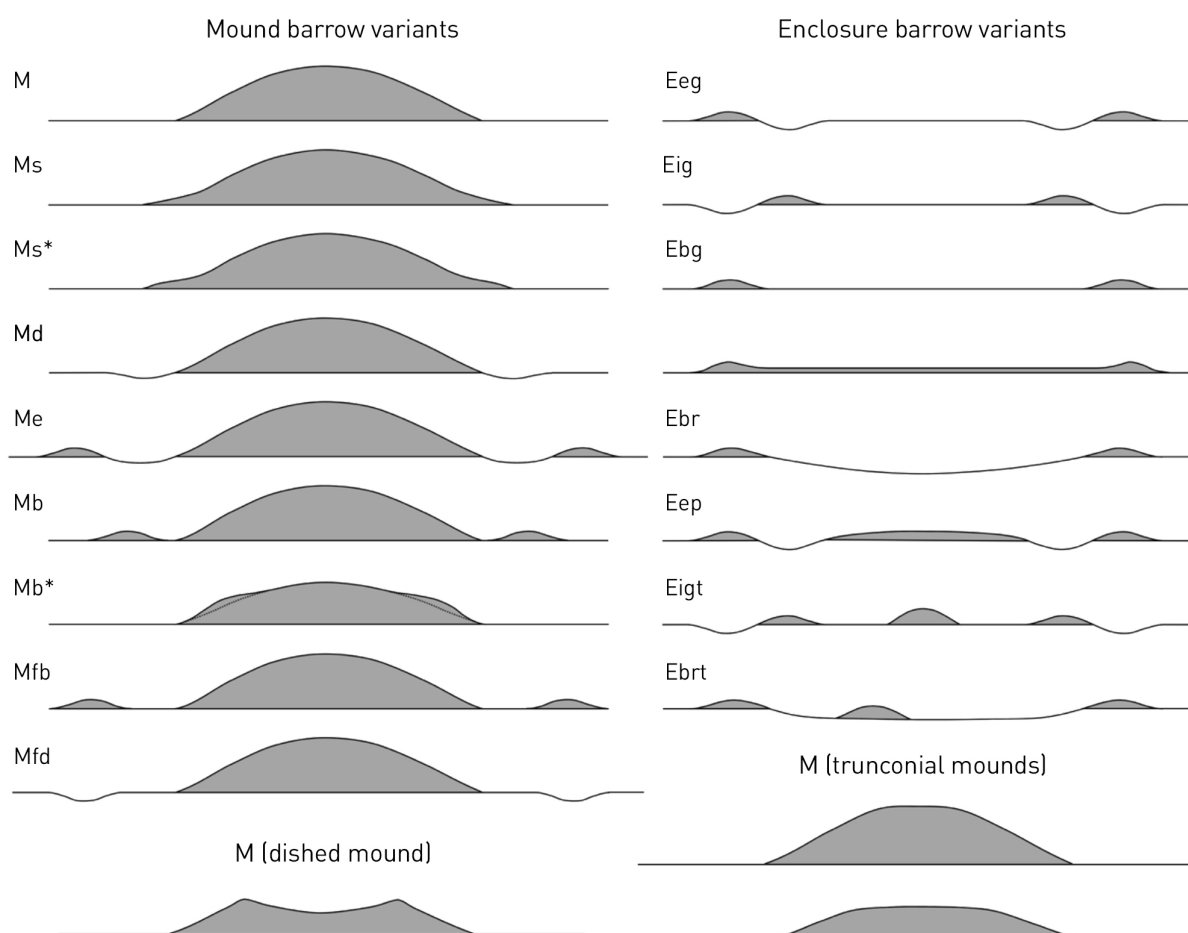


Figure 23: Profiles of mound and enclosure barrow variants. Notations: M: Mound barrow. E: Enclosure barrow. b: encircling bank; b*: embanked slope; d: encircling ditch; e: external bank and ditch; f: a flat, near-level annular zone - or berm - encircling the foot of the mound and enclosed by a ditch and/or bank; g: interior ground surface; s: skirt; s*: step. (after Needham and Anelay, 2021).

10. The use of ALS for archaeological survey

The use of ALS for archaeological field survey has been recognised as one of the most significant developments in such survey in the past 20 years (Corns and Shaw 2010) (Shaw and Corns 2011). ALS surveys have the potential to quickly gather significant volumes of new information about large areas of the landscape, which in turn has the potential to increase the number of archaeological sites and monuments recognised and recorded in the wider landscape. Fundamentally, the use of ALS data presents an opportunity for more expansive examination of the wider landscape beyond 'dots on maps' that are inevitably generated through monument-centric classifications.

As articulated in the tender information provided by the Project Sponsor for this project³,

archaeological examination of ALS data is increasingly used on TII road projects to reduce archaeological risk through the early identification of previously unknown, potentially significant monuments. In particular, risk can be reduced through the early identification and avoidance of monuments of potential national significance i.e., those that could be considered National

³ Transport Infrastructure Ireland TII268 Services Contract Lot 5 2021

Monuments under the definitions of the National Monuments Acts 1930–2014. The optimum time to undertake this examination is during Phase 2 Options Selection – the process of selecting a preferred route option – when there is maximum opportunity to avoid potentially significant monuments. However, current experience shows that it is often not feasible to undertake the archaeological examination of ALS data at this stage of a project due to the extents of the Phase 2 study areas, programme constraints and the limited capacity of the private sector to provide this service. As a result, archaeological examination of ALS data is more commonly undertaken on the preferred route only, during Phase 3 Design and Environmental Evaluation. This considerably increases the archaeological risk to a project as there is limited scope to avoid potentially significant sites at this stage.

The use of ALS datasets is an important and positive development in considering the potential impacts of new development upon archaeological remains. It moves away from simply looking at individual sites and monuments that have been recorded thus far and considers the archaeological potential of the wider landscape. Critically, it can assist in the identification of partially upstanding monuments and features that have otherwise gone undetected in previous field surveys, including monuments that present very limited surface indication at ground level. Not all of these sites or features will coincide with substantial buried remains, but in many instances they will; partially buried ditches, for example, may have little or no surface indication, but as reserves of below-ground remains may be both substantial and contain waterlogged deposits with high degrees of archaeological preservation of organic remains. When used with other remote sensing techniques such as geophysical prospection and aerial survey, ALS can provide new evidence about the nature, extent and relationships of features which require evaluation in terms of their significance. That new evidence can also facilitate a re-evaluation of the significance of other previously recorded sites or site types elsewhere.

Interrogation of the data arising from the surveys does, however, require an empirical gathering of information to describe the features identified. This can, on occasion, result in the formation of catalogues of features observed from or inferred by the data which ultimately become 'dots on maps' for ongoing use in the context of spatial planning. The processed data can, however, provide powerful visual representation of features in the landscape. When interpreted, that data will enable a fuller understanding of either the quantum of archaeological remains within any given area, or the likely survival of buried features in that area, or both.

10.1. Limitations and Bias

While the results of ALS surveys are often stunning, in terms of the quantity of information they provide or our realisation of the extent of archaeological survival revealed by the data, the methodology is not in itself a 'silver bullet' in terms of revealing the full extent of archaeological survival in any given area of land. There are three key limitations to what the survey methodology can provide:

1. *Subsurface features:* The methodology can only detect features which have a physical surface expression, that is, it maps the topography of the ground surface and can be used to identify elevated or sunken features on the ground surface. Processing tools, including the use of relief shading, sky-view factor and simple local relief models, can be used to give prominence to those features in an effort to make them more visible and 'legible'. The technology does not permit analysis of underlying features cut into the earth unless these have surface expression; geophysical prospection techniques (resistivity, magnetometry and GPR) and the identification of crop and soil marks from aerial prospection can detect such buried features and may frequently complement the data gathered through ALS survey. This is particularly the case in areas of intensive arable agriculture, where the present ground surface has been rendered flat as a result of heavy ploughing. Such ploughing activity is also likely to have had a negative impact on any buried remains that previously survived, but it is also possible that archaeological horizons and subsoil features will survive below the ground surface. Furthermore, intensive modern arable agriculture has the effect of levelling large areas of the landscape, removing any upstanding features or reducing those features to very slight remains. The banks of a ringfort that may once have stood over 2m high may be reduced to just a few centimetres.
2. *Nature of vegetation cover:* Ground cover and ground use is a significant element to be considered in the processing of data. While ALS can be frequently used to gather information about the ground surface that is covered in vegetation, the nature of the vegetation cover alongside the resolution of the survey (how many data returns are gathered per square metre) will affect the results obtained. This is not in itself a problem per se, since many shapes and patterns will still be observable, but it will affect the resolution of ground mapping that can occur. For example, a field with a dense sward of grass or cereal crop may present features such as parch marks that can be observed through aerial photography, and the features themselves may have some topographical expression. However, the resulting point density following processing of the ALS data may be such that the ground surface beneath the vegetation may not be of sufficient resolution to be able to identify the upstanding remains. In areas of specific high-density crops such as oilseed rape ALS the recorded points normally only model the top of the crop which may display a reflection of the underlying topographic surface, but other factors such as soil fertility and overall crop health could affect the height of the crop also.
3. *Continuous and close-contiguous features:* As noted by Gardiner (2018, 122) ALS is more effective for identifying certain features over others: earthworks may be evident on the ground, on aerial photographs or particularly visible on processed ALS data but

features such as stone alignments or settings are frequently much more difficult to identify, even if they are obvious features in the landscape. Continuous and close-contiguous features are often easier to identify through visual analysis of the processed data. These attributes tend to form more readily recognisable patterns in the ALS imagery, and so the sizes and proximity of those features will influence how they are perceived; larger features that are more than 1 square metre in size will always be more visible than smaller features in the processed data.

In general, therefore, one may state that ALS is most effective for identifying earthwork features greater than 1 square metre in extent, and which may survive as linear or discrete features on the ground surface. Analysis of land less than 1 square metre in extent can still be useful, provided there are sufficient data entries for that and adjacent ground which can form a coherent topographical representation of the ground surface. For example, a surface depression of 0.75m x 1.5m in area, 0.2m deep, in an otherwise featureless field may not appear to be particularly significant; the same feature within a discrete area of land with multiple repetitions of the same kind of features is more readily identifiable as significant. In either instance, the association of such a feature with other topographical features is important; in one instance other features may give rise to an interpretation of the feature as a pit; in another circumstance it could be interpreted as a grave, set within a collection of other graves.

The example above emphasises the shape of a feature that can be recognised through analysis of ALS data. The processed data can be used to identify patterns that may be considered anomalous compared to the natural topography of a place, and therefore features of potential archaeological and cultural interest. While there are limitations to what the ALS data can be used to identify, there are also likely to be biases in how the data was gathered and the interpretation of that data, based upon previous knowledge or perceptions about a place and the 'likely' archaeological features that may be found there.

The reasons why the data has been gathered, and the resolution of each survey, varies from instance to instance. In capturing large areas of landscape, the surveys cast a wide net, within which there will be different kinds of monuments that have been recorded in the past or remain to be catalogued. In this regard, much of the rationale for such data capture may be perceived to be 'neutral'; it gathers in a random collection of data within a particular area. This is correct, insofar as the data is inherently just that: data. It will be constrained by the methodology through which it was acquired and processed, and by other factors such as ground cover.

However, the area selections are not 'neutral'; they are deliberate selections, for a range of purposes, which do not always cover all of the landforms, topographical situations or drift geology, settlement or use. For example, data captured for the purposes of selecting a route for a new road will tend to focus on areas that link modern settlements, and therefore likely to be areas of continuing activity in the countryside, lands that are subject to agricultural activity or settlement. Surveys are unlikely to include significant areas of uplands, mountains or areas remote from modern population centres, and therefore it is unlikely that present records will include many examples of monument types more commonly associated with those more remote areas. That is not to say that monuments found in upland or remote areas were not previously built in lowland zones or lands which are intensively settled and used today, but to note that the subsequent history (the local taphonomy) means that such monuments are less likely to survive as upstanding monuments.

The area of land included within ALS surveys will have their own boundaries, connected to the geographical area of interest that spurred the survey in the first instance, or the limits of survey set for other reasons (such as affordability at the time the survey was commissioned). It is likely, therefore, that at least some monuments will straddle boundaries between surveyed and not surveyed areas. This means that, when considering individual monument classifications for use in a machine learning method, one should not include sites for which the survey data is incomplete. One may counteract this, to a degree, by use of buffer zones on the edges of survey areas, though this can produce the unintended consequence of limiting the number of individual monuments available for inclusion.

While ALS offers one the ability to produce recognisable 3D renders of the landscape, new visualisation and analysis methods offers a fundamentally different conceptual way of looking at the landscape. Landscape studies are generally rooted in theoretical and philosophical discourse, common to historical geography and increasingly in archaeology. This is noted in passing above, with regard to the work of Aston (1985), Rowley (1974), Tilley (2005) (1999) (1994) and Thomas (1993) and is likely to become a more significant component of archaeological analysis in the future. The step towards a machine learning model and use of computer algorithms to interrogate the landscape is a further development of that conceptual approach, and one which has potential for unforeseen consequences.

In the context of landscape studies, the use of ALS data may most readily fit into a discourse of Mitchell's and Henderson's 'apocryphal landscape', considering the "landscape as a way of seeing, especially a way of seeing that relishes the gaze, that asserts power by privileging perspectival vision, which, far from being a mere way of seeing, informs the actual, material making of places" (Mitchell (2003, 790), in part citing Henderson (2003, 192). Processed ALS data is a very particular kind of privileged perspectival vision, open to the kinds of manipulation noted above. Archaeologists are perhaps more comfortable with the concept of landscape as "the material revelation of human practice and thought" (*ibid.*), the 'epistemological' landscape, whereby the search through data gathered via ALS surveys can be used to illustrate change and development in the landscape, what people did, where and, through examination, why. Landscapes as 'palimpsests' can be decoded to achieve this 'epistemological' landscape, and ALS can be used for that decoding and analysis. In this process there is a more passive engagement with the landscape, or at least it can appear that way; as more information is gathered the more one may theorise about the relationships of features recorded. To give an example, a small rectilinear enclosure set at the end of series of linear or sinuous/curvilinear features may be interpreted as a plot of land that demarcated a place of dwelling, situated on top of the end of a selion or cultivation ridge. These cease to be a coincidence of features identified on a survey of an area of land, but becomes evidence of a change in landholding and agricultural practice at a particular point in time that is reflective of wider social, political and economic change, as suggested by O'Keefe (2018, 106)) with regard to the interpretation of ALS data at Newtown Jerpoint, County Kilkenny.

It is unfortunate, perhaps, to associate the use of developing a machine learning model with a discourse on an apocryphal view of the landscape or landscape studies. This is not an apocryphal 'rise of the computer' to take away human study of past human activity. Rather, it is to note that there is an inherent danger that one may expect too much from the machines or computers. In giving the data an elevated importance for archaeological survey, one may similarly give the computer an elevated position in terms of the outputs generated. Just as Walsh (1993, 101) noted with regard to the initial classification of monument that "*we must*

beware also that when comparing monuments to classify them, we do not impose patterns and thereby infer a relationship where, in fact, none existed", we must similarly take care in the use of computer-aided survey and especially in the development of algorithms to support that survey that we do not inadvertently create the circumstances where computers generate patterns and inferred relationship where none existed.

The use of deep learning algorithms requires the learning – or imposition – of patterns, the use of formulae that when repeated are also refined. Where this is used to identify features of archaeological interest, there is an inherent danger that in due course it is only those features identified through the algorithm that are seen as being of archaeological interest. Three challenges or questions thus arise:

1. Are the patterns or formulae valid as reliable indicators of features of archaeological interest?
2. Why would an algorithm discard or ignore other features of archaeological interest?
3. Why would the features identified through a computer algorithm be accepted more readily than those identified by a human?

With regard to the first two questions, it needs to be explicitly noted that these outcomes depend upon the inputs, the human interaction that formulates the formulae. The output will relate to the question asked, and human limitations to this. With regard to the third question, one pre-supposes that this would be the case, whereas this may not be what will happen. It is, however, a legitimate question, since the usefulness of the machine learning method in some respects anticipates that the machines would be better and more efficient at the process of site identification than people. At this stage, I would very strongly argue that this is not the case, but that machine learning models may help speed up the process to assist people in interrogating the data, to identify features of interest more quickly, and enable thoughtful consideration of what the features may be and provide a level of consistent focus and observation which may vary in the human inspection model.

