



Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach

Alexandre Guyot, Marc Lennon, Thierry Lorho, Laurence Hubert-Moy

► To cite this version:

Alexandre Guyot, Marc Lennon, Thierry Lorho, Laurence Hubert-Moy. Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach. *Journal of Computer Applications in Archaeology*, 2021, 4 (1), pp.1. 10.5334/jcaa.64 . hal-03166011

HAL Id: hal-03166011

<https://hal.science/hal-03166011>

Submitted on 11 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach

ALEXANDRE GUYOT 

MARC LENNON

THIERRY LORHO

LAURENCE HUBERT-MOY

**Author affiliations can be found in the back matter of this article*

RESEARCH ARTICLE

]u[ubiquity press

ABSTRACT

Until recently, archeological prospection using LiDAR data was based mainly on expert-based and time-consuming visual analyses. Currently, deep learning convolutional neural networks (deep CNN) are showing potential for automatic detection of objects in many fields of application, including cultural heritage. However, these computer-vision based algorithms remain strongly restricted by the large number of samples required to train models and the need to define target classes before using the models. Moreover, the methods used to date for archaeological prospection are limited to detecting objects and cannot (semi-)automatically characterize the structures of interest. In this study, we assess the contribution of deep learning methods for detecting and characterizing archeological structures by performing object segmentation using a deep CNN approach with transfer learning. The approach was applied to a terrain visualization image derived from airborne LiDAR data within a 200 km² area in Brittany, France. Our study reveals that the approach can accurately (semi-)automatically detect, delineate, and characterize topographic anomalies, and thus provides an effective tool to inventory many archaeological structures. These results provide new perspectives for large-scale archaeological mapping.

CORRESPONDING AUTHOR:

Alexandre Guyot

Laboratoire LETG – UMR 6554,
Université Rennes 2, FR

alexandre.guyot@univ-rennes2.fr

KEYWORDS:

archaeological prospection;
airborne laser system; transfer
learning; object segmentation;
remote sensing

TO CITE THIS ARTICLE:

Guyot, A, Lennon, M, Lorho, T and Hubert-Moy, L. 2021. Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach. *Journal of Computer Applications in Archaeology*, 4(1), pp. 1–19. DOI: <https://doi.org/10.5334/jcaa.64>

1. INTRODUCTION

The past decade has seen an increasing interest in remote-sensing technologies and methods for monitoring cultural heritage. One of the most relevant changes is the development of airborne light detection and ranging (LiDAR) systems (ALS). With the ability to measure topography accurately and penetrate the canopy, ALS has been a key tool for important archaeological discoveries and a better understanding of past human activities by analyzing the landscape (Bewley, Crutchley and Shell 2005; Chase et al. 2011; Evans et al. 2013; Inomata et al. 2020) in challenging environments.

Most archaeological mapping programs based on ALS do not use LiDAR 3D point clouds directly, but use instead derived elevation models that represent bare soil in the topographic landscape. Perception of the terrain is usually enhanced by specific visualization techniques (VT) (Bennett et al. 2012; Devereux, Amable and Crow 2008; Doneus 2013; Hesse 2010; Štular et al. 2012) that are used to visually interpret landforms and archaeological structures (Kokalj and Hesse 2017). These visualizations have resulted in better understanding of the human past in different periods and different regions of the world. For example, LiDAR-derived terrain combined with VT has been used to provide new insights into a prehistoric hillfort under a woodland canopy in England (Devereux et al. 2005), discover a pre-colonial capital in South Africa (Sadr 2019), supplement large-scale analysis of a human-modified landscape in a Mayan archaeological site in Belize (Chase et al. 2011) and explore long-term human-environment interactions within the former Khmer Empire in Cambodia (Evans 2016). However, these expert-based and time-consuming approaches are difficult to replicate in large-scale archaeological prospection projects.

A variety of (semi-)automatic feature-extraction methods have been developed to assist or supplement these visual interpretation approaches. Object-based image analysis (Freeland et al. 2016) and template-matching (Trier and Pilø 2012) methods, which rely on prior definition of purpose-built spatial descriptors or prototypical patterns, respectively, are difficult to generalize because they cannot include the high morphological diversity and heterogeneous backgrounds of archaeological structures (Opitz and Herrmann 2018). Supervised machine-learning methods have been assessed to address these limitations (Lambers, Verschoof-van der Vaart and Bourgeois 2019). Data-driven classifiers (e.g. random-forest, support vector machine) applied to multi-scale topographic or morphometric variables have provided interesting results for detecting archaeological structures (Guyot, Hubert-Moy and Lorho 2018; Niculiță 2020). However, detection was either performed at the pixel level without considering the target as an entire object (archaeological structure) with spatial aggregation and internal complexities, or was based on previous image

segmentation, which prevents them from being applied to complex structures. In recent years, deep learning Convolutional Neural Networks (deep CNNs) have resulted in a new paradigm in image analysis and provided groundbreaking results in image classification (Krizhevsky, Sutskever and Hinton 2012) or object detection (Girshick 2015). Deep CNNs are composed of multiple processing layers that can learn representations of data with multiple levels of abstraction (LeCun, Bengio and Hinton 2015). In the context of LiDAR-based archaeological prospection, they were first applied in 2016 (Due Trier, Salberg and Holger Pilø 2016) to detect charcoal kilns and were further evaluated in different archaeological contexts and configurations (Caspari and Crespo 2019; Gallwey et al. 2019; Kazimi et al. 2018; Trier, Cowley and Waldeland 2018; Verschoof-van der Vaart et al. 2020; Verschoof-van der Vaart and Lambers 2019). These studies focused on image classification (predicting a label/class associated with an image) (*Figure 1a*) or object detection (predicting the location (i.e. bounding box (BBOX)) of one or several objects of interest within the image) (*Figure 1b*). While these deep CNN methods have detected archaeological structures adequately, they could not provide information that (semi-)automatically characterized them because structures must be delineated to move from detection to characterization. Recent deep CNN methods, such as Mask R-CNN (He et al. 2017), have object-segmentation abilities (*Figure 1c*) that delineate objects. These deep CNN methods remain strongly restricted by the large number of samples required to train models and the need to define target classes before using the models. While the lack of ground-truth samples (reference data) is a known constraint in remote-sensing archaeological prospection, two strategies can address this limitation: transfer learning and data augmentation. The first strategy applies a pre-trained source domain model to initialize a targeted domain model (Weiss, Khoshgoftaar and Wang 2016), while the second strategy uses transformers that modify input data for training. These strategies are known to improve model performance for small datasets and to increase model generalization (Shorten and Khoshgoftaar 2019). Defining target classes before using a model is based on one-class approaches that define only a generic “archeological structure” class without dividing it into several sub-classes, assuming that the object characterization can identify types of archaeological structures.

Using deep CNN for archaeological prospection of LiDAR derived-terrain (Caspari and Crespo 2019; Gallwey et al. 2019; Küçükdemirci and Sarris 2020; Soroush et al. 2020; Trier, Cowley and Waldeland 2018; Verschoof-van der Vaart et al. 2020; Verschoof-van der Vaart and Lambers 2019) is in its infancy, and to our knowledge, these studies have not evaluated the object-segmentation abilities of the CNN, except the evaluation of Mask R-CNN for simple circular-based

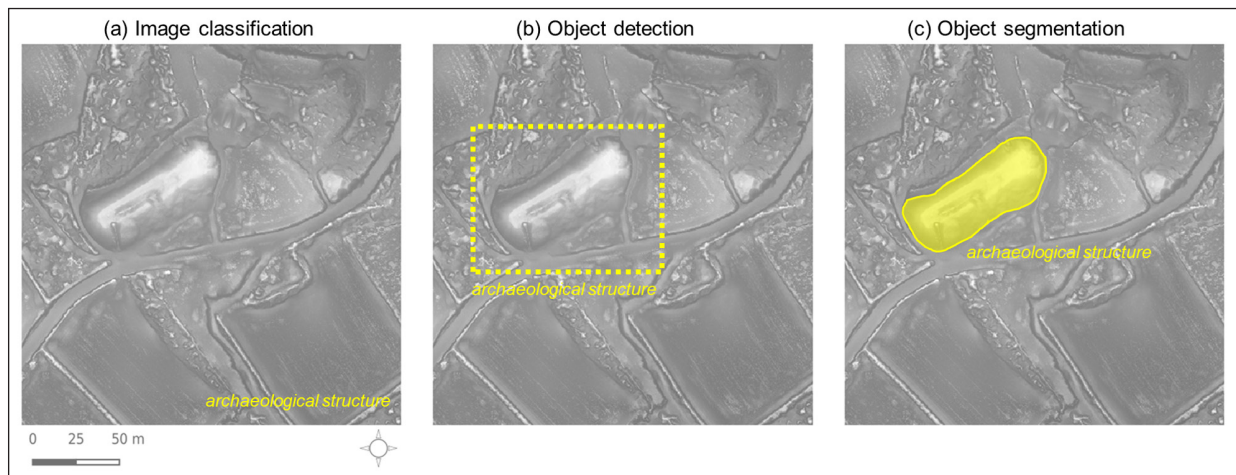


Figure 1 Image analysis using deep learning Convolutional Neural Networks for an archaeological site (Tumulus du Moustoir, Carnac, France). **(a)** Image classification: a class or label associated with the image, **(b)** Object detection: a labeled bounding box that locates the object of interest within the image, and **(c)** Object segmentation: a labeled footprint that locates and delineates the object of interest within the image.