The frequency of use of the algorithms, the size of the datasets upon which they rely, the frequency of updates for those datasets, and the accuracy of the information held on those datasets will all impact upon the usefulness of a machine learning model approach. This approach does not substitute machine for human interrogation of the data but is intended to enhance the human interrogation of the data by drawing attention to key features in the landscape that are of interest and provide an efficient method to processing the large volumes of remote sensing data which can be utilised within landscape archaeology studies.

The recognition of archaeological sites and monuments and their ongoing survival in the landscape means that they will continue to have a role in the concretisation and making of memory, as tangible places where the past can be witnessed, explored or understood. If machine learning model can be developed in respect of field monuments in Ireland, it will take on some of the role of memory maker for archaeologists and the wider public. It will enable the identification of features similar to those already known, with an automated function of highlighting key spatial information about the features in question. One may also expect that this technology could be used to revisit previously recorded sites, to make further records of those sites. The third question above, about the acceptance of features identified through computer algorithm rather than human recognition cuts to the heart of a question that will arise: whose memory is more accurate? Additional challenges may also arise based upon the acceptance of a new SMR based solely on remote sensing data and the potential time lag

between the discovery of a new monument within ALS data and its official update within the SMR.

		known by archaeologists	unknown by archaeologists
known by non-archaeologists	OPEN AREA	These are monuments that have been identified in the SMR and are well known by the wider public. It may include major or public monuments (i.e. State Care monuments), but may also include other sites that are not 'public' monuments per se but which are well known by the wider public.	In the context of State-led management of the historic environment, sites that are known by the wider public but not known by archaeologies may be considered to be blind spots. This assumes that both archaeologists and non-archaeologists use the same measurements for what they consider to be a monument. A good example is a ringfort known by a local community but not yet recorded. Such monuments can become 'Open' sites through the sharing of knowledge from non-archaeologists to archaeologists.
	HIDDEN AREA	This class of description shares some similarities with the 'Blind Area' monuments noted above, but this time retains an air of privileged or 'secret' knowledge about an individual site type or class of monument. Such monuments can become 'Open' sites through the sharing of knowledge from archaeologists to non-archaeologists.	UNKNOWN AREA
unknown by non-archaeologists			

Figure 24: A Johari Window model of historic monument recognition

In 2002 Donald Rumsfeld, then the Secretary of Defence in the United State of America, gave his 'known knowns, known unknowns and unknown unknowns' response to a question about the progress of conflict in Iraq, terrorism and weapons of mass destruction.⁴ The quote brought to the fore concepts of dealing with data and knowledge, and while it has been frequently mocked for comic effect it could equally have been used to describe what we know about the survival of field monuments in Ireland. Figure 23 below is a modified Johari Window, based upon this concept of known knowns etc.; while ALS is not in itself a heuristic self-help system, at times it shares some similarities. It can be used to facilitate a rapid identification of sites (the 'knowns' and previously 'unknowns') which is, practical but not guaranteed to be perfect, useful for reaching an immediate or short-term goal of identifying sites in the context

⁴ https://en.wikipedia.org/wiki/There_are_known_knowns

of road route-planning. As the discussion above should explain, it is not a perfect method, but it is useful. It is short-term insofar as it in part addresses an immediate need in this context around route selection but does not deal with the management of any features outside of that context.

Beyond this project, other discourses on landscape as identified by Mitchell (2003, 790) can make extensive use of ALS data, including the discussion of landscape in the context of a territory or political space, and the discussion of landscape as social space in or through time. Each discourse overlaps, but each requires an understanding of the landscape, its archaeology and the spaces between individual sites and monuments.

Ultimately, the interpretation of the data gathered will be subject to the conscious or unconscious bias of the person making an interpretation of features identified in the survey. Monument classifications are seldom intended to be so proscriptive that one cannot consider other potential interpretations of a site, either in terms of its overall composition or history, its primary and secondary uses or contemporary understanding. However, classifications are intended to be positive identifiers of a monument as one thing or another. And as a human judgement there will be times when an error is made, usually in the classification of a site as one kind of monument or another, but occasionally in the classification of a feature to be of archaeological interest that transpires to have no tangible archaeological remains. The ability to interpret the data for archaeological purposes is informed by experience and knowledge. It is emphasised here to note that the elements of human agency will influence the usefulness of the data and its interpretation, how it is gathered and what conclusions one draws from it. It is hoped that many features may be generally identified, however, more precise classification may require field inspection.

11. Applied use of lidar in Ireland on the study and protection of archaeology

11.1. Initial use of ALS in Ireland

The earliest use of ALS for archaeological prospection dates to the early 2000s with Ian Barnes' (2003) work at Salisbury Plain (i.e., the Stonehenge landscape). This, and subsequent surveys at Stonehenge (Bewley, Crutchley and Shell 2005) demonstrated the potential of the technology for both the identification of previously unidentified archaeological sites and for the rediscovery of sites previously considered to be lost/destroyed. While these early studies did not exactly open the floodgates for the use of lidar in archaeology, they certainly paved the way for other studies including the development of archaeological lidar work in Ireland.

The first application of ALS for archaeology in Ireland was the work of Shell et al. at Loughcrew (Shell and Roughley 2004) (Shell 2005). This survey comprised an area of c. 30 km² at an average ground point spacing of 60 cm, collected by the University of Cambridge Unit for Landscape Modelling. Initial publication of their prospection results was followed by a long hiatus, with subsequent publication of research on the open-air rock art (Twohig, et al. 2010) and most recently later field systems post-dating the initial capture by over fifteen years (Roughley, et al. 2021)

The Loughcrew surveys clearly demonstrated the potential of the method in an Irish context and laid the groundwork for a series of ALS-based research projects over the following few years covering some of Ireland's most iconic archaeological sites and landscapes. In 2007 the UK Environment Agency were engaged by Meath Co. Council to capture a large area of data centred on the Brú na Bóinne World Heritage Site (WHS). This extended well beyond the core and buffer areas of the WHS and arguably mirrored the earlier prospection projects at the comparable UNESCO WHS landscape of Stonehenge. Here, a DSM and DTM were provided ready-gridded at 1 m ground resolution. These data were subject to thorough archaeological investigation until 2010, although they were employed as a base layer for geomorphological studies in the INSTAR-funded Boyne Valley Landscape Project (Lewis, et al. 2009) (Lewis, et al. 2018); (Davis, et al. 2013). The Clare Island dataset, also captured in 2007 has also seen as much use as a geomorphological base map as for archaeological prospection. Despite its clear archaeological potential, especially as regards later ridge and furrow cultivation, this dataset has yet to see systematic prospection cf. (Manning, Gosling and Waddell 2007).

While undoubtedly valuable, the Loughcrew and Boyne surveys are characterised by relatively low spatial resolution and, crucially, by the non-availability of the raw ALS data. In the former surveys the data were considered under copyright to the University of Cambridge and so have remained unavailable, while in the latter the provision of the raw data was apparently not a requirement of the original data acquisition.



Figure 25: ALS derived DSM relief shaded model of the Brú na Bóinne landscape, enlargement around Newgrange passage tomb (Meath County Council/The Heritage Council/The Discovery Programme).

11.2. High-resolution surveys: Tara and beyond

The first high-resolution surveys specifically commissioned in Ireland for purely archaeological purposes were a series of experimental FliMap helicopter-mounted surveys undertaken by Fugro (now FugroBKS) and funded by the Heritage Council. These surveys, commissioned by the Discovery Programme (Corns, Fenwick and Shaw 2008); (Corns and Shaw 2008) (Corns and Shaw 2013) focused on relatively constrained areas in iconic landscapes: the Royal sites of Dún Ailinne and the Hill of Tara (the latter the focus of longstanding and ongoing research by the Discovery Programme – e.g., Newman (1997)) and the deserted medieval settlement of Newtown Jerpoint (Lambrick 2008). These datasets were captured at 40-60 points per square metre, primarily to test the efficacy and value of such high-resolution data capture. Arguably, as one of the least forested nations in Europe there are few landscapes where such high-resolution data would be a necessity, while conversely in areas of coniferous plantation there is no canopy penetration, even at very high data capture resolutions. However, while the previous surveys had brought ALS to the attention of archaeologists, it was these high-resolution surveys – especially of Tara - that began to bring the method to the attention of the wider public through the stunning visuals that such datasets make possible. Of the three sites investigated at the time only the Tara data has seen sustained and continued research. The Dún Ailinne dataset while showing the site in stunning detail revealed that there really was not very much topographic variation to see within the enclosure, and remains to be incorporated into ongoing research (e.g., Johnson et al. 2014)). The Newtown Jerpoint data has only recently been incorporated into a Teagasc-funded research project within UCD School of Archaeology

⁵ alongside large-scale geophysical surveys. These ALS surveys were followed by a similarly high-resolution capture of data at Ireland's other UNESCO WHS – Skellig Michael (Corns and Shaw 2008) (Corns and Shaw 2013) which again has served both as a source of spectacular imagery and an invaluable tool in site management and preservation (DoHLGH 2020).

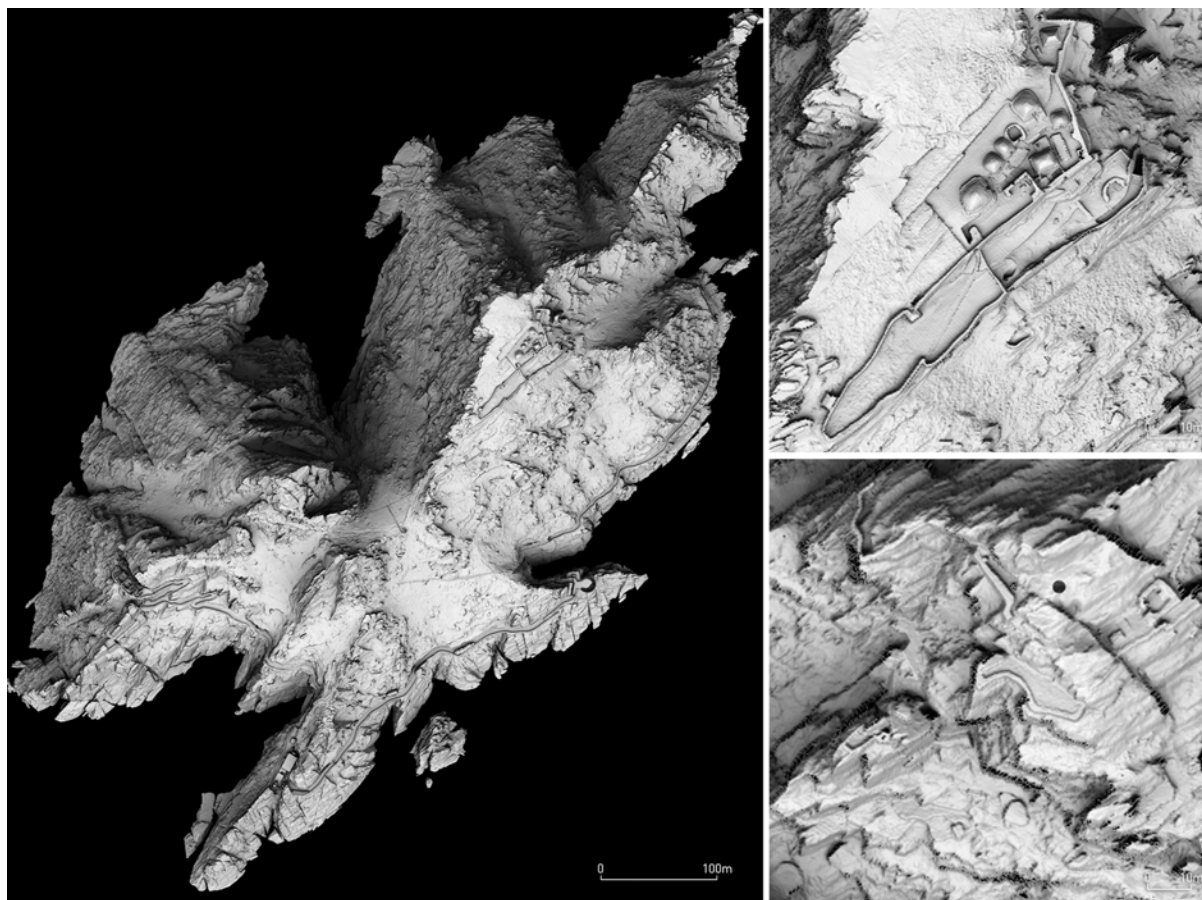


Figure 26: Sky view factor visualisation of FliMAP derived ALS data of the whole Skellig Michael with enlargements at the monastery and south peak (DoHLGH/The Discovery Programme). The use of high-resolution ALS surveys (commonly 20-40 points per square metre and occasionally more) in heritage research has become increasingly commonplace, and the commissioning of new datasets is now a frequent component of site management practice.

In the eastern part of Brú na Bóinne the private agrotechnology company Devenish commissioned FliMap data in order to assist in both in the archaeological management of their 'Lands at Dowth' and to allow carbon quantification on their farmlands. While this dataset represents a significant upgrade to the previous UKEA dataset, prospection revealed little that was not already evident from the earlier surveys (Ní Lionáin and Davis 2016). However, it has proved valuable in other regards, both in terms of quantifying above ground carbon and studying patterns of runoff within the estate e.g. Green et al. (2021). This represents an unusual situation where ALS captured for archaeological purposes has been repurposed to assist in land management, rather than vice versa.

In the last few years new datasets have been commissioned for Iniscealtra, Co. Clare (Clare Co. Council) – recently included on the candidate World Heritage Site list, - the Hill of Uisneach, Co. Westmeath (2020) another candidate WHS as part of the 'Royal Sites of Ireland' (Westmeath Co. Council) and The Curragh Plains, Co. Kildare (Davis 2021b). The significant

⁵ (<https://www.rte.ie/brainstorm/2022/0118/1274278-newtown-kilkenny-ireland-medieval-town/>)

body of hillfort research undertaken in recent years by Billy O'Brien and James O'Driscoll also incorporated a significant focus on remote sensing, including the use of ALS data (O'Brien 2017), (O'Brien and O'Driscoll 2017), as has research on the uninhabited island of Inis Airc, Co. Galway (e.g., Kuijt et al. (2010)). Meanwhile, the INTERREG-funded CHERISH project - a collaborative venture between the Discovery Programme, GSI, the University of Aberystwyth and the Royal Commission on the Ancient and Historical Monuments of Wales – has also funded significant new ALS collection, especially for the Welsh and Irish islands. This highlights the potential of ALS to inform policy as regards climate change and coastal erosion, a clear requirement of the national Built and Archaeological Heritage Sectoral Adaptation Plan (Daly 2019)

The new 'Community Monuments Fund' administered by the Department of Housing has offered a mechanism for funding the acquisition of such datasets which might previously have been beyond the operating budget of most local authorities. The future of small-scale data capture is likely to be through the deployment of unmanned aerial vehicle (UAV) ALS platforms. BlueSky currently operate such missions within Ireland, as do Western Aerial Surveys, and these have seen use in some community projects already. Paradoxically this may lead projects away from the current trend of large area landscape studies towards more site-based narratives.

11.3. ALS Data captured for non-heritage purposes

There are some significant datasets that have been captured within Ireland for reasons other than heritage that nonetheless have significant archaeological potential.

For several years, the primary provider of ALS data in the Republic of Ireland was the Ordnance Survey Ireland. The majority of these data were provided 'off the shelf' as DTM/DSM, gridded as a standard to 2 m or upon request at 0.5 m ground resolution. These data were used in a number of research projects including at the Hill of Ward (Davis 2011); (Davis 2013), New Ross, Co. Wexford (Marshall and McMorran 2016), the Black Pig's Dyke, Co. Monaghan (O'Driscoll, et al. 2014) and in the Lough Gur area (Motherway 2012). Most of these data were captured in the mid- to late-2000s and are, at this stage, showing their age. That is not to say that they are not of value. In fact, some of their value might lie in the fact that they represent a historical record of topography from over a decade ago, including areas that have since been lost to erosion or development. OSI policy generally prohibited the provision of raw las data except in the case of research agreements (e.g., surveys at Rathcroghan (Dempsey 2016)); in fact much of the data captured by OSI have never been processed – this was usually undertaken only when data were purchased – and it is hoped that these datasets which currently remain unavailable might be brought into the public domain in the coming years.

Transport Infrastructure Ireland (TII) captured significant noise mapping datasets along many of Ireland's national road corridors (Davis 2012). These data, flown by Fugro, have been available as unprocessed xyz .txt flightlines and as 2 m raster DTM for a number of years but have seen limited exploration, in part owing to the limited nature of the data themselves: because of the rationale for their capture, they represent narrow (1-1.5 km) strips of data centred on major road networks. There has been no systematic treatment of these datasets and while they are clearly not so valuable as more recent large area surveys, they still possess considerable untapped archaeological potential.

Similarly, the Office of Public Works (OPW) have collected data over a number of years over multiple phases. Initial surveys comprised both coastal data and narrow corridors focused on the Rivers Suir and Lee. These early datasets were only provided at 1 m DTM at best, with the coastal data at 2 m and frequently of too narrow strips to be especially useful archaeologically. The Suir dataset was the focus of significant PhD research, focused on Anglo-Norman agricultural systems (Corcoran 2018) and clearly demonstrated its value beyond the original capture objectives. Later phases of OSI data focused on a large number of areas deemed liable to flooding issues. OPW have also recently commissioned large-scale coastal surveys in the west of Ireland which are likely to feed directly into debates on coastal erosion and climate change. Unlike the earlier data these were available as classified .las files. Again, some of these data have been incorporated into PhD research, focused on counties Roscommon and Monaghan by Curran (2013); (2020) alongside some more extensive TII datasets (see below). The scale and quality of these OPW datasets makes them a potential goldmine of underexploited heritage data.

In Northern Ireland a similar programme of river basin ALS data was collected between 2004-2014, in addition to some limited road corridor collection. The majority of other data collected in Northern Ireland has been specifically in relation to managing archaeological heritage, with predominantly monument-centred datasets collected between the years 2008-2014 to assist with the interpretation and protection of archaeological landscapes. More recently data were commissioned by the National Trust NI, with a similar objective: to help in the management of their Mount Stewart and Castle Ward properties (<https://www.nationaltrust.org.uk/mount-stewart/news/800-year-old-norman-motte-uncovered-at-mount-stewart->). As in the South, some of these data have been repurposed to good effect, providing significant research benefit. This has ranged from recent large-scale surveys at Navan Fort (O'Driscoll, Gleeson and Noble 2020) and Clogher, to significant prospection projects such as those by McNeary (2011) (2014).

Most recently the Geological Survey of Ireland (GSI) have also captured large areas of data especially in counties Galway, Roscommon and Sligo. These data are currently available through the Open Topographic Viewer as 1 m raster files and again have seen no systematic exploration for archaeological purposes. Given the success of previous surveys in Co. Roscommon this is likely to be an exceptionally rich archaeological area and certainly warrants further investigation.

Other state and semi-state bodies within Ireland certainly use and commission ALS data, but do not generally make their data available. These include the likes of Teagasc, who commissioned data for their hedgerows research and agricultural catchments programme (Green, Martin, et al. 2019) (Shortle and Jordan 2017)) and Bord na Mona who use such data to monitor peatland extraction. It is unfortunate that significant bodies of data remain within such organizations and are unavailable to the heritage community.

11.4. ALS for infrastructure in Ireland

The first application of ALS as part of infrastructural development in Ireland was the assessment undertaken by Megarry (2010) (2011) as part of the N4 Carrick-on-Shannon to Dromod road scheme. This initial prospection focused on the proposed preferred option corridor and a buffer zone of 50-100 m around it (Megarry 2010, 5); in total, 64km² of the total 140km² was assessed. This assessment made use of existing OSI ALS data, supplied as 2 m

gridded rasters. The N4 report built upon the work Megarry (2011) had undertaken on the Brú na Bóinne ALS dataset as part of the 2010 Boyne Valley INSTAR project, working in a grid wise fashion and using multiple visualizations, although some of these were very novel at this time ((Hesse 2010); (Kokalj, Zakšek and Oštir 2010); (Kokalj, Zakšek and Ostir 2013); (Hesse 2012); (Kokalj and Hesse 2017)). As a geographic location the selection of rural Roscommon was fortuitous: subsequent research has demonstrated that this is an exceptionally rich archaeological area that is ideally suited to ALS-based prospection. Other areas with more significant arable cultivation (e.g., Wexford; eastern Meath/Louth) might well have proved disappointing and consequently slowed the adoption of ALS-based prospection by TII.

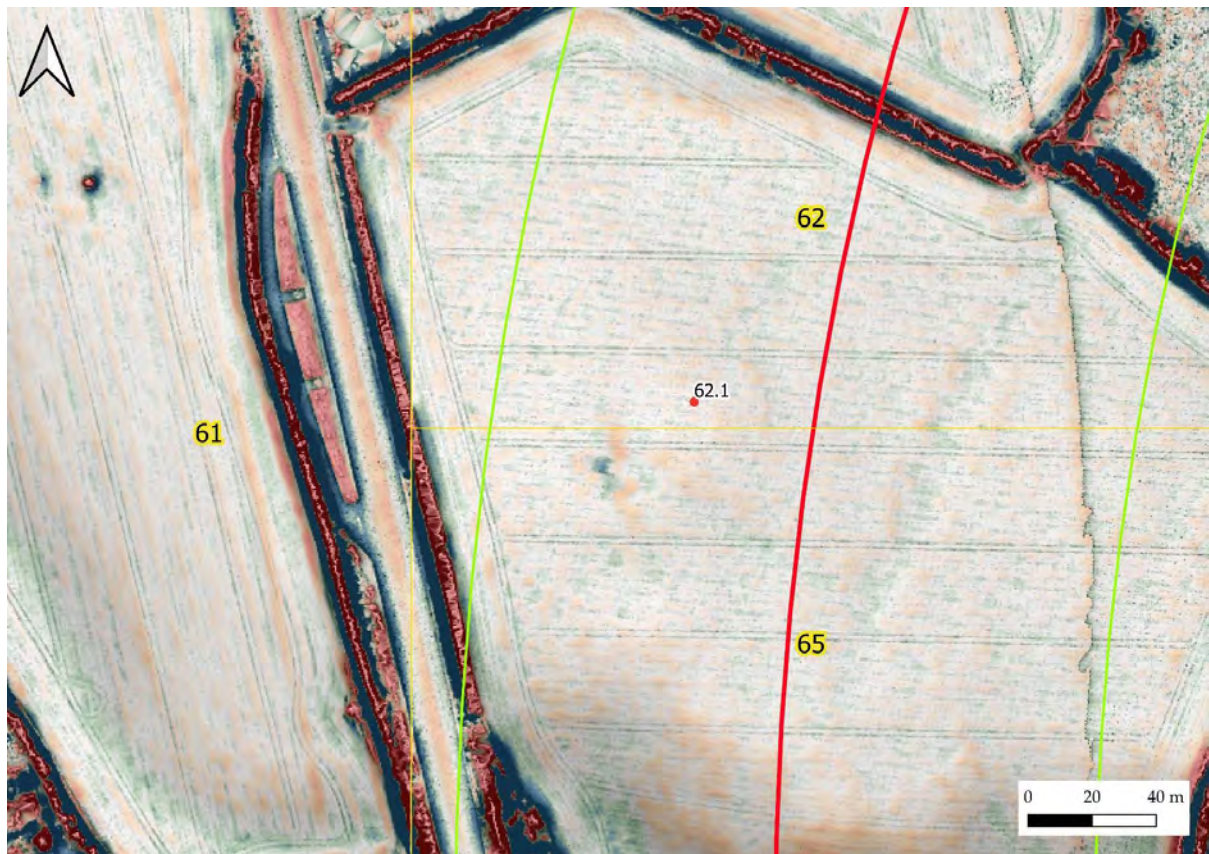


Figure 27: Example of a potential low-profile enclosure (62.1) identified as part of the N11-N25 Oilgate-Rosslare Harbour scheme archaeological assessment of ALS data. Visualisation scheme used LRM over SVF (Davis 2021b).

The success of the N4 Carrick-on-Shannon project led to the first commissioning of data specifically for infrastructural work. Once again this was based in Roscommon – the N5 Strokestown-Ballaghaderreen road scheme (Davis 2015); (Davis 2016). On the one hand this was close enough to the Carrick-on-Shannon scheme that it was deemed to be of similar archaeological potential to the earlier project; on the other the scheme also ran close enough to the Rathcroghan complex that a highly detailed prospection survey was a necessity. Once again, the ALS-based prospection revealed a host of potential archaeological features that informed later stages of field survey and testing. By this point the method had become established as potentially a valuable part of road construction, in particular within Phase 3 – post route selection.

Not all assessments have proved so informative: for example, in the recent N11-N25 Oilgate-Rosslare Harbour scheme (Davis 2021a) there were very few potential new sites identified (only 13 'high confidence' sites in a survey area of c. 60 km²) and the majority of the recorded

structures had also been levelled by centuries of pastoral agriculture. To an extent we are yet to really understand the degree to which monument distributions are artefacts of past agricultural practice or reflect real archaeological patterns. This could be better informed by incorporating a feedback loop, whereby a post-survey/excavation assessment of the ALS-based survey is incorporated, providing critical information on where this prospection is succeeding and where it is not.

There is an increasing trend currently towards large area data capture. This is, in part, owing to the demonstrable potential of automated approaches (e.g. (Sevara, et al. 2016); (Due Trier, Cowley and Waldeland 2018); (van der Vaart and Lambers 2021)). It is to be hoped that in the foreseeable future at least a preliminary assessment of these datasets might be automated. A second reason for large area capture is the potential to move prospection earlier in the planning process (again, coupled with automation) to inform route selection. In the recent Virginia Bypass scheme, a total ALS area of c. 120 km² was captured with archaeological prospection only required for an area of 20 km². Similarly, in the recently tendered Cahir-Tipperary Road scheme the total dataset comprises c. 320 km² while only 37 km² is to be prospected.

To date, the overwhelming majority of ALS studies in Ireland have essentially used the data as a 2-dimensional image or series of images. The 3D nature of the data is used for site profiles (usually requested for existing SMR sites) but there have been few uses of ALS for viewsheds or least cost path for example. Exceptions for this have been both in the N5 Strokestown-Ballaghaderreen scheme (Davis 2015) and the N2 Slane Bypass ALS assessment (Davis 2018). In both of these cases there were concerns regarding visual impacts on significant heritage landscapes – respectively Rathcroghan and Brú na Bóinne. In the latter case a number of alternative route options were examined. This perhaps foreshadows the incorporation of ALS assessment at an earlier stage of road planning.

11.5. Future uses of ALS within Irish archaeology

In 2018, the 'Open Topographic Data Viewer' was launched, an online data viewer and portal that acts as a repository for freely available ALS datasets within Ireland. These include the existing OPW flood relief data, GSI data, some TII data (TII noise mapping data), Discovery Programme datasets and some Co. Council data (e.g., Hill of Uisneach). At present these are provided as raster only: this means that in some cases (e.g., TII noise mapping data) these models are 2 m ground resolution only. An obvious next step here would be making the raw data freely available. The free availability of historic datasets is also becoming more of a possibility. While these are dated, they still have enormous archaeological potential. In fact, the mid-2000s capture date actually provides clear baseline data from which to explore long-term landscape change.

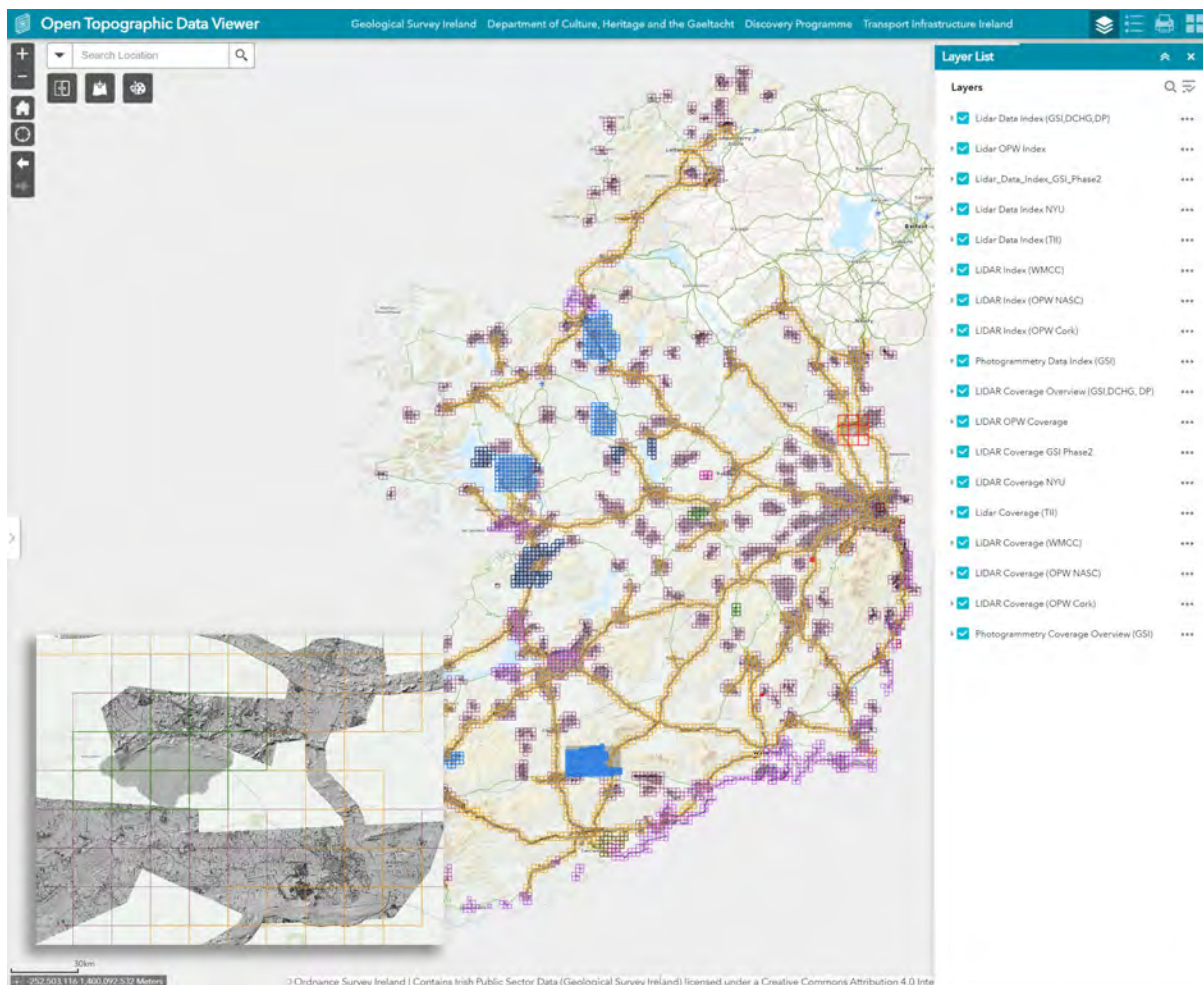


Figure 28: Screen image of the GSI Open Topographic Viewer. Inset highlights three separate organisational ALS datasets (OPW NASC/NPWS/TII) across the same region of County Offaly.

(<https://dcenr.maps.arcgis.com/apps/webappviewer/index.html?id=b7c4b0e763964070ad6>)

Increasing data capture areas are clearly becoming more common, and we are now facing a situation where the amount of data exceeds our capacity to prospect it. These large area datasets are increasingly making it possible to consider landscape-scale narratives of monument distribution beyond the narrow corridors previously available. Numerous countries have moved towards national open ALS coverage (Kakoulaki, Martinez and Florio 2021); this is not as yet planned for Ireland; however, national open data has been suggested as part of the Northern Ireland Sustainable Agricultural Land Management Strategy (DAERA 2018) and is becoming increasingly commonplace.

To an extent this data overload demonstrates the need for the current project, and further research focused on automated feature detection. An alternative or parallel approach might be incorporating crowdsourcing of feature detection, as explored by DigVentures' 'Deep Time' project in the UK (<https://digventures.com/product/deep-time-2021/>). These approaches are critical to effectively using these data to create the archaeological baselines which are critical for monitoring environmental impacts, whether these be from climate change or changing land use.

12. Methodology for monument classification selection

In order to identify the relevant and appropriate monument classes to include with the research project a staged analysis process was conducted as illustrated in Figure 29. A first step for the identification of monument classifications to be considered in this project was to review the current classifications used by the Archaeological Survey of Ireland⁶. 'At present, the ASI uses some 480 individual monument classification labels to describe recorded sites and monuments.'. Each monument classification was considered with regard to:

1. The physical attributes of the monuments described;
2. the nature of landscapes in which those monuments are most often encountered;
3. whether the monuments were *in situ*;
4. whether ALS survey was likely to be useful in identifying individual examples of the monument type;
5. whether the monument type had sufficient consistency in morphology, including spatial arrangement and dimensions, to enable a standardised description for use in ALS;
6. how many individual examples of the monument classification are known at the present time.