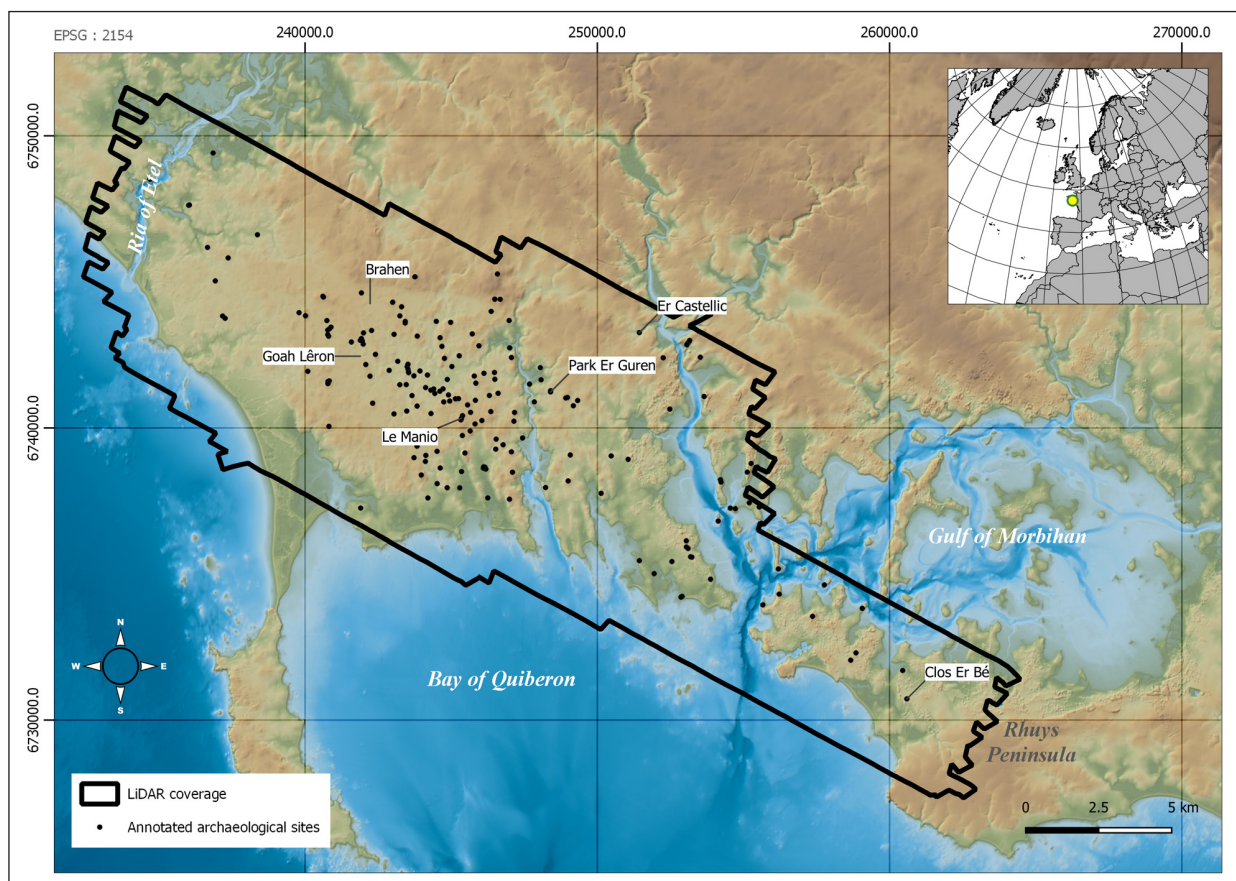


Figure 2 The study area with the location of the 195 annotated archaeological sites used in the study. Areas mentioned in the text are labeled.

landforms (Kazimi, Thiemann & Sester 2019; Kazimi, Thiemann & Sester 2020). In the present study, we assess the contribution of deep CNN to the combined detection and segmentation of archeological structures for further (semi-)automatic characterization.

More specifically, we aim to provide new insights into object segmentation using deep CNN for archaeological prospection to address two key issues: i) the extent to which the approach is sensitive to the amount of sample

data, since data are a sparse resource in archaeology, and ii) after object detection, the utility of object segmentation for characterizing archaeological structures.

2. MATERIALS AND METHODS

2.1. STUDY AREA

The study area (*Figure 2*) is located in southern Morbihan (Brittany, France) and covers an area between the Ria

of Etel and the Rhuys Peninsula on the Atlantic coast. The region is a complex and fragmented mosaic of landscapes. The hinterland is composed of woodlands, moorlands and farmland that form a rural environment oriented to agriculture. The coastal area is also diverse, with estuaries and small islands near the intricate Gulf of Morbihan and large open, sandy areas in the Bay of Quiberon that concentrates most of the economic activities of tourism and fisheries.

The area is home to a unique megalithic heritage. Erected between the 5th to 3rd millennia BC, the Neolithic architecture (standing stones and megalithic tombs) represents an exceptional corpus of archaeological sites that are candidates for the UNESCO World Heritage List. Beyond this emblematic heritage, the coast of Morbihan includes a wide variety of archaeological sites that marked the gentle topography of the area and encompass different prehistorical and historical periods.

2.2. DATASET

2.2.1. LiDAR-derived visualization image

The workflow for processing LiDAR data consisted of several steps (Figure 3). The image dataset was derived from a LiDAR point-cloud collected over the area in 2016 (200 km², excluding water area). The raw point-cloud was collected from a bispectral (1064 and 532 nm) Optech Titan LiDAR sensor operated from a fixed-wing vector 1300 m above ground level at a pulse repetition frequency of 300 kHz per channel and a 26° field of view to obtain a nominal point density of 14 points/m². The 3D point-cloud recorded was processed with LasTools (rapidlasso GmbH, Gilchin, Germany) to perform ground-filtering and gridding operations to create a Digital Terrain Model (DTM) at a spatial resolution of 50 cm (Guyot, Hubert-Moy and Lorho 2018). The terrain model was then used to perform two VTs.

First, a multiscale topographic position (MSTP) image (Lindsay, Cockburn and Russell 2015) was created based on a previous archaeological prospection study

(Guyot, Hubert-Moy and Lorho 2018). The MSTP image was generated from a hyperscale datacube (30 bands corresponding to 30 window sizes) of the topographic metric DEV (deviation from mean elevation) (Wilson and Gallant 2000) and reduced to three dimensions by extracting the absolute maximum value from micro, meso, and macro scale ranges, which had window sizes of 3–21, 23–203 and 223–2023 px, respectively. Second, a morphological VT was created by combining a red-toned elevation gradient (slope) and a greyscale positive/negative topographic openness based on Chiba, Kaneta and Suzuki (2008). Finally, MSTP and morphological VT were blended into a single composite image using a soft-light blending mode with 100% and 70% opacity, respectively.

The resulting enhanced multiscale topographic position (eMSTP) image (Figure 4) was proposed as an optimal VT for this study. It provided effective and informative multiscale visualization of archaeological structures (Guyot, Hubert-Moy and Lorho 2018) and enhanced perception of local morphological characteristics of the terrain (a known limitation of MSTP (Guyot, Hubert-Moy and Lorho 2018)). A 3-channel image was used as input of the network to facilitate transfer-learning from models trained on natural RGB images.

eMSTP images were cropped from the overall mosaic as 150 images, 512 px × 512 px in size, to be input into the deep CNN architecture and cover the annotated archaeological sites.

2.2.2. Archaeological annotated reference data

The reference dataset consisted of 195 georeferenced polygons that represented footprints of known archaeological sites in the study area. The sites were selected from the regional archaeological reference dataset provided by the *Service Régional de l'Archéologie* (SRA Bretagne). Only archaeological structures of which topographic characteristics could be perceived on the LiDAR-derived DTM were kept (thus excluding sites related to small-object deposits, such as potsherds,

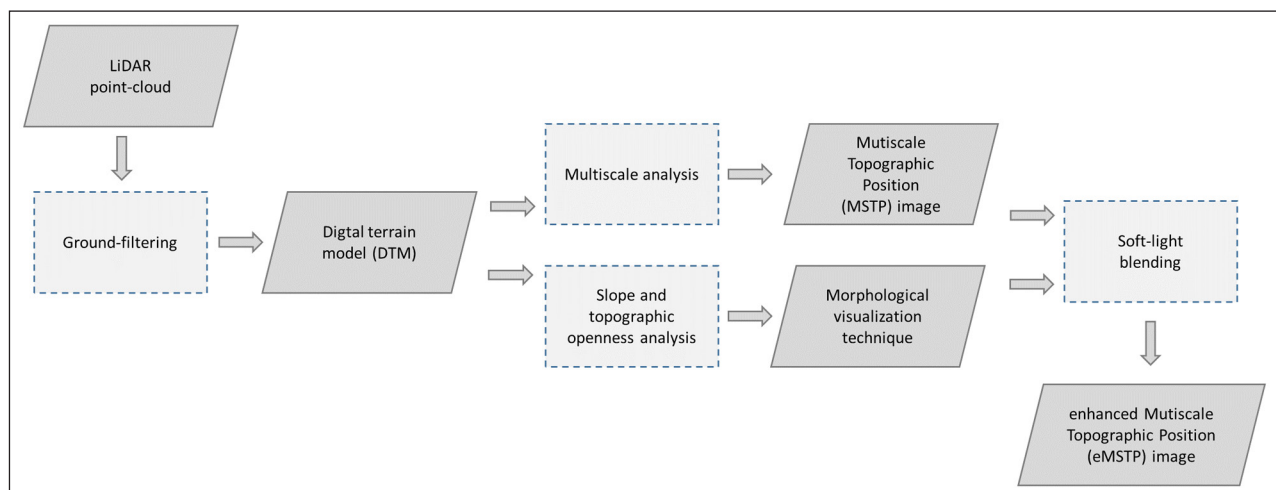


Figure 3 Image dataset's workflow from DTM to enhanced Multiscale Topographic Position (eMSTP) image.

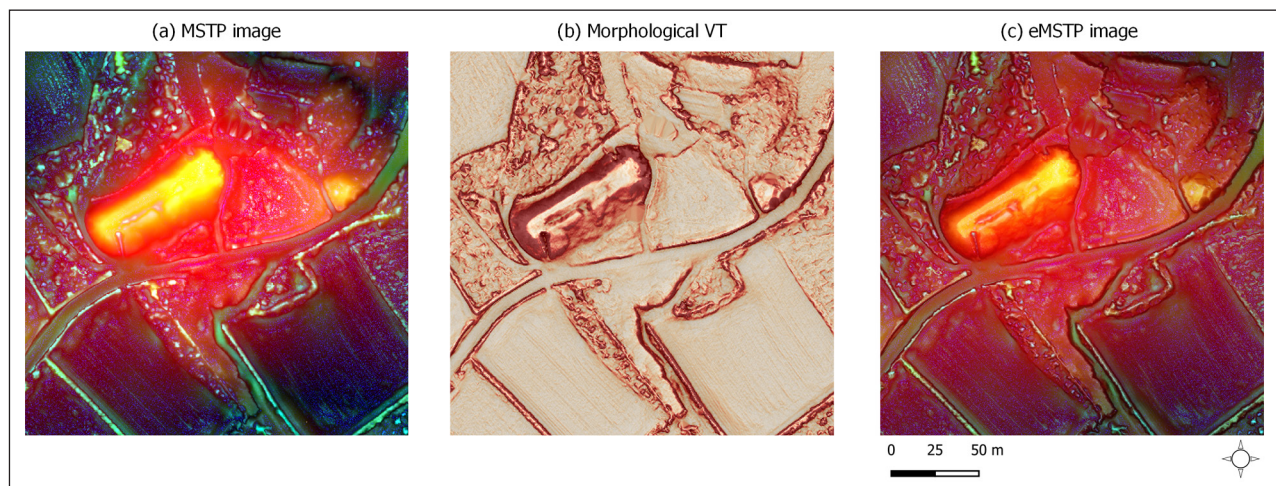


Figure 4 (a) Multiscale Topographic Position (MSTP) image, (b) Morphological visualization technique (VT) and (c) resulting enhanced multiscale topographic position (eMSTP) image of a Neolithic monument (Tumulus du Moustoir, Carnac, France).

and sites considered as above-ground structures with no influence on the bare-earth topography, such as standing stones).

The selected archaeological sites had diverse chronologies, morphologies, and landscape contexts. Their state of conservation also varied greatly, from long-known restored monuments to unexcavated little-documented structures. The reference dataset included 195 archaeological structures, including 176 funeral structures attributed to the Neolithic, 10 funeral structures attributed to protohistoric periods, 1 motte, 3 promontory forts and 5 ruined windmills.

Given the highly imbalanced dataset (over-representation of Neolithic structures) and the tasks to evaluate (object detection and segmentation), the annotations were intentionally grouped into a single “archaeological structure” class with no further distinction. The reference dataset was converted from a geospatial format to an annotation one (json COCO) in which each annotation was associated with its corresponding eMSTP tile to be input into the deep CNN architecture. Due to spatial proximity between some archaeological sites, 150 eMSTP images covered the 195 annotations (a mean of 1.3 annotations per image).

2.3. METHODS

2.3.1. Overall workflow

From the eMSTP images input, the overall workflow (Figure 5) of the approach consisted of two main parts:

- Object detection and segmentation
- Object characterization

2.3.2. (Semi-)automatic object detection and segmentation

We used the open-source implementation of Mask R-CNN developed by Matterport (Abdulla 2017). The feature-extraction phase (backbone) was performed using the Resnet-101 deep CNN initialized with weights pre-trained

on the COCO dataset (Lin et al. 2014) for the transfer-learning strategy.

To limit overfitting due to the small training dataset, data augmentation (DA) was activated in the Mask R-CNN workflow using the *imgaug* library (Jung et al. 2020). For each epoch, input images were randomly augmented with affine transformations (scaling: 80–120% of the original image size; translation: –20% to 20% of the original image position; 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 eMSTP images. 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.

A specific sampling strategy was used to assess the model’s stability (varying training/validation/test draws) and sensitivity to the number of training samples (varying training size). The initial dataset of 150 images was randomly split into 110, 20 and 20 images for training, validation and testing, respectively. This random split was performed 10 times to create 10 different experimental datasets (different draws). For each experimental dataset, the training dataset was divided into 11 sub-training datasets with 10–110 images, with an increment of 10. Given the number of experimental datasets and sub-training datasets, a total of 110 experimental configurations were available (see Appendix A.1). Each experimental configuration was checked to ensure that no leaks occurred between validation, test and training datasets. Many hyperparameters can be calibrated in Mask R-CNN. To reduce specific effects and focus on the generalized behavior of the model, only a few hyperparameters were configured. The Region Proposal Network (RPN) was configured to consider the size and aspect ratios of objects of interest by setting $RPN_ANCHOR_SCALES = [16, 32, 64, 128]$ (in px) and $RPN_ANCHOR_RATIOS = [0.5, 1, 2]$ (width/height ratio).