⁶ Accessible here: [Monument Class \(archaeology.ie\)](https://webgis.archaeology.ie/NationalMonuments/WebServiceQuery/Lookup.aspx)
<https://webgis.archaeology.ie/NationalMonuments/WebServiceQuery/Lookup.aspx>

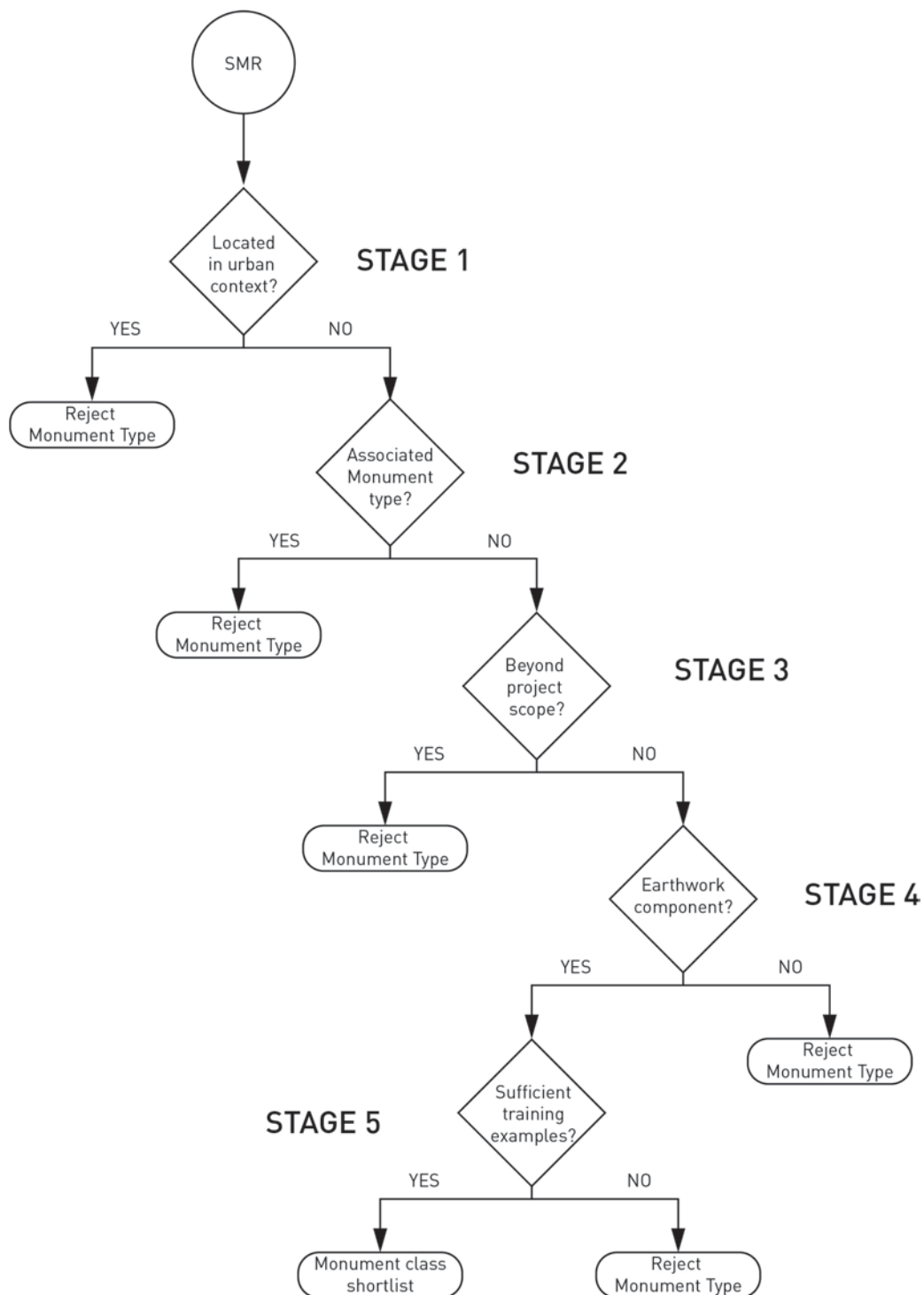


Figure 29: Flowchart illustrating the staged processes involved in the analysis and proposal of the sites and monument classification scheme to use within the project.

12.1. Stage 1 Filter

An initial filter was then applied to discount (a) those monument classifications normally associated with urban contexts today, or which (b) were no longer in their original location. Some 18 classifications are normally associated with urban context, and 50 of the classifications relate to monuments which have been moved, or are likely to have been moved, from their original location. In total 68 classifications were discounted in this filter exercise.

12.2. Stage 2 Filter

A substantial number of monument classifications relate to structures or features that one might normally expect to find in close association one to the other, e.g., 'Cross', 'Church' and which, as noted above, are discrete features of the same place. Many of these features may be useful indicators of larger sites and monuments that have not yet been recorded (for example, an isolated stone cross may simply be a wayside cross without attendant structures but could equally be the last upstanding or well-known component of an extensive ecclesiastical site). In addition, there are a substantial number of classifications relating especially to sculptural and architectural pieces, prehistoric and historic, for which ALS is unsuitable as a survey methodology. Again, future ALS surveys focusing on the wider landscape around such features may provide useful information about associated features and remains, as discrete research projects.

Classifications relating to these kinds of sites were also discounted in a second filter exercise for this current project. This resulted in a further 258 classifications being discounted from further consideration for this project.

12.3. Stage 3 Filter

As part of a further review, all of the monument classifications were considered again vis-à-vis the usefulness of ALS as a primary identification survey methodology. In reviewing the sites, it was evident that there are some classifications that are likely to be clearly identifiable through ALS, some that are not, and a large number that 'might be'. It was clear that, in addition to the sites noted at the Stage 2 Filter, there are significant monument classifications in the 'might be' category for which ALS would be very useful, such as mapping the extent of former designed estate landscapes or understanding the wider landscape of sites of industrial activity, especially mining and milling. However, in the context of this current project, most of those sites have been discounted at this time. Individually these would be interesting projects, and potentially a research area that can be considered in the future. At the present time mapping the extent of such landscape-level cultural components is beyond the scope of this current project. It should be noted, however, that individual features associated with such classifications may be picked up through MLM, since the morphology of individual features are common to more than one kind of monument.

Any classification that had not already been discounted at Stage 1 and Stage 2, and which was considered as unlikely to be appropriate in this current project, was discounted at Stage 3. Only three classifications were discounted at Stage 3.

12.4. Stage 4 Filter - Shortlist

Following the processes of reducing the numbers of sites for consideration at Stages 1-3 above, some 151 monument classifications remained for consideration. At Stage 4 it was intended that a final filter may be applied to the remaining classifications to achieve a much shorter list of sites to be brought forward for consideration in this project. Particular attention was paid to what MLM may achieve in the context of this project, based upon the monument classifications identified either as enclosures, conglomerations (or accretions) but also taking into account sites that were composites, i.e., both enclosures with conglomerations or accretions.

Attempting to reduce this to a shortlist of three proposed classifications for development of a machine learning model presents a number of questions around site selection and bias in terms of what one may expect to achieve from the process. While 151 individual monument classifications remained, these included some 31+ classifications that related to specific buildings or structures that may be expected to have above-ground remains, not including religious houses (which would bring the overall figure of classifications that relate to non-funerary structures to at least 55 individual entries), some 22 different classifications that relate primarily (but not exclusively) to earthwork enclosures, including ringforts, ringworks, henges, ceremonial enclosures and motte and bailey castles, at least eleven classifications relating to landscape organisation and division, including linear earthworks, field boundaries and cultivation, and there also remained various classes of prehistoric tombs and structures.

It is evident, too, that while at least some of the separations in the classifications used relate to slight differences in size and morphology, in many instances the classification of a site to be one monument or another involves a matter of professional judgement, based on inspection of the site and an understanding of its setting and relationship to other features. Included in the list of monuments for consideration at Stage 4 are a considerable number that, in general, are unlikely to be considered as 'earthwork' monuments. These sites have been purposefully included for consideration at this stage; few masonry or stone monuments in Ireland exist divorced from some form of earthwork or earthen component. A stone cairn is frequently covered and appears as a grassy knoll, while the collapsed walls of a masonry building will become engulfed in vegetation and eventually soil deposits.

As a result, Stage 4 was initially split into two sections, 4 (a) and 4 (b). The list of sites at 4 (a) removes all of those monuments which are not normally considered to be earthworks and isolates those sites that (as per the project specification) may be discounted solely based on the material used to compose the monument types. The list at Stage 4 (b) is derived from a different filter, essentially a re-classification of the final list of sites into a number of categories that reflect the dominant kinds of features one may associated with each monument: curvilinear enclosures; rectilinear enclosures; contiguous earthworks; mounds; and composites. In this instance a composite site is identified only where there is expected to be a fairly even balance between two or more features, with an attempt to confine the wider classifications tightly into these five categories. A final category for 4 (b) is that of 'Discount', that is, on further consideration of the classification it should not be included at this time.

The filter at Stage 4 (a) thus resulted in a potential 88 classifications being discounted, including ringfort – cashel, and all of the religious houses. Of the 63 classifications that remained, many of the same challenges identified above persist, with the various barrows dominating the list of potential classifications for further investigation. The Stage 4 (b) filter produced a different kind of result. One classification could be discounted, with the remaining 150 classifications as outlined in Table 2.

Overview Classification	Number of Classification types	Notes
Composite	3	Sites covered by these classifications include a combination of enclosures, upstanding elements and spatial arrangement that are likely to be encountered for each element individually under other classifications (e.g., a motte and bailey includes a mound, ditches, enclosures etc.).
Contiguous Earthworks	12	This group includes Linear Earthworks, features of landscape organisation and use, and earthwork defences encountered at promontory forts.
Curvilinear Enclosures	54	This covers a significant number of monuments, including all classes of Barrows, Ringforts and most Enclosures, all of which share similar aspects of morphology in terms of the character of their enclosing elements.
Mounds	22	This collection includes a significant number of classifications which may require subsequent refinement, since the footprint of the various classification types vary considerably in shape.
Rectilinear Enclosures	59	This group included all of the classifications that included monuments generally composed of structures and complexes laid out in straight lines and approximate right angles. It includes nearly all building forms, as well as military earthworks.

Table 2: Summary of the number of monument classifications selected following stage 4 filtering process.

The project team was concerned, following the filter processes applied thus far, that the actual numbers of individual sites available to establish a machine learning model would be insufficient. The numbers involved may have been low in the first instance, since individual examples of certain monument classifications are few in terms of their total occurrences in Ireland, while the numbers located within the ALS survey extents (as noted above) for other classifications could be low because the lands surveyed do not have many sites in the first instance. This stage in the process echoed Ó Ríordáin's words, seeking this time to identify "a typical class of antiquity widespread throughout Ireland". The results were somewhat surprising when one considers the Sites and Monuments Record as a whole, as there are clearly monument types that are very well known (such as passage tombs) that are widely understood as being rare, but there are also monuments which have a widespread distribution which, while numerous, are not necessarily as common as one might have expected. Further, as it is widely recognised that not all monuments have an even distribution across the whole island of Ireland, there was concern that the specific survey areas might not have sufficient numbers of the most common monuments to provide a statistically valid sample group.

A further filter, Stage 4 (c), was thus applied to the process to establish the total numbers of monuments located within available ALS surveys in order to calculate the potential number of monuments available as machine learning training candidates. A geospatial query identified the monument point data which were contained within the polygons of the ALS survey extents. The relevant ALS surveys utilised within this evaluation include:

1. TII Noise mapping ALS survey
2. TII Road Scheme ALS survey
3. LiDAR Coverage GSI/DCHG/DP Phase I – Various surveys collated from several organisations (Geological Survey Ireland (GSI), Department of Culture, Heritage and the Gaeltacht, Discovery Programme and the Heritage Council)
4. LiDAR Coverage Office of Public Works (OPW)
5. LiDAR Coverage Office of Public Works (OPW) National Aerial Survey Contract (NASC)
6. LiDAR Coverage Office of Public Works (OPW) Cork
7. LiDAR Coverage Westmeath County Council

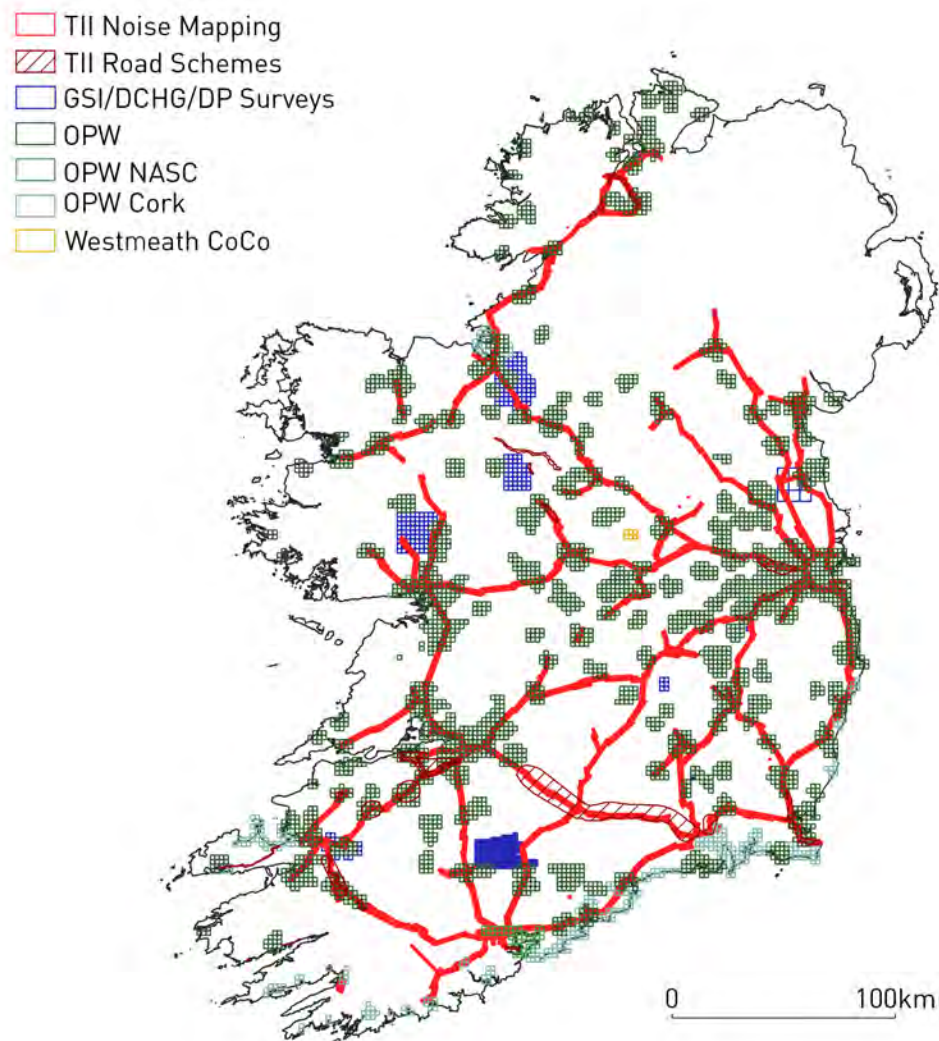


Figure 30: Map illustrating the survey extents of the relevant ALS datasets which could be potentially used within the project.

The accuracy of the polygons which represent the ALS survey extents differs slightly depending upon where they have been sourced. For the ALS data provided by the TII (noise mapping data), the extent polygons were calculated from the original .las file extent and are therefore a true reflection of the ground coverage. For additional ALS data which may be available for use within the project the Geological Survey of Ireland provided a polygon dataset which represented the indexes of the raster tiles available through the Open Topographic Viewer. Where an index tile is fully populated with ALS tiles they will act as a true representation of the ALS survey extent, however, where indexes are not fully populated there will be an overestimation of the number of monuments contained (see Figure 31). The effect of this should not be too significant in the classification type selection process, but should be noted at this stage. Once the project has received all the ALS data to generate the training datasets this will be rectified.

Although available through the Open Topographic Viewer, the Photogrammetry Coverage GSI Ireland datasets were not considered for inclusion due to the different nature in terms of resolution and capture methodology and the focus of the survey area being on coastal and intertidal areas.