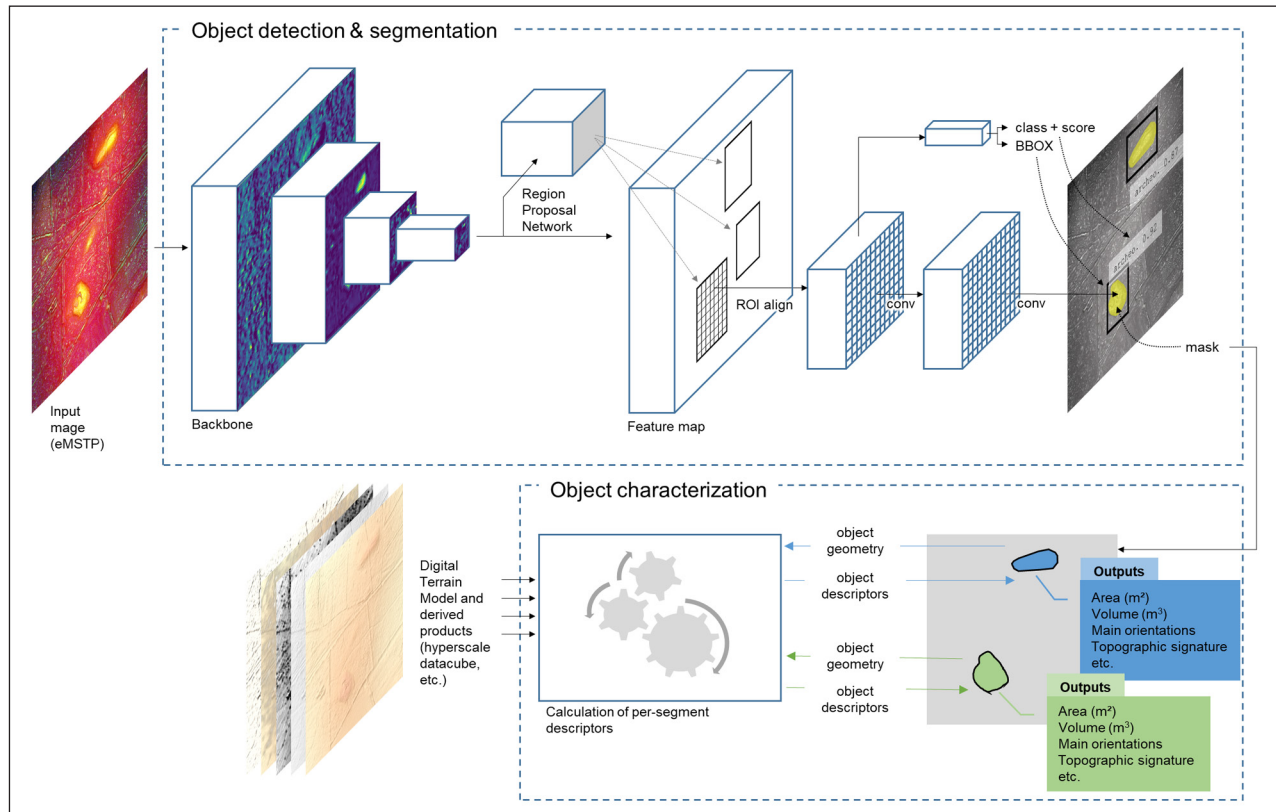


Figure 5 Overall workflow of (semi-)automatic object detection, segmentation, and characterization of archaeological structures. eMSTP = enhanced multiscale topographic position, ROI = region of interest, BBOX = bounding box.

The training was performed on 60 epochs with a decaying learning rate (LR) schedule (training stage 1:20 epochs at LR 10^{-3} , training stage 2:20 epochs at LR 10^{-4} , training stage 3:20 epochs at LR 10^{-5}). To consider the variability in training size (10–110 images, depending on the experiment), the number of iterations per epoch (*STEP_PER_EPOCH* parameter) was 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 ensured that the deep CNN observed each image (or its augmented version) only once per epoch.

The training process was set to fine-tune the head layers of the network (RPN, classifier and mask) (the other layers were frozen) to maximize use of transfer learning within the backbone network. The validation dataset was used to monitor the loss at the end of each epoch. For each experimental configuration, the model was run in inference mode to predict results from the test dataset (20 images). The inference returned a BBOX, confidence score and binary mask (or segment) for each object detected in the images of the test dataset.

Model performance was evaluated both statistically and visually. Predictions were assessed statistically per experiment by using metrics adapted to object detection and segmentation. The AP (average precision) for an IoU (intersection over union) threshold of 0.5 was used to assess each image and averaged as mAP to assess each dataset of the experimental configurations.

IoU refers to the overlapping score of the predicted mask compared to the reference data:

$$IoU = \frac{\text{area of intersection}}{\text{area of union}} \quad (1)$$

AP refers to the area under the precision-recall curve, with:

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

$$\text{recall} = \frac{TP}{TP + FN} \quad (3)$$

with *TP* and *FP* and *FN* the true positives, false positives and false negatives, respectively.

$mAP@IoU_v$ refers to the mean APs at a IoU threshold *v* for a given dataset with:

$$mAP@IoU_v = \frac{1}{n} \sum_{i=1}^n AP_i \quad (4)$$

with *n* the number of images *i* for a given dataset.

Visual analysis was then performed to compare reference data and model predictions for each image for three case studies.

To assess the approach within an archaeological prospection scheme, we trained an additional deep CNN model (the deployment model) using all possible reference data (i.e. 150 images). The deployment model was applied to an independent set of images of the study area that did not contain any known archaeological

structures that are topographically visible. The model was evaluated through human-interpretation and field survey.

2.3.3. Characterization of segmented objects

The results of (semi-)automatic detection and segmentation (i.e., predicted masks) were used to evaluate object characterization (morphological and contextual characterization). Predicted masks (polygons) were used as base units to calculate simple morphometric descriptors (Table 1) and extract hyperscale topographic position signatures of the segmented objects (see the *LiDAR-derived visualization image* section for details on the hyperscale datacube).

3. RESULTS

3.1. OBJECT DETECTION AND SEGMENTATION PERFORMANCES

3.1.1. Sensitivity of deep CNN to the amount of sample data

The overall performances of the deep CNN approach applied to 110 experimental datasets (i.e. 10 datasets $\times$ 11 training sizes) were measured using the mean average precision (mAP) metric. The creation of the experimental datasets from 150 images and the evaluation metric (mAP@IoU.5) used to assess performance are described in the *Materials and methods* section.

The mAP@IoU.5 ranged from 0.29 (experiment $F_{train10}$) to 0.77 (experiment $A_{train80}$), with a mean of 0.50 and standard deviation of 0.10 (Figure 6a and 6b). The

sensitivity analysis of the number of training images available (Figure 6b) showed that mean mAP@IoU.5 increased from 0.37 to 0.55 as the number of training images increased from 10 to 110, respectively. Mean mAP@IoU.5 varied greatly among datasets (Figure 6c), with the mean mAP@IoU.5 ranging from 0.40 (dataset E) to 0.69 (dataset A).

3.1.2. Detailed analysis of three case studies

Predictions for object detection and segmentation compared to the reference dataset from a per-image analysis are illustrated (Figure 7) for three areas (Area 1, Area 2, Area 3). Models $A_{train110}$ (maximum training size) and $A_{train10}$ (minimum training size) were used as contrasting examples.

3.1.2.1. Area 1: Le Manio, Carnac

Area 1, located at Le Manio (Carnac, France), has three Neolithic burial mounds under a dense canopy composed mainly of coniferous vegetation and brush undergrowth (Figure 7a and 7b). These archaeological structures are identified as Manio 4 (56 034 0113), Manio 5 (56 034 0114) and Manio 8 (56 034 0259) on the national archaeological map.

The low-trained model ($A_{train10}$) and high-trained model ($A_{train110}$) performed well in this area, with 3/3 matches (AP@IoU.5 = 0.92 and 1.0, respectively) (Figure 8). $A_{train10}$ predicted five objects (Figure 7c) that corresponded to three known archaeological structures. However, for the two objects with the lowest IoU values (obj. 3 (0.66) and 5 (0.31)) the predicted BBOXs influenced the predicted

NAME	TYPE	CALCULATION
Area	Morphology	Mask area
Perimeter	Morphology	Mask perimeter
Major axis	Morphology	Orientated mask BBOX major-axis length (m)
Minor axis	Morphology	Orientated mask BBOX minor-axis length (m)
Hyperscale topographic signatures	Context	See the <i>LiDAR-derived visualization image</i> subsection

Table 1 Characterization metrics calculated for the objects detected.

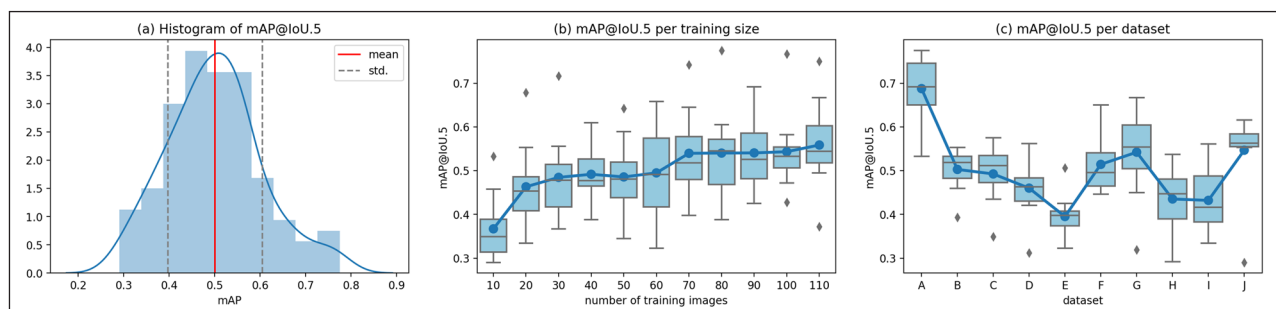


Figure 6 Statistical performances of the 110 models. (a) histogram of mAP values, (b) boxplots of mAP per training size and (c) boxplots of mAP per dataset. Whiskers represent 1.5 times the interquartile range.

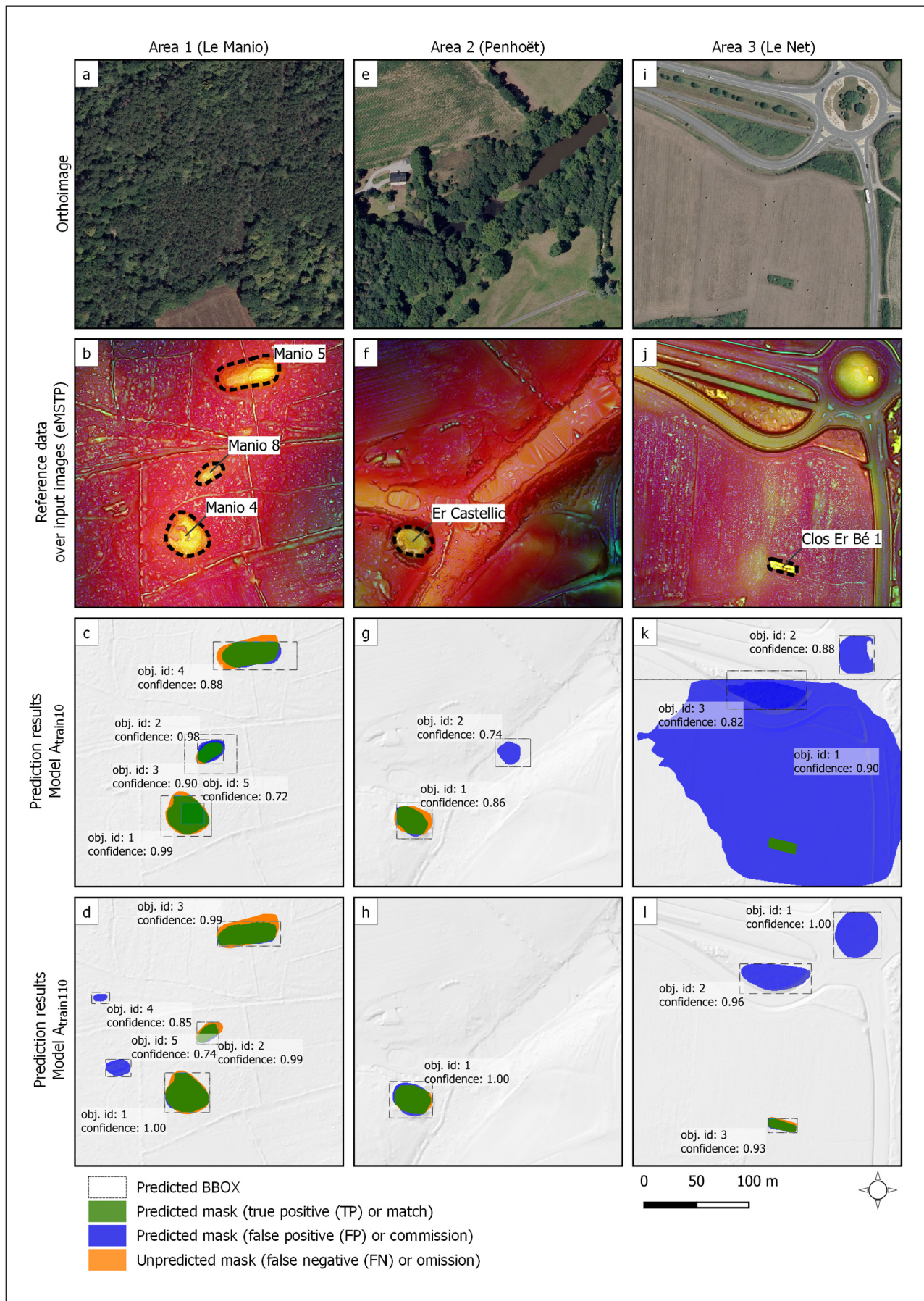


Figure 7 Detected (BBOX) and segmented (mask) objects predicted by $A_{train10}$ and $A_{train110}$ models. Results are shown for Area 1 (Le Manio), Area 2 (Penhoët) and Area 3 (Le Net), France.

.mask. While obj. 3 converged to a correctly adjusted segment by leveraging the segmentation phase within a BBOX larger than the target, obj. 5 resulted in an excessively small segment bounded by an excessively small predicted BBOX. $A_{train110}$ also predicted five objects (Figure 7d); the three with the highest confidence scores corresponded to the three known archaeological structures. The other two objects (obj. 4 and 5), which had lower confidence scores (0.85 and 0.74 respectively), were local topographic anomalies assumed to be due to recent (contemporary period) forestry operations. The quality of the predicted segments was confirmed using available archaeological documentation and in-situ photos (Figure 9).

3.1.2.2. Area 2: Penhoët, Crac'h

Area 2, located at Penhoët (Crac'h, France), has an archaeological structure that is considered to be a motte (Brochard 1994; Cayot-Délandre 1847) that dominates the valley of Le Plessis near the confluence of the Auray River. The archaeological structure, identified as *Er Castellic* (56 046 0015) on the national archaeological map, has never been excavated and it is scarcely documented.

Both the low-trained model ($A_{train10}$) and high-trained model ($A_{train110}$) were able to predict the presence of the archaeological structure ($AP@IoU.5 = 1.0$). $A_{train10}$ predicted two objects (Figure 7g); the BBOX with the highest confidence score (0.86) corresponded to the motte's location. The second BBOX (confidence score 0.74) was a false positive most likely due to an irregularity in the interpolated DTM that was visible on the enhanced multiscale topographic position (eMSTP) image on the surface of a lake.

$A_{train110}$ predicted a single object with a confidence score of 1.00 at the motte's location (Figure 7h). While the predicted mask (770 m²) was slightly larger than the object that had been drawn manually based on the reference data (690 m²), it represented the compact ovoid shape (Figure 10a) of the archaeological structure better. Topographic analysis across the predicted mask identified a visible external ditch and an internal embankment (Figure 10c and 10d).

3.1.2.3. Area 3: Le Net, St Gildas de Rhuys

Area 3, located at Le Net (Saint Gildas de Rhuys, France), has a Neolithic passage grave 21 m long registered as a National Historic Monument since 1923

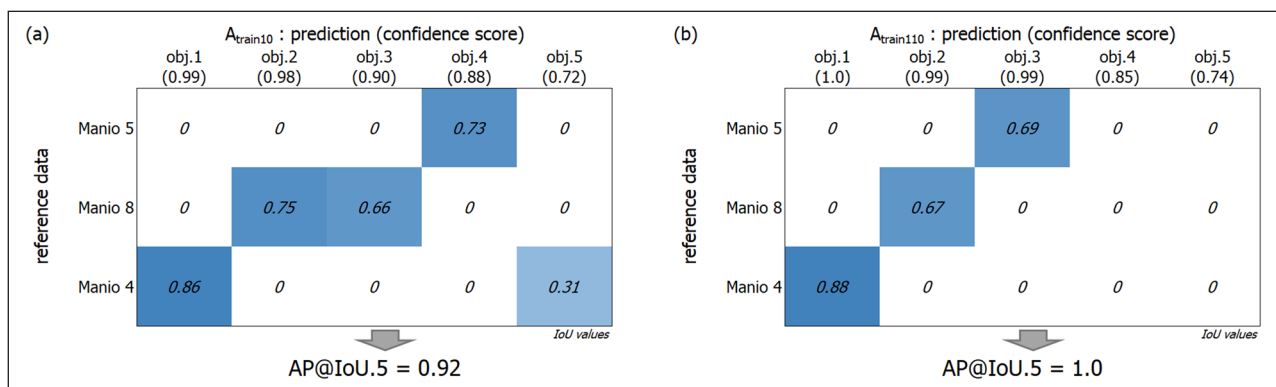


Figure 8 Prediction matrix for (a) $A_{train10}$ and (b) $A_{train110}$ models in Area 1 (Le Manio, France).



Figure 9 Burial mounds at Le Manio (France). (a) Perspective view of the predicted masks (highlighted over an enhanced multiscale topographic position image) and corresponding archaeological structures. (b) Manio 5 viewed from the southwest.