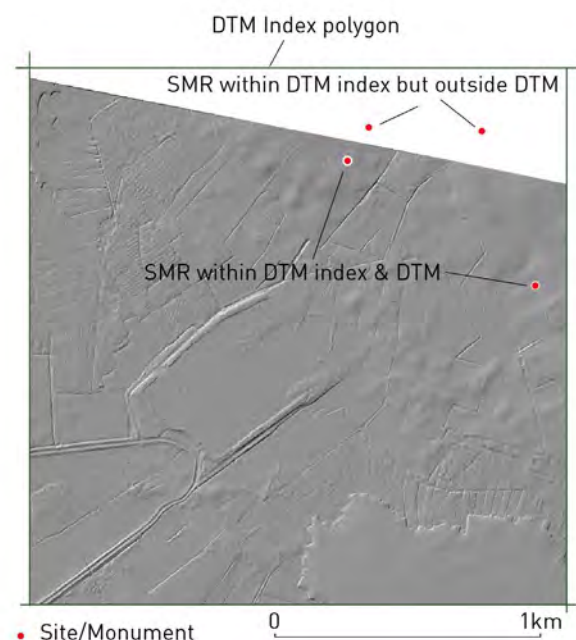


Figure 31: Illustration of the overestimation of the number of monuments contained within the ALS indexes provided by the GSI.

This filter considered the total number of monuments contained within the five classifications at Stage 4 (b). Examination of the numbers of individual monuments involved indicated that some 62 individual classifications could be discounted at this stage, leaving some 88 classifications for further consideration in monument selection as detailed in Table 3. The full processes were also repeated against the original SMR classification scheme to identify the total number of monuments available for training as detailed in Table 4.

Overview Classification	Number of previously recorded individual monuments in the classification types within the ALS survey sets	Number of Classification types as per Stage 4 (b)	Notes
Composite	239	3	The total number of sites thus identified is very low, and unlikely to provide a sufficient sample to be used for a Machine Learning Model at this time. Individual component elements may be picked up under other classification categories.
Contiguous Earthworks	1,015	12	The number of individual sites here is low, and likely to include a number of monument classifications that have significant differences in terms of scale and arrangement, though sharing similar morphological elements.
Curvilinear Enclosures	19,208	54	This is a significant number of monuments, relating to 54 individual monument classifications. The average value per classification is approximately 350 individual sites, which is much less than one would require for a Machine Learning Model, but within this group there are individual monument classifications that account for 100 or fewer individual monuments. Within this classification Ringforts, which are the most common classification overall and enclosures are both numerous enough to be included within this research.
Mounds	3,605	22	The total number of monuments would be sufficient for the purposes of developing a Machine Learning Model, though the average value per classification is low, at c.160 individual monuments. However, some of the individual classifications account for 50 or fewer individual monuments. In practice most mounds that are archaeological are already recorded. The model would find them but I this may be of limited value.
Rectilinear Enclosures	478	59	The total number of sites thus identified is very low, and unlikely to provide a sufficient sample to be used for a Machine Learning Model at this time. Overall, within the Class List Definitions, there are comparatively few individual monuments of this form, of which Moated sites are the most common but account for less than 1% of the total number of recorded monuments.
Total	24,545	150	The total number of monuments represents approximately 14% of the total number of recorded sites on the island of Ireland, and about 17% of the total recorded sites within the jurisdiction of the State of Ireland.

Table 3: Summary the number of observed monuments within each overview classification present within the available ALS footprints.

SMR Classification	Overview Classification	Total instances
Barrow - bowl-barrow	Curvilinear enclosures	59
Barrow - ditch barrow	Curvilinear enclosures	126
Barrow - embanked barrow	Curvilinear enclosures	22
Barrow - mound barrow	Curvilinear enclosures	153
Barrow - pond barrow	Curvilinear enclosures	9
Barrow - ring-barrow	Curvilinear enclosures	424
Barrow - stepped barrow	Curvilinear enclosures	17
Barrow - unclassified	Curvilinear enclosures	246
Bastioned fort	Rectilinear enclosures	45
Boundary mound	Mounds	25
Burial ground	Curvilinear enclosures	417
Burial mound	Mounds	29
Burnt mound	Mounds	582
Castle - motte	Mounds	108
Castle - motte and bailey	Composite	68
Castle - ringwork	Curvilinear enclosures	50
Ceremonial enclosure	Curvilinear enclosures	20
Children's burial ground	Curvilinear enclosures	411
Churchyard	Curvilinear enclosures	7
Concentric enclosure	Curvilinear enclosures	36
Crannog	Mounds	257
Cultivation ridges	Contiguous Earthworks	11
Cursus	Contiguous Earthworks	5
Earthwork	Curvilinear enclosures	591
Ecclesiastical enclosure	Curvilinear enclosures	309
Ecclesiastical site	Curvilinear enclosures	55
Embanked enclosure	Curvilinear enclosures	15
Enclosure	Curvilinear enclosures	5,563
Enclosure - large enclosure	Curvilinear enclosures	151
Field boundary	Contiguous Earthworks	109
Field system	Contiguous Earthworks	483
Fortification	Rectilinear enclosures	27
Fulacht fia	Mounds	1,986
Graveyard	Curvilinear enclosures	1,367
Habitation site	Curvilinear enclosures	175
Henge	Curvilinear enclosures	10
Hillfort	Curvilinear enclosures	28
Hilltop enclosure	Curvilinear enclosures	46
Inauguration site	Curvilinear enclosures	6
Linear earthwork	Contiguous Earthworks	65
Linkardstown burial	Mounds	6
Moated site	Rectilinear enclosures	394
Mound	Mounds	393
Platform	Rectilinear enclosures	9
Platform - peatland	Mounds	34
Promontory fort - coastal	Contiguous Earthworks	98
Promontory fort - inland	Contiguous Earthworks	26
Ringfort - rath	Curvilinear enclosures	8,183
Ringfort - unclassified	Curvilinear enclosures	711
Road - hollow-way	Contiguous Earthworks	13
Road - road/trackway	Contiguous Earthworks	204
Settlement cluster	Composite	69
Settlement deserted - medieval	Composite	102
Settlement platform	Mounds	185

Table 4: Summary the number of observed monuments within each SMR classification present within the available ALS footprint (note: SMR classifications where the total instances are less than 5 have been omitted from the table).

12.5. Stage 5 Filter – Final Selection

Following on from the analysis and filtering at stages 4a and 4b, the final selection of monument classifications for inclusion within the project is dependent upon there being a critical number of monument examples within the ALS data to develop a successful MLM. Generally, the larger the training dataset the greater statistical power for pattern recognition (Raudys and Jain 1991) (Varoquaux 2018). Based upon the calculation of the number of monument classification types as detailed in Table 4 there are only a few classes which would have a sufficient number of examples within the ALS data to ensure the creation of a representative and successful training model, as detailed in Table 5. Within several of the classes (barrows, enclosures and ringforts) amalgamation has taken place to provide a larger critical mass of training examples.

SMR Classification	Overview Classification	Total instances
Barrow (all classes excluding pond)	Curvilinear enclosures	1,047
Burial ground	Curvilinear enclosures	417
Earthwork	Curvilinear enclosures	591
Enclosure (including large enclosure)	Curvilinear enclosures	5,714
Moated site	Rectilinear enclosures	394
Ringfort (rath & unclassified)	Curvilinear enclosures	8,894

Table 5: list of monument classification types and number of potential training dataset instances

Of the monument classifications outlined above there are several clear candidates based upon the significantly larger number of examples (enclosure and ringfort). For the remaining classifications barrows would have a significant number of examples and could potentially be augmented through the inclusion of barrow examples within the UK (see 13. Supplementary ALS data)

In addition, the quality of the monument examples and annotations is essential for supervised training, validation, and testing of machine learning models. Within the potential monuments identified it can be assumed therefore that not all will be able to be used, therefore the numbers calculated within Table 4 will be reduced. Several factors will call for the inspection and potential reduction of the training set, including:

1. Sites or monuments which are located towards the edge of an ALS survey are maybe truncated or may not contain sufficient background terrain which enable the features (banks, slopes depressions etc.) to be identified
2. Sites of monuments which have no physical presence within the surface topography (e.g. site only evident in crop mark). In some cases, this may be noted within the description field within the SMR database

Recommendation: It is proposed that quality control of the dataset will be carried out before the training samples are created

13. Supplementary ALS data

In order to supplement the number of barrow models within the training datasets additional examples could be included from other geographic regions where applicable monuments are evident which display similar morphology and comparable ALS data exists. As there are many similarities between several of the monument types of Ireland and the UK, additional training examples could be sourced. As there is little ALS data capture within Northern Ireland and

Wales, effort in this task was focused on the potential monument data which exists in England and Scotland.

13.1. Evaluation of UK sites and monument data

The first stage of this procedure was to carry out the process of monument classification harmonisation between Ireland, Scotland and England as each jurisdiction would have adopted their own standardised system. Access to the Scottish and English monument vocabularies was obtained through the Heritage Data – Linked Data Vocabularies for Cultural Heritage website (<https://www.heritagedata.org/>). As both the Scottish (Monument Type Thesaurus (Scotland)) and English (FISH Thesaurus of Monument Types) vocabularies utilise “broader” and “narrower” semantic relationships it was adopted that the broader terms should be used to qualify any relationship. Equivalent terms for monument classifications are summarised in Table 6. If it is required that additional training examples are used from both Scotland or England a more detailed evaluation of each narrower term may be required and certain classifications excluded.

Ireland	Scotland	England
Barrow - bowl-barrow Barrow - ditch barrow Barrow - embanked barrow Barrow - mound barrow Barrow - pond barrow Barrow - ring-barrow Barrow - stepped barrow Barrow - unclassified	Barrow - Long barrow - Round barrow - Bank barrow - Square barrow - Chambered barrow	Barrow - Long Barrow - Round Barrow - Bank Barrow - Square Barrow - Chambered Barrow - Ring Barrow - D Shaped Barrow
Enclosure Embanked enclosure Enclosure - large enclosure	Enclosure - Pit enclosure - Rectilinear enclosure - Causewayed enclosure - Curvilinear enclosure - Circular enclosure - Palisaded enclosure - Turf enclosure - Burial enclosure Henge	Enclosure - Pit Defined Enclosure - Rectilinear Enclosure - Causewayed Enclosure - Curvilinear Enclosure - Rectangular Enclosure - Defended Enclosure - Stockaded Enclosure - Ditched Enclosure - Ring Enclosure - Banjo Enclosure - Subrectangular Enclosure - Tor Enclosure - Henge Enclosure
Earthwork	Earthwork - Boundary earthwork	Earthwork - Bank (Earthwork) - Embankment Cross - Linear Earthwork

Table 6: Potential equivalency between the monument classification vocabularies of Ireland, Scotland and England for selected monument types.

Based upon the classifications identified in Table 6 a formal request for an extraction of Historic Environment Records (HERs) data was submitted to Historic Environment Scotland through the CANMORE (<https://canmore.org.uk/>) system. Both polygon and centroid data for all the relevant sites contained within the Scottish HER were provided to the project under license.

For the English data, as HERs are formally managed by each local authority, access was more challenging. For many local authority areas, HER data is made available through the Heritage Gateway (<https://www.heritagegateway.org.uk/>) or through their own host authority's website. None of these services offer the ability for bulk download of HER data at present. A formal request was submitted to Wiltshire and Swindon HER to test the availability of data, as this local authority's area would contain a significant number of equivalent monuments, particularly barrows. The request was rejected as this would have incurred additional cost to the local authorities as the management of the data is carried out by an external organisation.

Additional data does exist in the form of the English Monuments GIS Data where scheduled monuments are represented by a polygon defining the extent of the statutorily protected area. Although this would not be an exhaustive list of potential sites, there is still the potential to utilise this as the number of sites contained is still significant (e.g., 5,457 barrow polygon definitions exist within the data). Before this dataset is used, an evaluation of the sites selected must be carried out as there may be bias introduced into the MLM as the preservation and maintenance of these sites may be relatively better in comparison to the full range of English sites.

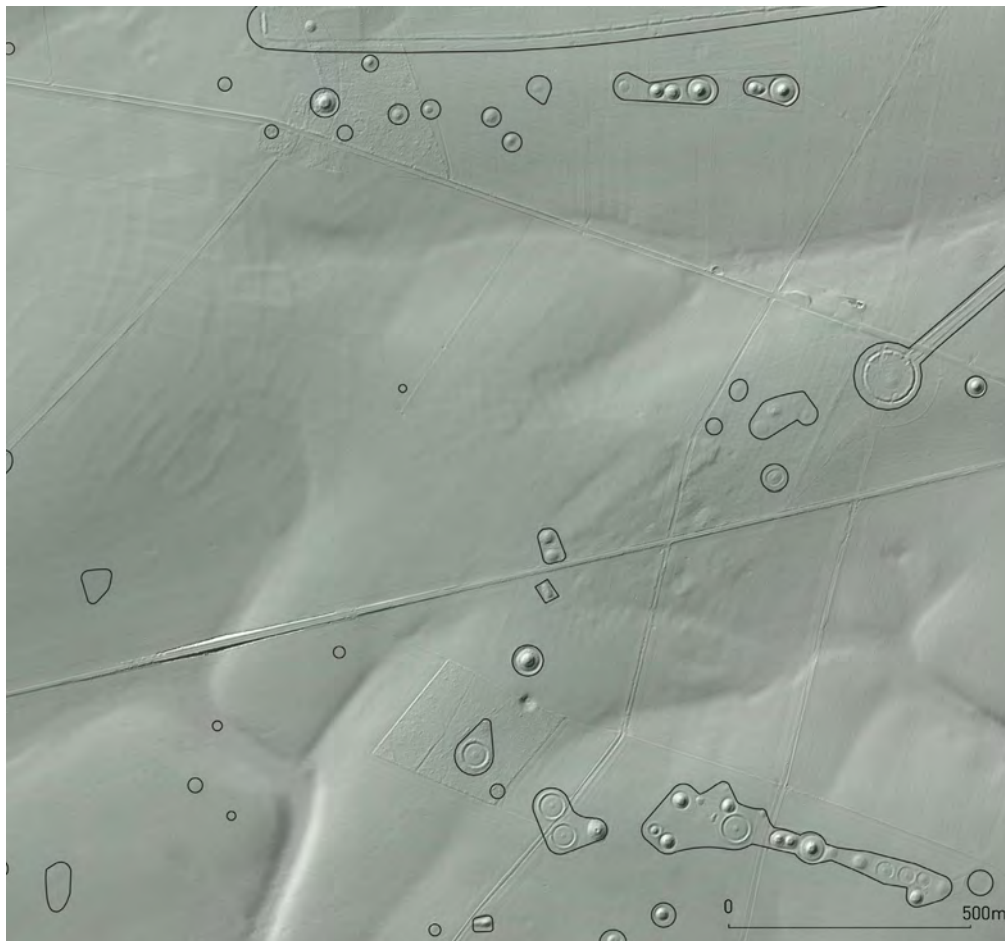


Figure 32: Image displaying part of the ALS derived relief shaded model for GB grid tile SU14SW located in County Wiltshire. Each of the enclosing polygons represents the scheduled area which contains at least one example of a barrow site. For this tile there are 89 scheduled barrow polygons

Within the English Scheduled monuments polygon data each polygon has the potential to enclose several monuments (see Figure 32). This indicated that each tile has the potential to contain many more training examples than initially estimated, however, care must be taken to

ensure individual monuments are selected for training purposes rather than monument complexes.

13.2. Evaluation of UK ALS data

Within the UK acquisition, management and publishing of ALS data is devolved to each Nation. Of the four nations, England and Scotland have carried out significant ALS campaigns in the past and make the data available through the relevant portals.

Within England, The Environment Agency National LIDAR Programme provides accurate elevation data at 1m spatial resolution for all of England up to the end of 2021 with survey data available in a range of formats (DEM, DTM, *.laz point cloud). Data is available at several resolutions including 1m surface model where the point cloud will have an average point density across the survey of at least 1 point per square metre (ppsqm) and 50cm resolution data, which increases the point density to 4 ppsqm and 25cm resolution is 16ppsqm. Access to the data is provided through the Department for Environment Food & Rural Affairs Data Services Platform (<https://environment.data.gov.uk/>).

Within Scotland the Scottish Remote Sensing Portal (<https://remotesensingdata.gov.scot/>) provides access to ALS data which has been collected across the country since 2009 on a phased basis, with the latest phase (Phase V) occurring between 2020 to 2021. Data is available across a range of formats (DEM, DTM, *.laz point cloud) with resolution varying between 1m and 0.25cm for various areas of the country, however full national coverage does not exist to date. Both countries datasets are delivered as 5km tiles aligned to the ordnance survey grid with the tiles named by the unique survey id, OS grid reference and the first and last survey date of the survey id (PXXXXXOSOSOSSDFLOWNEDFLOWN.*).