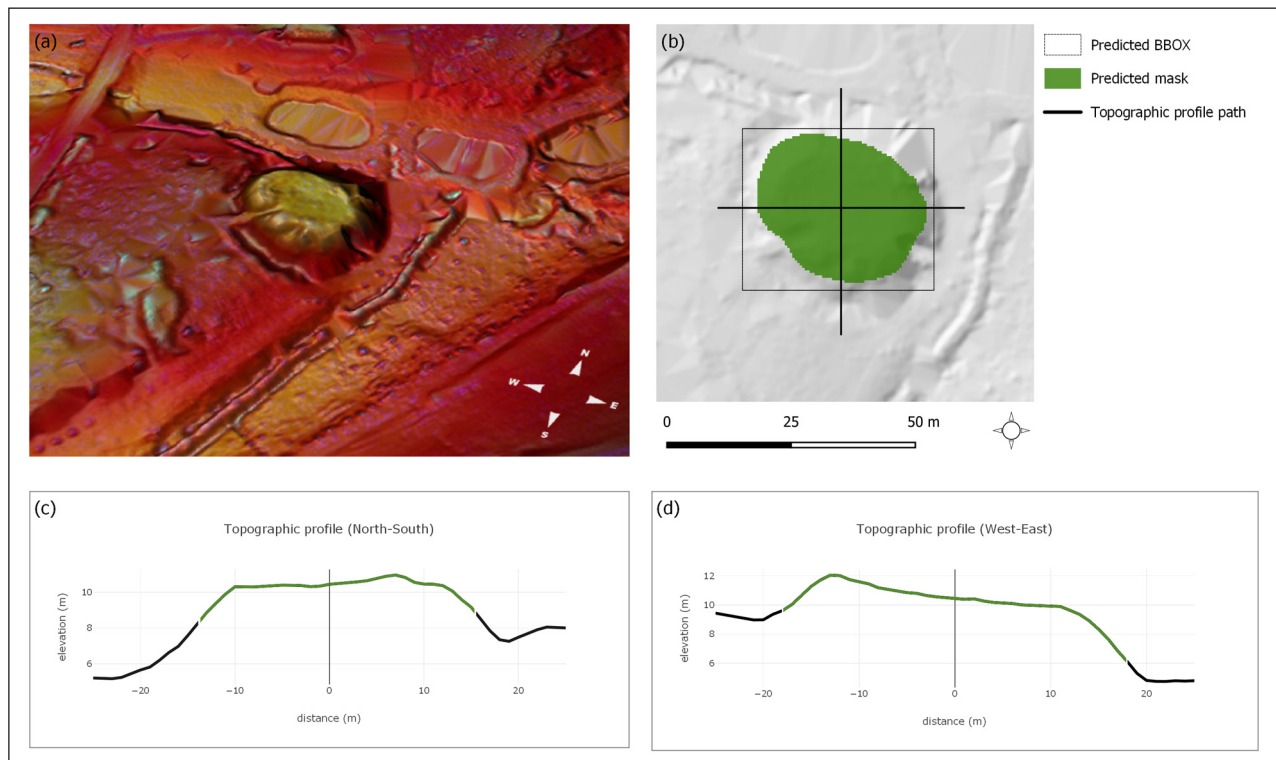


Figure 10 (a) Perspective view of Er Castellic (France) with enhanced multiscale topographic position image overlay, (b) 2D view of the predicted object over the hillshade DTM and (c and d) topographic profiles generated across the predicted object.



Figure 11 (a) Plan and cross-sectional views of the Clos Er Bé passage grave (France) (Le Rouzic, Péquard and Péquard 1922)). (b, c) Current state of the monument covered by bushes and ferns.

(Figure 11a). The site, located in an agricultural field and covered by vegetation and bushes (Figure 11b, 11c), is identified as Clos Er Bé 1 (56 214 0004) on the national archaeological map.

A_{train10} predicted that the monument was contained in one (obj. 1) of the three objects detected (Figure 7k). However, visual analysis revealed that obj. 1 was a large (>3 ha) irregular stain that covered most of the image.

The commission error associated with this single object was 99%.

A_{train110} predicted also three objects (Figure 7l). The passage grave was predicted (obj. 3) with a confidence of 0.93 and an IoU of 0.79, indicating that it corresponded to the footprint of the archaeological structure provided by the reference dataset. The other two objects (obj. 1 and 2), which had higher

confidence scores (1.0 and 0.96, respectively), are perfect examples of false positives. Obj. 1 is a traffic roundabout with a perfectly circular mound landscape design as the central element, while obj. 2 is a recent elongated embankment that protects the bicycle lane. Both objects have topographical and morphological characteristics that resulted in the model making inaccurate predictions.

3.2. OBJECT CHARACTERIZATION: INITIAL RESULTS

As mentioned, the (semi-)automatic process of the deep CNN provided two levels of information: (i) the location of the objects of interest (BBOX, associated with

a confidence score) and (ii) a mask that describes the shape of each predicted object. The latter information was used to characterize the context and morphology of the detected and segmented objects.

This approach was applied to the archaeological site of Park Er Guren (*Figure 12*), which is located east of the Bay of Saint Jean in the commune of Crac'h. The site contains two dolmens separated by 25 m in a north-south orientation that were registered as National Historic Monuments in 1926. The model predicted the presence of two objects (*Figure 13*). Hyperscale topographic position signatures (*Figure 14*) and morphometric descriptors (*Table 2*) were calculated for the masks of both objects.



Figure 12 Dolmens of Park Er Guren (France), view from south. Dolmen A lies in the background, while Dolmen B lies in the foreground.

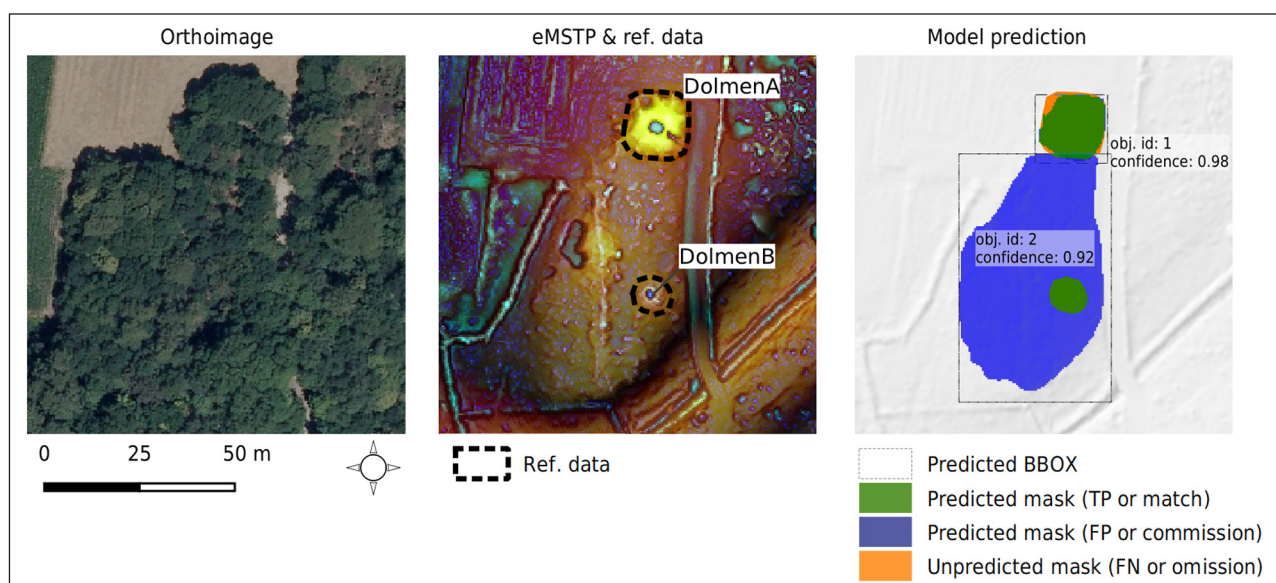


Figure 13 $A_{train110}$ model predictions for the area of Park Er Guren (France). eMSTP = enhanced multiscale topographic position, BBOX = bounding box, TP = true positive, FP = false positive, FN = false negative.

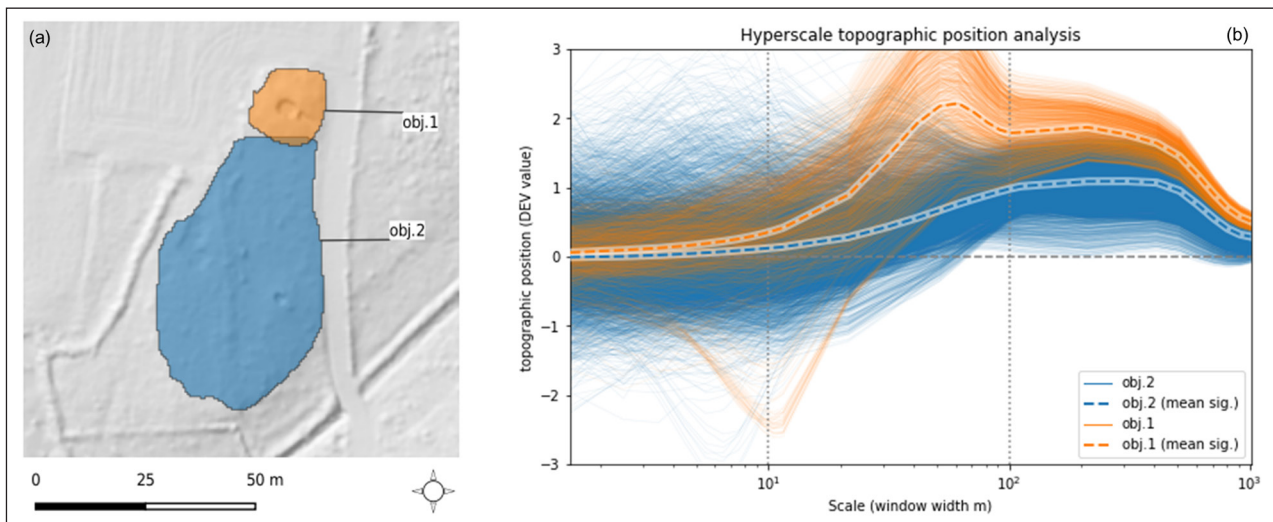


Figure 14 Characterization of the predicted objects for Park Er Guren (France) based on the resulting (semi-)automatic approach. **(a)** Object segmentation results for object 1 (orange) and 2 (blue), **(b)** hyperscale topographic position signatures (each corresponds to a pixel included in the object footprint). Dashed curves indicate the mean signature of each object.