This project does not have the time or resources to evaluate all the lidar tiles available therefore only those tiles with a high density of monument examples will be used if required. Geospatial analysis was conducted to establish the monument density within each 5km OS grid tile to identify those areas which would produce the most benefit from further investigation. The results of this analysis are illustrated in Figure 33 and Figure 34. Those tiles which contain a significant number of examples (e.g., more than 20 sites) will be selected for evaluation.

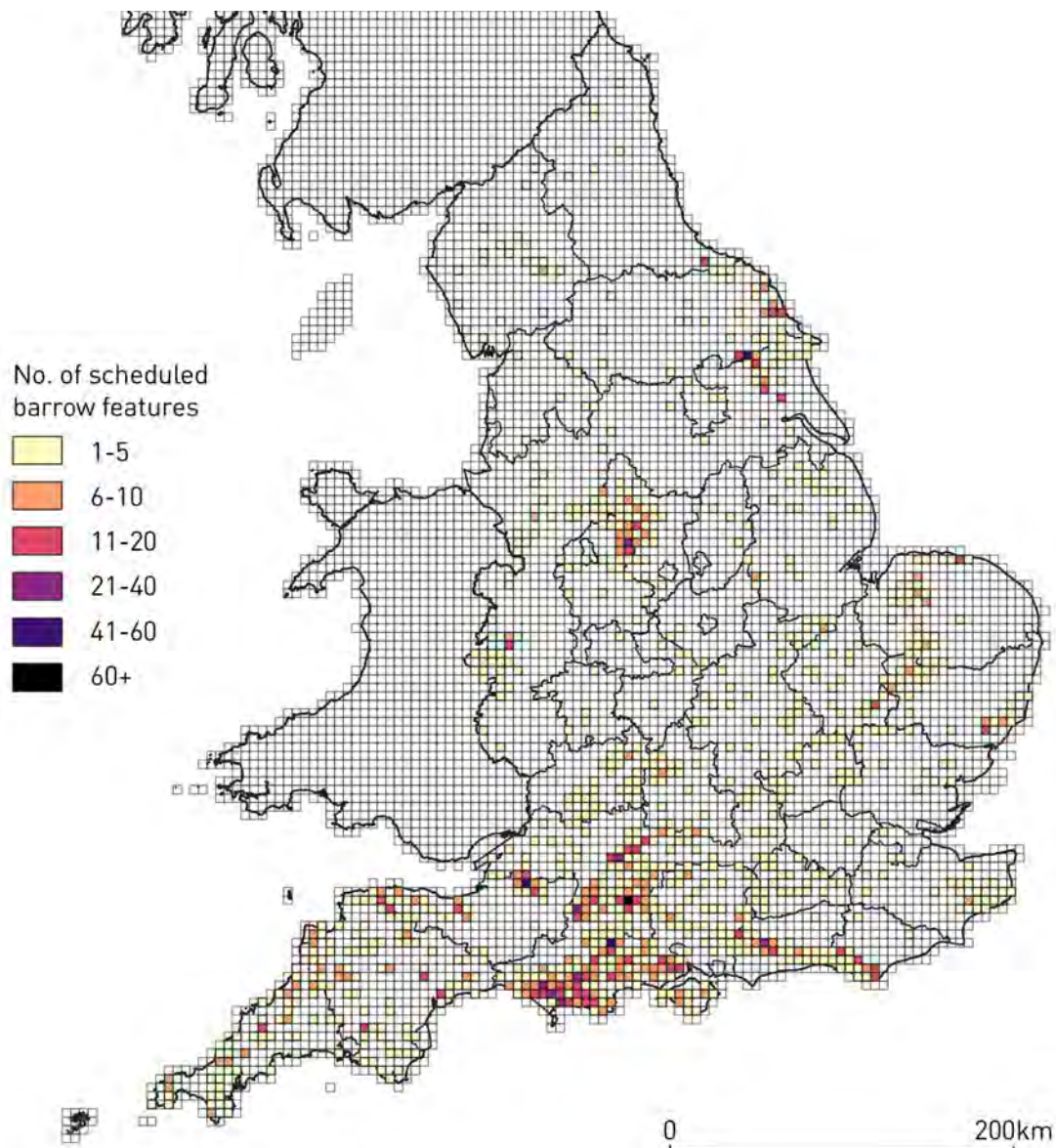


Figure 33: Map illustration of the density of scheduled barrows within The Environment Agency's National LIDAR Programme. Each grid cell corresponds to individual ALS tiles (5km tiles aligned to the ordnance survey grid).

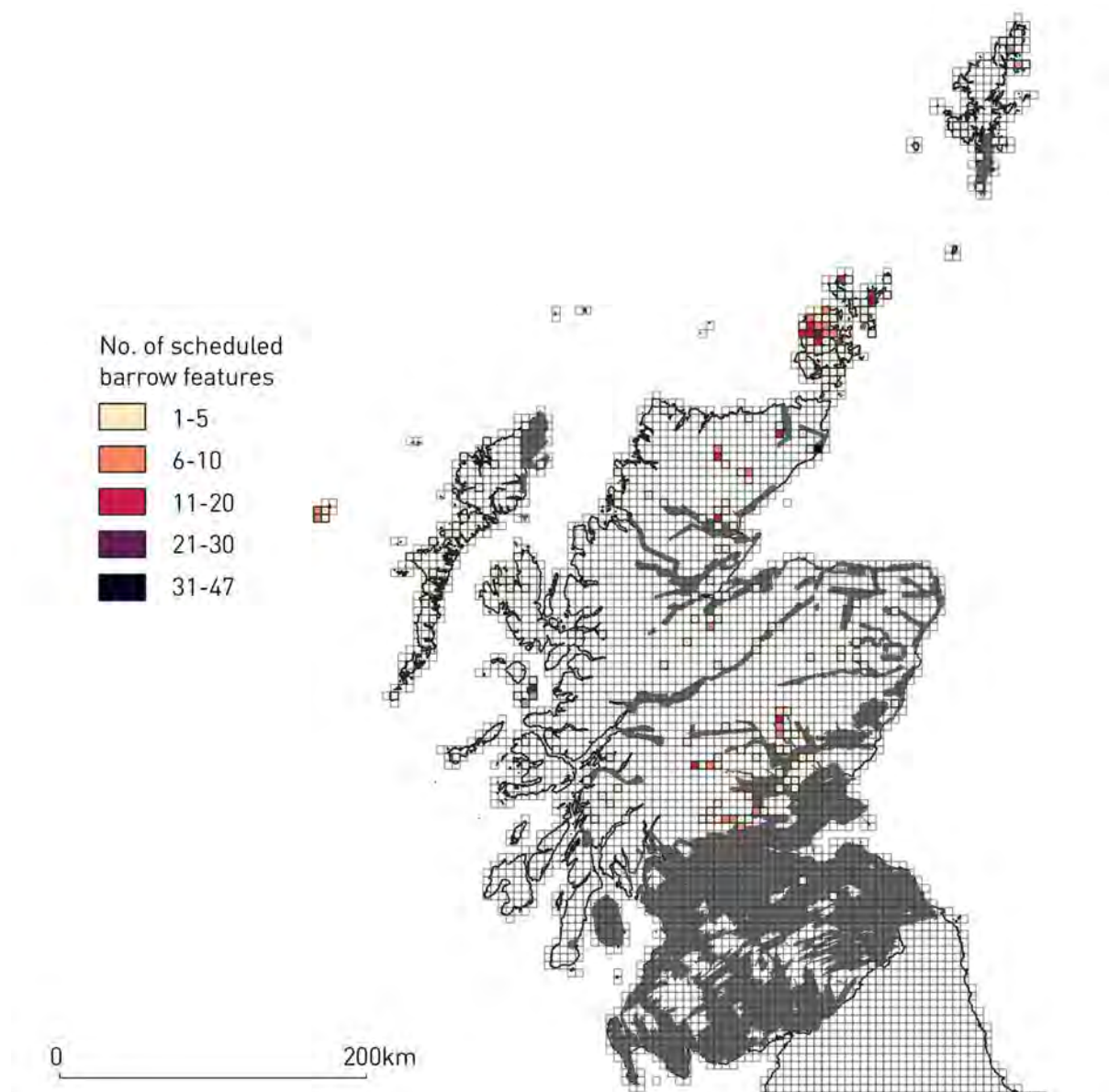


Figure 34: Map illustration of the density of scheduled barrows within Scotland. The grey area corresponds to the ALS survey extent captured through Phase I to Phase V of the Scottish Lidar acquisition 2009 – 2021. Each grid cell corresponds to individual ALS tiles (5km tiles aligned to the ordnance survey grid).

Recommendation: If required, the training datasets created for barrows will be augmented through the inclusion of selected sites within England and Scotland and their corresponding ALS datasets following quality control measures. Full specifications and details of all the LAS datasets utilised within the project will be provided within the Final Technical Report

14. General observations

Having conducted an analysis on the current profile of Irish monuments, using the classification used in the formation of official records, there are issues about the current classifications that may inhibit realisation of the effectiveness of a machine learning model if it is tightly bound to the current monument classifications in use. These issues do not mean that the concept cannot be taken forward, but it does mean that those issues need to be taken into account.

One of the first issues that arises is the total number of individual monument classifications provided in the Class List Definitions. One may anticipate that a machine learning model could potentially be developed for each monument classification, but this is unlikely ever to be a practical approach unless processing capacity of most computers is upgraded to allow multiple re-processing and searching of data for each of the 480 types. Furthermore, there are many classifications that account for a very small number of monuments overall. Perhaps the greatest challenge is the variability in monument scale, composition and survival, especially within the largest single monument classification (Ringforts-raths). While there is a drive to consistency in the record, it is the case that errors occur, while the quality of fieldwork can, at times, be variable. One may find, for example, that there are slightly different approaches to recording individual sites, estimating or measuring diameters based on slightly different parameters (e.g. for the recording of a ringfort, can one be certain that the dimensions recorded are across the same features in each instance, top of bank to top of bank, toe to toe, ditch centre to ditch centre, or outside edge to outside edge?).

A question under consideration by the Project Team is whether or not individual feature types, rather than whole monument classifications, would be a more valuable approach in the development of a machine learning model. While there may be similar variation in terms of scale, it may be possible to set clear parameters around what the machine learning model is designed to recognise, though this does have an inherent risk that the model would identify monuments as one might wish to find them but misses sites as they really are. The intention at Stage 4 (b) is intended to avoid the tyranny of classification in terms of monument or site types, but to retain a focus on the component parts of individual sites that are ultimately used to describe and classify each monument.

If one takes the most common class of field monument, the Ringfort-rath, it is immediately apparent that, at the present time, the considerable number of monuments included in the Sites and Monuments Record could be reduced into multiple sub-categories. For the purposes of study such classification would be useful, but likely to be problematic. O'Sullivan *et al* (2014, 55-57) outline the challenges in achieving consensus amongst archaeologists about new classifications relating to enclosure forms and functions associated with the early medieval period, challenges that are not unique to that period. Making sense of those features, in terms of their form, function and date, requires flexibility and thought. However, in the Northern Ireland Sites and Monuments Record this classification is broken into several components, based upon certain characteristics of the monument (including Platform Rath, Counterscarp Rath, Raised Rath, Bivallate, Trivallate, Multivallate), with at least nine distinct classifications, not including either cashels or 'enclosures' that do not have a specific or likely identification as a rath. These monument types are found all over Ireland, so the classifications used in one jurisdiction could equally be used in the other.

Within the individual classifications used in Northern Ireland for Ringforts-Raths, it is clear that the classifications associate sites that are of similar characteristics in plan or their present composition. They share a conceptual description, where measurement is a factor but not necessarily a defining feature. While most raths are found to have a single penannular ditch, set around a penannular bank (with the break in the ditch and bank aligning and leading one to the other), enclosing a circular or sub-circular space, there continues to be significant variation in scale of each of those elements. Arguably, variations in scale for each element – if the variation is a reflection of the original construction rather than its subsequent history and survival – may be indicative of different functions and potentially different dates. A shallow,

wide-bottomed ditch, narrow bank and large enclosed circular space may, for example, lead one to suspect the enclosure is not a rath but may be an ecclesiastical enclosure, even if no church ruin or other folklore that the site was used for ecclesiastical purposes. In terms of upstanding remains, archaeological fieldwork aimed at identifying previously unrecorded church sites involved trying to identify any traces of enclosing elements such as banks and ditches. Swan (1983, 269, 274), has previously noted that curvilinear enclosures at church sites is a recurring phenomenon that appears to signify the remains of landscape organisation in the first millennium AD associated with the church foundations themselves. In essence, such enclosures may be taken as evidence of an early (pre-Norman, most likely first millennium AD) ecclesiastical foundations. Thus, while the difference between one site and another will be immediate and obvious, it should be clear that in some instances the differences and attributes will be subtle and may well have been missed in previous surveys.

Just as most of our classifications of monuments derive from a conceptual differentiation of one site from another, most of the monuments that dot our landscape were built according to a concept, not a series of universal, pre-ordained and measured plans. Stout (2000, 14) observed that most ringforts are close to being a 'true' circle, suggesting "the plan of each new ringfort was originally laid out using a measuring line pivoted from a central stake", further noting that this could only occur if the site was clear of any other obstacle, including earlier buildings, the inference being that ringforts were normally constructed on greenfield sites (though it may also involve the remodelling of an earlier site where buildings were levelled). In this respect there is clearly a shared technological approach to setting out the perimeter and internal arrangement of the ditches and banks that enclose a central space. But while the concept here may be shared throughout the island, the size of the individual sites is not, with internal diameters varying from 15.5m to 75m (*ibid*, 15). Ironically, while there is a large number of separate categories or site types for raths used in Northern Ireland, these sites would all most likely be included in the category 'Rath', just as they would all be included in the category 'Ringfort-rath' by the Archaeological Survey of Ireland. There has not yet been a sub-division based on scale, morphology of enclosing elements or the relative configuration and sizes of those enclosing elements.

It is also clear that, for certain other circular earthwork monuments a similar foundation approach was undertaken as that described by Stout for ringforts. Ceremonial enclosures, large hilltop enclosures and many other sites with enclosing earthworks may be noted as having a consistent radius from a common point of origin, but local topography (such as steep slopes or undulations in the ground) may account for sub-circular or irregularly-outlined monuments which may otherwise be considered to be curvilinear, annular or arcuate. In those instances where a consistent radius can be identified one may suspect that, at the time the enclosing element was set out, there were either no obstacles to tracing such a radius (that is, a person or stake at the point of origin held one end of a rope or chain, while at the other end was used to trace a circumference of the site), or there was very special regard for maintaining the line of the circumference with minimal interruption.

Where sites have multiple forms of enclosure, especially concentric enclosures, the form of the enclosing elements may similarly reflect something of the site's history, especially if some of those elements are accretions to or reductions of an existing site. Harney (2020b, 34) notes that additional enclosures were added to some ecclesiastical monuments over time. Where multiple enclosures occur, especially where concentric enclosures occur, it would be instructive to establish if these are happening on sites that are already in use, accretions to the original

enclosures, or new arrangements on a fresh or levelled site. Harney (2020a, 23) citing (Henry 1965, 81), and (Doherty 1982, 302) further noted that such multiple enclosures, either double or triple enclosure of a church site, was permissible (though not obligatory) by ecclesiastical authorities by at least the eighth century (see also Lawlor (1895)). In this case there is a *schema* for site development that is illustrated in the eighth century AD, but this remains a conceptual plan not a measured blueprint.

From our analysis of the sites and ALS datasets, including datasets outside of TII's commissions, some, 88 monument classifications have been identified for further consideration, representing some 26,466 individual monuments. While Ringfort-rath monuments dominate the selection, these are arguably the kinds of sites that are most readily identifiable through field inspection, previous mapped histories and aerial photography, whereas the complexity of individual examples of such sites is perhaps more significant in terms of dealing with new route selections or mitigation measures for new roads.

15. Conclusions and Recommendations

As outlined at the start of this paper, the purpose of the present project is to carry out research leading to the development of a user-friendly software tool capable of automatically detecting archaeological features from ALS data using machine learning, specifically deep learning algorithms. It is evident that there are features that are readily detected using ALS data, and it is likely that ALS can be 'trained' to detect such features.

The next stage of the process, to develop a 'trained' machine learning model that can automatically detect three classes of earthwork monuments from ALS data is more challenging, especially in terms of developing a learning model that is useful vis-à-vis landscape management in the context of spatial development and landscape-level change. While it is important to ensure that the concept of using ALS to automatically detect historical monuments is tested, it appears a priori that this may immediately run into difficulty in terms of:

- the datasets available, in terms of quality of survey data and resolution;
- the datasets available, in terms of the numbers of discrete monuments already recorded within the survey footprints;
- the highly variable nature of monument classification and differentiation used in the present Class List Definitions in Ireland;
- the empirical data gathered in respect of individual monument examples, within individual monument classifications; and
- a reliance upon identifying the kinds of monuments that are already known (either by archaeologists or non-archaeologists) or the identification of examples of monuments within existing classifications that are not currently known, without consideration of identifying either new or previously-unrecognised monument types, monument variants or monument examples.

Furthermore, having reviewed the system of classification in use, it is important to emphasise that the process of feature identification is not the same as monument classification. The presence of recognised features may lead one to identifying that place as having archaeological potential, with a suspected monument evidenced by that feature or features

revealed through ALS data, but this does not in itself determine that the feature is, in fact, of archaeological interest nor can it attribute an immediate classification. That process of classification is likely to require data inputs out with those currently logged as key site data entries within the Sites and Monuments Record. Within the datasets available it is clear that not all monuments have precise written dimensional description, though at least some of that information may be obtainable in hard copy records and published material, or through interrogation of maps or illustrations of the sites.

Recommendation: The tender specification notes that the classes of monuments to be included in the model will be agreed between TII and the Project Team. Based on the above review and evaluation of the potential application of automated detection processes on Irish monument forms and shortlisting of suitable earthwork monument types for the training process, the following monument classifications are recommended

1. Ringforts
2. Enclosures
3. Barrows

16. Section II References

- Archaeological Survey (Northern Ireland). 1980. *Recognising Irish antiquities*. Belfast.
- Aston, M. 1985. *Interpreting the Landscape: Landscape Archaeology and Local History*. 1st. Routledge. doi:10.4324/9780203442128.
- Aston, M., and T. Rowley. 1974. *Landscape archeology : an introduction to fieldwork techniques on post-roman landscapes*. Newton Abbot: David & Charles.
- Barnes, I. 2003. "Aerial remote-sensing techniques used in the management of archaeological monuments on the British Army's Salisbury Plain Training Area, Wiltshire, UK." *Archaeological Prospection* 10: 83–91.
- Barrett, G. F., B. J. Graham, and C. J. Lynn. 1975. "Some Considerations concerning the Dating and Distribution of Ring-Forts in Ireland." *Ulster Journal of Archaeology*, Third Series ed.: 33-47. <https://www.jstor.org/stable/20>.
- Barry, T. B. 1987. *The Archaeology of Medieval Ireland*. London: Routledge.
- Bewley, R.H., S. P. Crutchley, and C. Shell. 2005. "New light on an ancient landscape: lidar survey in the Stonehenge World Heritage Site." *Antiquity* 79: 636–647.
- Breen, C. 2005. *The Gaelic Lordship of O'Sullivan Beare: A Landscape Cultural History*. Dublin: Four Courts Press.
- Corcoran, M. 2018. "Land use in medieval and early modern Ireland: An integrated archaeological, geospatial and paleoenvironmental approach. Unpublished PhD Thesis, UCD School of Archaeology."
- Corns, A., and R Shaw. 2010. "Cultural heritage Spatial Data Infrastructures (SDI)-unlocking the potential of our cultural landscape data." *30th EARSeL Symposium*. 1-8.
- Corns, A., and R. Shaw. 2008. "High resolution LiDAR for the recording of archaeological monuments & landscapes." Edited by R. Lasaponara and N Masini. *Advances in Remote Sensing for Archaeology and Cultural Heritage Management, Aracne*. Roma. 99–102.
- Corns, A., and R. Shaw. 2013. "Lidar and World Heritage Sites in Ireland: why was such a rich data source gathered, how is it being utilised, and what lessons have been learned?" In *Interpreting Archaeological Topography*, edited by R. S. Opitz and D. Cowley, 146-160. Oxford: Oxbow Books.
- Corns, A., J. Fenwick, and R., Shaw. 2008. "More than Meets the Eye." *Archaeology Ireland* (Wordwell Publishing) 22 (3): 34–38.
- Curran, S. 2020. "Hidden depths and empty spaces? A remote sensing approach to the exploration of settlement patterns, identity and social hierarchy in early medieval Ireland (AD 400 - 1100). Unpublished PhD Thesis, UCD School of Archaeology."
- Curran, S. 2013. "LiDAR and Early Medieval Settlement in Counties Roscommon and Leitrim: a New Angle on Early Medieval Landscape Studies." Co. *Roscommon Historical and Archaeological Society Journal, 30th Anniversary Edition*.
- DAERA. 2018. "Sustainable Agricultural Land Management Strategy report and Executive summary." ni.gov.uk/sites/default/files/publications/daera/16.17.079%20Sustainable%20Land%20Management%20Strategy%20final%20amended.PDF.
- Daly, C. 2019. "Built & Archaeological Heritage Climate Change Sectoral Adaptation Plan. Prepared by the Department of Culture, Heritage and the Gaeltacht." <https://assets.gov.ie/75639/a0ad0e1d-339c-4e11-bc48-07b4f082b58f.pdf>.
- Darvill, T. 1993. "Can Nothing Compare 2 U?" Resources and philosophies in archaeological resourcemanagement and beyond." Edited by Swain. *Rescuing the Historic Environment, Archaeology, The Green Movement and Conservation Strategies for the British Landscape, Proceedings of a Conference at LeicesterUniversity, January 1993, RESCUE*. Hertford: The British Archaeological Trust. 5-8.
- Davis, S.R. 2013. "Heritage Guide No. 63: The Hill of Ward: A Samhain site in County Meath." *Archaeology Ireland*.
- Davis, S. R. 2011. "The Unseen Hill Of Ward: new insights from LiDAR data." *Archaeology Ireland*, 4 ed.: 36–40.
- Davis, S.R. 2016. "Foynes-Limerick Road Improvement Scheme LiDAR assessment. Unpublished report on behalf of Transport Infrastructure Ireland/Limerick Co. Council."
- Davis, S.R. 2020. "Hill of Uisneach: LiDAR assessment report. Unpublished report on behalf of Westmeath Co. Council."
- Davis, S.R. 2021a. "N11-N25 Oilgate-Rosslare Harbour Archaeological Lidar assessment report. Unpublished report for Aegis Archaeology."
- Davis, S.R. 2015. "N5 Strokestown-Ballaghaderreen corridor: LiDAR Assessment. UCD School of Archaeology (unpublished report for Roscommon County Council and TII."
- . 2012. "Getting to the point(s)." *Seanda*, 6-9.

- Davis, S.R. 2018. "Slane Bypass: Lidar assessment report. Unpublished report for Courtney-Deery Heritage Consultancy Ltd."
- Davis, S.R. 2021b. "The Curragh Plains, Co. Kildare Archaeological Lidar Assessment Preliminary Report. Unpublished report on behalf of Kildare Co. Council."
- Davis, S.R., C. Brady, K. Barton, and W. Megarry. 2013. "Lidar survey in the Brú na Bóinne World Heritage Site." In *Interpreting Archaeological Topography: Airborne Laser Scanning, 3D Data and Ground Observation*, edited by R. Opitz and D. Cowley. Oxford: Oxbow Books.
- Dempsey, G. 2016. "Rathcroghan Field System Survey. Unpublished report for the Royal Irish Academy 'Directed Archaeological Research for World Heritage Sites' scheme."
- Doherty, C. 1982. "Some aspects of hagiography as a source for Irish economic history." *Peritia* 1: 300-328.
- DoHLGH. 2020. "Sceilg Mhichíl World Heritage Site Management Plan 2020–30." https://www.worldheritageireland.ie/fileadmin/user_upload/documents/2022/Sceilg_Mhich%C3%ADL_Management_Plan_2020-2030.pdf.
- Due Trier, Ø, D. Cowley, and A.U. Waldeland. 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." *Archaeological Prospection* 26: 165-175.
- Gardiner, M, and K. O'Connor. 2017. "The later medieval countryside lying beneath." Edited by Stanley et al, 133-152.
- Gardiner, M. 2018. "Landscape and farming in the north of Ireland in the late Middle Ages and early modern period: the evidence from the uplands." *Journal of Irish Archaeology* 27: 117-134. <https://www.jstor.org/stable/10.2307/26564147>.
- Green, S., S. Martin, S. Gharechelou, F. Cawkwell, and K. Black. 2021. "BRIAR: Biomass Retrieval in Ireland Using Active Remote Sensing (2014-CCRP-MS.17)." https://stor.teagasc.ie/bitstream/handle/11019/2654/Research_Report_305.pdf?s.
- Green, S., S. Martin, S. Gharechelou, F. Cawkwell, and K. Black. 2019. "BRIAR: Biomass Retrieval in Ireland Using Active Remote Sensing (2014-CCRP-MS.17)." EPA. <http://hdl.handle.net/11019/2654>.
- Gwynn, A, and R. N. Hadcock. 1970. *Medieval religious Houses of Ireland*. London: Longman.
- Harney, L. 2020. *The early medieval ecclesiastical enclosures of Dublin: exploring their character, chronology and evolving function in the light of excavations across Ireland*. Vol. XVII, in *Medieval Dublin*, 17-127.
- Harney, L. 2020. "Cemetery demographics as evidence for the monastic and pastoral character of the early Irish church." *Ulster Journal of Archaeology* 75, 2019-20 (Third Series): 34-51.
- Henderson, G. " in (eds) 2003,. 2003. "What (else) we talk about when we talk about landscape: for a return to the social imagination: Cultural Landscape Studies After J.B. Jackson." In *Everyday America: Cultural Landscape Studies After J.B. Jackson*, edited by P. Groth and C. Wilson, 178-198. University of California Press.
- Henry, F. 1965. *Irish Art in the Early Christian Period (to 800 AD)*. London: Methuen.
- Hesse, R. 2010. "LiDAR-derived Local Relief Models (LRM) – a new tool for archaeological prospection." *Archaeological Prospection* 17: 67–72.
- Hesse, R. 2012. "Using lidar-derived Local Relief Models (LRM) as a new tool for archaeological prospection." By Landscape & Heritage Series, Proceedings Landscape Archaeology between Art and Science – From a multi- to an interdisciplinary approach, edited by S.J. Kluiving and E.B. Guttman-Bond, 369- 378. Amsterdam: Amsterdam University Press.
- Johnston, S. A., P. J. Crabtree, and D. V Campana. 2014. "Performance, place and power at Dún Ailinne, a ceremonial site of the Irish Iron Age." *World Archaeology* 46 (2): 206–223.
- Kakoulaki, G., A. Martinez, and P Florio. 2021. *Non-commercial Light Detection and Ranging (LiDAR) data in Europe*. JRC Technical Report. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC126223/jrc126223_jrc126223_lidar_opensourcedata.pdf.
- Kokalj, Ž., and R. Hesse. 2017. *Airborne laser scanning raster data visualization: A Guide to Good Practice*. Ljubljana: Založba ZRC.
- Kokalj, Ž., K. Zakšek, and K. Oštir. 2010. "Archaeological Application of an Advanced Visualisation Technique Based on Diffuse Illumination." Edited by R. Reuter. *Proceedings of the 30th EARSeL Symposium: Remote Sensing for Science, Education and Culture*. Paris: EARSeL. 113-120.
- Kokalj, Z., K. Zaksek, and K. Ostir. 2013. "Visualizations of lidar derived relief models." In *Interpreting Archaeological Topography. 3d Data, Visualization and Observation*, edited by R.S. Opitz and D.C Cowley, 100–114. Oxford: Oxbow Books.
- Kostyrko, M., and G. Kiarszys. 2019. "Cultural landscape as palimpsest revisited." *Aerial Archaeology Research Group News*, April: 12-14.
- Kuijt, I., R. Lash, M. Gibbons, J. Higgins, N. Goodale, and J. O'Neill. 2010. "Reconsidering early medieval seascapes: new insights from Inis Airc, Co. Galway, Ireland." *The Journal of Irish Archaeology* 19: 51–70.
- Lambrick, G. 2008. *Air and Earth: Aerial Archaeology in Ireland*. The Heritage Council.

- Lawlor, H. J. 1895. "Notes on some non-Biblical matter in the MS of the Gospels known as the Book of Mulling." *Proceedings of the Society of the Antiquaries of Scotland* 29: 11-145.
- Lewis, H., C. Gallagher, S. R. Davis, J. Turner, G. Foster, and et al. 2009. *An integrated, comprehensive GIS model of landscape evolution and land use history in the River Boyne valley: Phase 2 report. Unpublished report for the Heritage Council (INSTAR aw. Unpublished report for the Heritage Council (INSTAR award 16711).*
- Lewis, H., C. Gallagher, S. R. Davis, J. Turner, G. Foster, and et al. 2018. *Developing a model of landscape evolution and land-use history of the River Boyne*. Vol. 6, in *Excavations at Knowth*, by G. Eogan and K. Cleary, 570-579. Dublin: Royal Irish Academy.
- Manning, C., P. Gosling, and J. Waddell. 2007. *New Survey Of Clare Island: Vol.5: Archaeology*. Dublin: Royal Irish Academy.
- Manning, Conleth. 1991. . Dublin : Stationery Office.
- Marshall, S., and T. McMorran. 2016. "People and places: a bird's-eye view of medieval Old Ross." In *Medieval Wexford: Essays in memory of Billy Colfer*, edited by I. Doyle and B. Browne, 241-261. Four Courts.
- McNeary, R. W. A. 2014. " Lidar Investigation of Knockdhu Promontory and its Environs, County Antrim, Northern Ireland." *Archaeological Prospection* 21: 263-276.
- McNeary, R..W.A. 2011. "Archaeological feature detection using airborne laser scanning at Shane's Castle Park, County Antrim." *Ulster Journal of Archaeology* 70: 111-125.
- Megarry, W. 2010. *LiDAR (Light Detection and Ranging) and GIS Survey Report: Proposed N4 Carrick-on-Shannon to Dromod Roadscheme, County Roscommon and Leitrim*. Unpublished report for Leitrim County Council.
- . 2011. "Etched in the Landscape." *Seanda*, 32-35.
- Mitchell, D. 2003 ' in. 2003. "Cultural Landscapes: just landscapes or landscapes of justice?." *Progress in Human Geography* 27 (6): 787-796.
- Motherway, E. 2012. "LiDAR and its potential contribution to archaeological survey in Ireland. Unpublished MPhil. thesis, Department of Archaeology, University College Cork."
- Needham, S., and G. Anelay. 2021. *Barrows at the core of Bronze Age communities*. Leiden: Sidestone Press.
- Newman, C. 1997. *Tara: An archaeological survey*. Vol. Discovery Programme Monographs No.2. Royal Irish Academy.
- Ní Lionáin, C., and S.R. Davis. 2016. "Conservation plan for the cultural heritage of the Devenish Lands at Dowth. Unpublished report for Devenish Nutrition."
- Ó Carragáin, T. and Turner, S., ed. 2017. *Making Christian landscapes in Atlantic Europe. Conversion and consolidation in the early Middle Ages*. Cork: Cork University Press.
- Ó Carragáin, T. 2021. *Churches in the Irish Landscape AD 400-1100*. Cork: Cork University Press.
- Ó Ríordáin, S. P. 1942. *Antiquities of the Irish Countryside*. 5th. Edited by R de Valera. London: Routledge.
- O'Brien, W. (2017). 2017. "The development of the hillfort in prehistoric Ireland." *Proceedings of the Royal Irish Academy: Archaeology, Culture, History, Literature* 117C: 3-61.
- O'Connor, K. D. 1998. *The Archaeology of Medieval Rural Settlement in Ireland*. Dublin:: Discovery Programme/Royal Irish Academy.
- O'Drisceoil, C., M. Leenene, S.R. Davis, B. Fitzgibbon, and M. Teehan. 2014. "The Black Pig's Dyke Regional Project 2014." https://www.longfordlibrary.ie/heritage/projects/black-pigs-dyke/bpd_phase_one_report2014_vol1.pdf.
- O'Driscoll, J., P. Gleeson, and G. Noble. 2020. "Re-imagining Navan Fort: new light on the evolution of a major ceremonial centre in Northern Europe." *Oxford Journal of Archaeology* 39: 247-273.
- O'Keeffe, T. 2000. *Medieval Ireland: an archaeology*. Stroud: Tempus.
- O'Keeffe, T. 2018. "Theory and medieval archaeology in Ireland and beyond: the narrative tradition and the archaeological imagination." *The Journal of Irish Archaeology* 27: 99-116. <https://www.jstor.org/stable/10.2307/26564146>.
- O'Brien, W., and J. O'Driscoll. 2017. *Hillforts, Warfare and Society in Bronze Age Ireland*. Oxford: Archaeopress.
- Ordnance Survey of Ireland, Nualláin, S.Ó. and De Valera, R., . 1961. *Survey of the Megalithic Tombs of Ireland*. Dublin: Stationery Office.
- O'Sullivan, A., McCormick, F., Kerr, T. R., & Harney, L. 2014. *Early Medieval Ireland AD 400-1100 : the Evidence from Archaeological Excavations*. Portland: Royal Irish Academy.
- Roughley, C., E. S. Twohig, C. Shell, and G. Swanton. 2021. "The changing fieldscapes of Loughcrew: New Insights from airborne lidar." In *Europe's Early Fieldscales. Themes in Contemporary Archaeology*, edited by S. Arnoldussen, R. Johnston and M. Løvschal. Springer.
- Sevara, C., M. Pregesbauer, M. Doneus, G. Verhoeven, and I. Trinks. 2016. "Pixel versus object - A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data." *Journal of Archaeological Science: Reports* 5: 485-498.
- Shaw, R, and A Corns. 2011. *High resolution LiDAR specifically for archaeology: Are we fully*