OBJECT	OBJ.1	OBJ.2
Area (m ²)	243	1809
Perimeter (m)	71	203
Major-axis (m)	20	64
Minor-axis (m)	16	37
Hyperscale topographic signature	Figure 14b	Figure 14b

Table 2 Hyperscale topographic position signatures and morphometric descriptors calculated for the predicted objects for Park Er Guren (France).

The hyperscale topographic position signatures and morphometric descriptors were then used to provide a data-driven description of the predicted objects, which was then compared to the archaeological reference dataset and additional archaeological documentation (Gouezin 2017; Le Rouzic 1933) as follows:

- Object 1 was a pseudo-circular element 16–20 m in diameter composed of two main topographical units (groups of signatures). The first unit largely dominated its environment at the mesoscale (10–100 m) and, to a lesser extent, macroscale (100–1000 m). The second unit, with only few pixels, had a negative value of topographic position at the micro-/meso-scale, indicating the presence of a pit or trench. This object corresponded to Dolmen A and described its visible inner structures (e.g. corridor, central position of the chamber). The dolmen's topographically dominant position is characteristic of other Neolithic funeral monuments in the area.
- Object 2 was a large piriform element 64 m long and 37 m wide that varied greatly in topographic positions. Its mean topographic position became progressively

dominant at the meso- and macro-scales, while not being the most dominating element within windows wider than 500 m. A few signatures were highly negative at the microscale, indicating the presence of local depressions within the object's footprint. The complex combination of signatures reflects the multiple topographical units in this piriform mound. The reference data did not describe this complex structure (thus making it statistically a false positive or commission error), but the object suggested an elongated tumulus associated with the dolmens. In addition to the mound, analysis of the hyperscale topographic position signatures suggested topographically visible pits that may correspond to (i) the chamber of Dolmen B and (ii) modern excavation areas visible on the western flank of the mound (**Figure 13**). Locally (micro- and meso-scales), dominating signatures highlighted the presence of the north-south oriented embankment on top of the mound.

3.3. EVALUATION OF THE APPROACH WITHIN AN ARCHAEOLOGICAL PROSPECTION SCHEME

The results of the deployment model showed predicted potential structures with confidence scores ranging from 0.5 to 1. These prediction results highlighted the pixel to object aggregation capability of the deep CNN approach, and predicted object sharing shape and size characteristics with the reference data used to train the model. The predicted objects were visually interpreted on the eMSTP image using two additional study sites using the eMSTP image that were not included for model training, validation or testing.

3.3.1. Analysis on the area of Goah Leron, Carnac (France)

Objects A and B were considered as interesting structures for further field verification. Object A with a circular shape

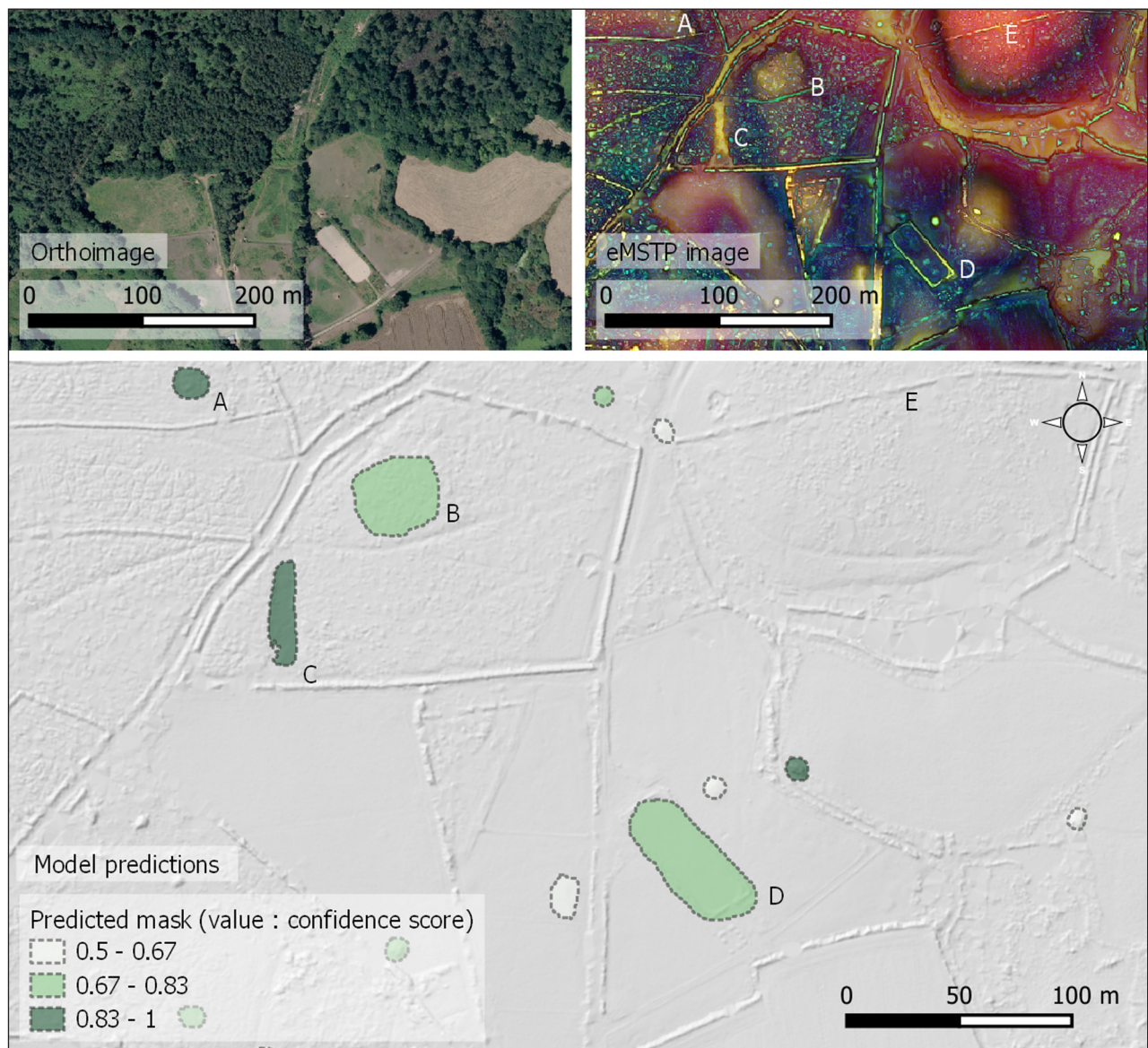


Figure 15 Example of prediction results outside the reference dataset, Goah Léron area, Carnac (France). Objects A and B were considered as interesting structures for further field investigation based on human-interpretation of the eMSTP image. Objects C and D were considered as false positives. Point E highlighted the fact that no potential structure was predicted on the hill. The remaining objects (small isolated mounds) would require further investigation.

(16 m diameter) and low positive elevation (less than 0.3 m above surrounding terrain) showed a rough texture on the eMSTP image, typical of undergrowth vegetation under dense canopy (Figure 15). Object B with a pseudo-circular shape (36 m diameter) and a positive elevation of 0.8 m above surrounding terrain, shared the same eMSTP characteristics. It is to be noted that the presence of standing stones (not visible on the LiDAR-derived DTM) is attested between object A & object B, thus supporting the idea of the possible presence of Neolithic burial mounds nearby.

Object C was considered as a false-positive. This object corresponded to a north-south orientated terrain depression of 12 m wide, 46 m long and 40 cm deep that shared similarities with the representation of some elongated tumulus in the eMSTP image. This was mostly

due to the conversion of the topographic metric DEVs from relative to absolute values during the calculation of the eMSTP image.