- exploiting this valuable resource. Vol. EAC Occasional Paper No. 5, in *Remote Sensing for Archaeological Heritage Management*, edited by David C Cowley, 77-86. Belgium: Europae Archaeologia Consilium (EAC).
- Shell, C. 2005. "The Loughcrew Landscape Project." 1-3. <https://www.arch.cam.ac.uk/system/files/documents/past051.pdf>.
- Shell, C., and C. Roughley. 2004. "Exploring the Loughcrew Landscape: A New Airborne Approach." *Archaeology Ireland*, 22-25. <http://www.jstor.org/stable/20558973>.
- Shortle, G., and P. Jordan. 2017. *Agricultural Catchments Programme*. Teagasc. https://www.teagasc.ie/media/website/publications/2018/ACP_Phase_2_Report.pdf.
- Stalley, R. 1987. *The Cistercian Monasteries of Ireland: an account of the history, art and architecture of the White Monks in Ireland from 1142 to 1540*. New Haven; London: Yale University Press.
- Stout, M. 1992. "Plans from Plans: An Analysis of the 1:2500 OS Series as a Source for Ringfort Morphology." *Proceedings of the Royal Irish Academy*, 37-53. <https://www.jstor.org/stable/25516094>.
- Stout, M. 2000. "Settlement and environment. A history of settlement in Ireland." In *A History of Settlement in Ireland*, edited by T. Barry, 81-109. London: Routledge.
- Swan, R. 1983. "Enclosed ecclesiastical sites and their relevance to settlement patterns in the first millennium A.D." In *Landscape archaeology in Ireland*, edited by T. Reeves-Smyth and F. Hamond, 269-294. Oxford: BAR Publishing.
- Sweetman, D. 1999. *The Medieval Castles of Ireland*. Woodbridge: Boydell.
- Thomas, J. 1993. "The Politics of Vision and the Archaeologies of Landscape." In *Landscape: politics and perspectives*, edited by B. Bender, 19-48. New York: Berg.
- Tilley, C. 1999. *Metaphor and Material Culture*. Oxford: Blackwell.
- Tilley, C. 2005. "Phenomenological Archaeology." In *Archaeology: The Key Concepts*, edited by Colin Renfrew and Paul Bahn, 151-154. London: Routledge.
- . 1994. *Phenomenology of Landscape*. Oxford: Berg.
- Traviglia, A., D.C. Cowley, and K. Lambers. 2016. "Finding common ground : human and computer vision in archaeological prospection." *AARGnews*, 11-24. <https://a-a-r-g.eu/wp-content/uploads/2019/02/AARGnews53-1.pdf>.
- Twohig, E. S., C. Roughley, C. Shell, C. O'Reilly, P. Clarke, and G. Swanton. 2010. "Open-air rock art at Loughcrew, Co. Meath." *The Journal of Irish Archaeology* 19: 1-28.
- van der Vaart, W.B., and K. Lambers. 2021. "Applying automated object detection in archaeological practice: A case study from the southern Netherlands." *Archaeological Prospection* 29 (1). doi:10.1002/arp.1833.
- Walsh, P. 1993. "In Circle and Row: Bronze Age Ceremonial Monuments." In *Past perceptions: the prehistoric archaeology of south-west Ireland*, edited by E and Roynane, M. In Shee Twohig, 101-113. Cork University Press Cork.

Appendices

17. Appendix i: Additional machine learning resources

Method – RCNN	
Paper	https://arxiv.org/abs/1311.2524v5
Resources	https://paperswithcode.com/paper/rich-feature-hierarchies-for-accurate-object
Implementation	https://github.com/rbgirshick/rcnn

Method – Faster RCNN	
Paper	https://arxiv.org/abs/1506.01497v3
Resources	https://paperswithcode.com/paper/faster-r-cnn-towards-real-time-object
Implementation	https://github.com/facebookresearch/detectron

Method – YOLO	
Paper	https://arxiv.org/pdf/1506.02640v5.pdf
Resources	https://paperswithcode.com/paper/you-only-look-once-unified-real-time-object
Implementation	https://github.com/AlexeyAB/darknet

Method - YOLOR	
Paper	https://arxiv.org/abs/2105.04206
Resources	https://paperswithcode.com/paper/you-only-learn-one-representation-unified
Implementation	https://github.com/WongKinYiu/yolor

Method - DeepLabV3	
Paper	https://arxiv.org/pdf/1706.05587v3.pdf
Resources	https://paperswithcode.com/method/deeplabv3
Implementation	https://github.com/tensorflow/models/tree/master/research/deeplab

Method - HRNET	
Paper	https://arxiv.org/abs/1908.07919v2
Resources	https://paperswithcode.com/method/hrnet
Implementation	https://github.com/HRNet/

Method - Segmentation Transformers based on HRNET	
Paper	https://arxiv.org/abs/1909.11065v6
Resources	https://paperswithcode.com/paper/object-contextual-representations-for
Implementation	https://github.com/open-mmlab/mmdetection ; https://github.com/HRNet/HRNet-Semantic-Segmentation

Method - Swin Transformers	
Paper	https://arxiv.org/abs/2111.09883v1 ; https://arxiv.org/abs/2103.14030v2
Resources	https://paperswithcode.com/paper/swin-transformer-v2-scaling-up-capacity-and ; https://paperswithcode.com/method/swin-transformer
Implementation	https://github.com/microsoft/Swin-Transformer

Method - SpotNet	
Paper	https://arxiv.org/abs/2002.05540v2
Resources	https://paperswithcode.com/paper/spotnet-self-attention-multi-task-network-for
Implementation	https://github.com/hu64/SpotNet

18. Appendix ii: Monument classifications or features that may be considered to be 'Enclosures'

Monument type	Primary Period of Construction	Primary Structural Elements
Prehistoric ceremonial enclosures: Henge or hengiform monuments	Neolithic/Bronze Age Iron Age	Typically a ditch enclosing a circular or sub-circular/oval space, which in turn is enclosed by a bank (and frequently considered a non-defensive structure in any traditional sense), though in Irish examples the ditch may be absent. Size varies, with examples in excess of 100m diameter, and banks up to 4m+ in height. Few 'new' upstanding examples have been discovered in recent years, but several demolished/levelled sites known
Ring Ditch (burial)	Bronze Age; Potentially Iron Age	Complete or segmented shallow ditch enclosing a circular area which may present as a mound or flat space; size varies from 2-12m diameter, ditch width from .3-1.5m
Ring Ditch (structure)	Bronze Age; Potentially Iron Age Early Medieval	Primarily associated with features that delineate circular houses, whereby the ditch may be a deliberately-constructed drainage feature within a building, or around the perimeter of a circular building; size may vary from 4-8m diameter, ditch width 0.2-1m A further type are ditches or foundation trenches for timber structures of similar dimension, but which may include larger structures such as those identified at the major 'Royal' Iron Age sites of Ireland
Ceremonial, Religious or Major/large social spaces	Multiple periods, but primarily Neolithic, Bronze Age and Iron Age	Ranges from henges/hengiform monuments noted above, hill forts from the Bronze Age and later, to the enclosures at Iron Age and Early Medieval tribal centres, comprising banked and ditches enclosures in excess of 100m in diameter, and frequently multiple circuits of enclosing elements.
Church Enclosure	c.400AD onwards	Primarily curvilinear enclosures associated with early church foundations, pre-twelfth century reform, often reused as burial grounds, settlements or industrial areas associated with the ecclesiastical foundation. Scale can vary from c.20m-100m+ in diameter circular/sub-circular areas enclosed by a bank or drystone which may vary in size from 1-5m wide and 0.5-3m high, and on occasion by an external ditch. Some larger sites have multiple, usually concentric, enclosing elements of banks and ditches with open ground in between.

Monument type	Primary Period of Construction	Primary Structural Elements
		Late medieval and post-medieval examples may have rectilinear enclosures demarcating the extent of formal graveyards.
Ringfort/Ráth	c.600AD-c.1100AD	The most common field monument of the early medieval period, typically 20-40m diameter circular space enclosed by a bank 5m wide and 3m high, and an external ditch 2-5m deep. Multiple concentric closely-spaced enclosing elements may be present at some sites. There are, however, examples that are significantly larger than the 'average' dimensions noted above, up to 75m in diameter. Ploughed-down examples may only present as the filled-in remains of the enclosing ditches and present a very subtle difference in levels of the modern ground surface.
Longfort (Viking and later)	c.800AD-c.1550AD	Frequently constructed along water edges and thus formed as an arc of an enclosing bank or three sides of a rectilinear enclosure, the fourth/remaining arc composed by the water edge (river, lake)
Earthwork and masonry fortifications or derivations thereof.	c.1169AD-c1700AD	Present as mounds (mottes), with enclosures either being the enclosing ditches around those mounds, or as attendant enclosures that form appendages to the mound (as formal baileys or otherwise enclosed spaces). Sizes vary, though primary enclosing elements may be some 20-40m in diameter, with attendant sub-circular/rectilinear enclosures 40-60m in diameter/breadth-length. Ringworks and moated sites follow similar morphologies, though may be difficult to distinguish from ringforts or grange enclosures; similar dimensions. Late- and post-medieval examples are frequently composed of a primary building (tower house or fortified house) with an attached walled/bawn enclosure.
Precinct	c.1100AD-c.1550AD	Similar to the enclosures identified at early church sites, and perhaps of same function, but primarily the rectilinear enclosures associated with conventual religious orders and foundations; may be several hundred meters in length or breadth, with enclosing elements of earthworks or masonry
Burial Ground	c.400AD-today	Similar to the enclosures identified at early church sites, and frequently presented as circular/sub-circular enclosed spaces. In many instances these are early church foundations, re-used ringforts, or sites that follow a similar morphology by design.

Monument type	Primary Period of Construction	Primary Structural Elements
Military structures	c.1550AD-1700AD (Post-dissolution of monasteries in Ireland; Tudor Wars/Nine Years War; Jacobean settlement; Confederacy; Williamite Wars)	Multiple site types are included here, from campaign forts and simple rectilinear enclosures formed of earthworks, palisades or walls (or composites of all of the above). Key features that facilitate identification are straight line components that form structures such as bastions at the corners of the enclosures.
Town defences	c.850AD-c.1690AD	Similar to military structures above, with composition varying from earthworks of early town defences or defences erected at times of particular stress, to architecturally accomplished masonry structures including defensive plans especially at corner locations.
Farmland field enclosures	Any period	Any form of linear, rectilinear, circular or curvilinear land divisions which divide the landscape into a network of fields, routeways and settlement nodes.

19. Appendix iii: Classifications of Barrows as contained in the current Archaeological Survey of Ireland's Class List Definitions

<i>Barrow - bowl-barrow</i>	A circular or oval raised area (generally over 1m above the external ground level) with an external fosse and sometimes an outer bank. The name 'bowl-barrow' refers to the mound element which is like an inverted bowl. They contain and/or cover burials and were in use from the Bronze Age to the Iron Age (c. 2400 BC - AD 400).
<i>Barrow - ditch barrow</i>	A circular or oval level or slightly raised area (less than 1m above the external ground level) defined by a fosse; generally less than 20m in diameter. They are often found in clusters or in association with other barrow types. They are funerary monuments that may date to the Neolithic (c. 4000-2400 BC) but more usually date to the Bronze/Iron Ages (c. 2400 BC - AD 400). See also Ring-ditch.
<i>Barrow - embanked barrow</i>	A circular or oval raised area, generally less than 20m in diameter, enclosed by a continuous broad/large bank with a level or a concave interior. Generally there is no external fosse and no entrance. The proportion of the bank size (large) in relation to the internal diameter (small) is important in identifying this site type. These are part of the Bronze/Iron Age burial tradition (c. 2400 BC - AD 400).
<i>Barrow - mound barrow</i>	A circular or oval earthen or earth and stone mound with no external features. Mounds found in association with other barrow types are likely to be mound barrows. They are funerary in nature and contain and/or cover burials. Excavated examples have been dated to the Bronze and Iron Ages (c. 2400 BC - AD 400).
<i>Barrow - pond barrow</i>	A shallow, man-made, circular depression enclosed around its rim by an earthen bank. These are part of the Bronze/Iron Age burial tradition (c. 2400 BC - AD 400).
<i>Barrow - ring-barrow</i>	A circular or oval raised area (generally up to 1m above the external ground level or level with it) enclosed by fosse(s) and outer bank(s), with or without an entrance. These are part of the Bronze/Iron Age burial tradition (c. 2400 BC - AD 400).
<i>Barrow - stepped barrow</i>	An oval or circular platform with a raised, flat-topped or rounded central area, giving the monument its characteristic 'stepped' profile, sometimes with a bank on the outer edge of the platform. These are part of the Bronze/Iron Age burial tradition (c. 2400 BC - AD 400).
<i>Barrow - unclassified</i>	An artificial mound of earth or earth and stone, normally constructed to contain or conceal burials. Used where it is not possible to identify the specific type. See also Barrow - bowl-barrow, Barrow - ditch barrow, Barrow - embanked barrow, Barrow - mound barrow, Barrow - pond barrow, Barrow - ring-barrow and Barrow - stepped barrow. These are part of the Bronze/Iron Age burial tradition (c. 2400 BC - AD 400).

20. Appendix iv: Classifications of Ecclesiastical enclosures, Enclosures, Large Enclosures and Ringforts as contained in the current Archaeological Survey of Ireland's Class List Definitions

Classification	Definition
<i>Ecclesiastical enclosure</i>	A large oval or roughly circular area, usually over 50m in diameter, defined by a bank/banks and external fosse/fosses or drystone wall/walls, enclosing an early medieval church or monastery and its associated areas of domestic and industrial activity. These date to the early medieval period (5th-12th centuries AD).
<i>Enclosure</i>	An area defined by an enclosing element (e.g., bank, wall, fosse, scarp), or indicated as such cartographically, and occurring in a variety of shapes and sizes, possessing no diagnostic features which would allow classification within another monument category. These may date to any period from prehistory onwards. Enclosures with a diameter greater than 70m should be classed as Large Enclosure.
<i>Enclosure - large enclosure</i>	An enclosed area over 70m in diameter defined by an earthen bank or stone wall, sometimes with an external fosse. They display no preferred siting criteria, and those on hilltops can be classified as hilltop enclosures. They can date to any period from prehistory to the late medieval period (i.e., 4000 BC - 1700 AD). For other large enclosures see also: Hillfort, Hilltop enclosure, Causewayed enclosure, Ceremonial enclosure, and Embanked enclosure.
<i>Ringfort - cashel</i>	Usually a roughly circular or oval area, though some examples are rectangular, which is surrounded by a stone wall or walls. They functioned as residences and/or farmsteads and broadly date from 500 to 1000 AD. See Ringfort - rath for earthen equivalent.
<i>Ringfort - rath</i>	A roughly circular or oval area surrounded by an earthen bank with an external fosse. Some examples have two (bivallate) or three (trivallate) banks and fosses, but these are less common and have been equated with higher status sites belonging to upper grades of society. They functioned as residences and/or farmsteads and broadly date from 500 to 1000 AD. See Ringfort – cashel.
<i>Ringfort - unclassified</i>	A roughly circular or oval area surrounded by an earthen bank with an external fosse (see Ringfort - rath) or a stone wall (see Ringfort - cashel). The term Ringfort - unclassified is used in instances where the surviving remains are insufficient to determine whether the monument was originally a rath or cashel. They functioned as residences and/or farmsteads and broadly date from 500 to 1000 AD.