Object D was also considered as a false positive. This object, which corresponded to a horse training arena with flat elevation and surrounding embankments, shared shape characteristics with reference data, but not topographic or texture characteristics.

The model did not predict any potential structure on the hill located North-East of the area (point E). While the yellow-reddish color in the eMSTP image -associated to the meso-macro dominating topographic signature-corresponded to the specific position of many tumulus in the study area, the model did not predict any object, which was probably due to the absence of local morphological anomalies.

The remaining predicted objects were isolated small mounds (4 to 6 m in diameter) less than 1 m high, most of them being located in open agricultural areas. While it was not possible to determine their nature only based on the interpretation of the eMSTP image further investigation would be required to identify them.

3.3.2. Analysis on the area of Brahen, Carnac (France)

Objects A and B were identified as archaeological entities. Object A was a circular mound (26 m diameter) with positive elevation of 0.8 m above the surrounding terrain (*Figure 16*). The field verification confirmed the probable archaeological nature of this structure as a tumulus, with a possible attribution to the Bronze

Age based on its morphology. Object C corresponded to a dominating terrain covered by dense vegetation with a morphological anomaly on its highest position (Object B). In the field, remaining elements of a possible megalithic stone alignment were identified at this position.

Object D was considered as a false-positive. This object corresponded to a narrow ditch with east-west orientation that shared similarities with the representation of some elongated tumulus in the eMSTP image. This detection error could be due to the conversion of the topographic metric DEVs from relative to absolute values during the calculation of the eMSTP image.

The remaining predicted objects corresponded to local morphological anomalies that would require further investigation.

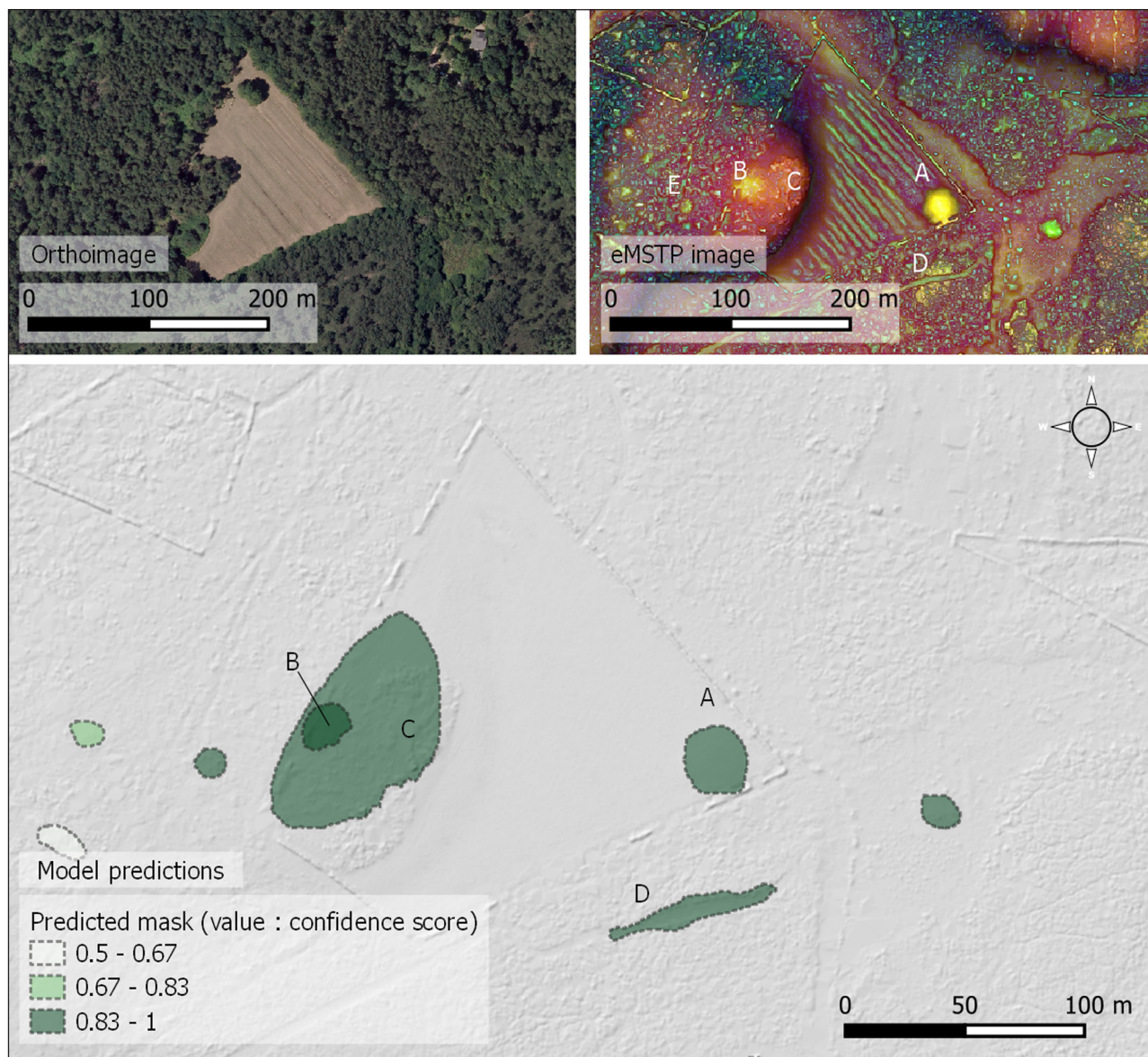


Figure 16 Example of prediction results outside the reference dataset, Brahen area, Carnac (France). Objects A and B corresponded to archaeological structures confirmed by field verification, object C being a dominating terrain including object B. Object D was considered as false-positive. Remaining objects were local morphological anomalies that would require further investigation.

4. DISCUSSION

4.1. SENSITIVITY OF THE APPROACH AND GENERALIZATION ABILITY

The deep CNN approach resulted in high detection and segmentation performances (mAP up to 0.77) with relatively small training datasets. The largest training dataset contained 110 images, which is small training set for deep learning. This confirms the approach's ability to perform well in archaeological contexts in which sparse reference data are a common limitation.

Nonetheless, the model's sensitivity to the images selected for the training and test datasets (with mAP@IoU.5 varying from 0.29 (model E_{train110}) to 0.77 (model A_{train110}) for the same number of training images) raises some concerns. A previous study that focused on (semi-) automatic archaeological mapping (Verschoof-van der Vaart et al. 2020) also mentioned this sensitivity. 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).

The deep CNN approach showed adaptability in detecting and segmenting different archaeological structures within the region. However, model training and evaluation were limited to a region that has particular topographic and archaeological characteristics. Most of the archaeological structures contained in the reference dataset have a topographically dominant position (burial mounds, hillforts, wind mills), but their local dominance is highly variable in magnitude and scale. While the trained models detected most above mean elevations (e.g. roundabout), they differed from local maximum detectors on their ability to consider the following archaeological landscape characteristics: the multiscale topographic position of the sites (maxima at specific local neighborhood or scale) and the local morphological patterns of archaeological structures. As confirmed by the results obtained using the deployment model applied on an independent set of images of the study area, these characteristics were learned during the training phase and used for prediction. This demonstrated the generalization capabilities of the approach in the geo-archaeological context of the study area.

The limits of the deep CNN approach were also identified. Beside prediction errors that were expected (e.g. roundabout), errors were also observed for objects sharing few or no similarities with the reference dataset (e.g. horse training area, large ditch). Such undesired behavior of the deep CNN models raised the question of negative training (i.e. providing the model with negative examples during training). While this was not implemented in the Mask R-CNN framework used in this study, it should be addressed in future works to improve prediction performances, for example using software

frameworks that handle negative training for instance segmentation, such as Detectron2 (Wu et al. 2019).

More generally, results showed that a particular attention should be paid to the selection of training examples. The sample selection strategy is still a challenging concern especially with the hidden and non-intuitive phenomena related to deep CNN. Tools that facilitate insights into model successes and failures such as Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2020) could be used to tackle such concern.

Further investigation of the multiple hyper-parameters and model configurations of deep CNN architectures would be helpful to assess the scope and limits of the approach. As an example, data augmentation (DA) was empirically used to improve model performances and generalization capabilities (Shorten and Khoshgoftaar 2019). The evaluation of DA was not included in this study, because a comprehensive assessment would involve a full-fledged study (evaluation of performances with and without DA, and with multiple DA configurations involving various combinations of DA techniques). Although we did not perform this comprehensive evaluation, we evaluated DA effect on a single model (A_{train110}) trained without and with data-augmentation using a performance test. Results showed an increase of the mAP@IoU.5 performance from 0.64 to 0.75.

Assessing the overall generalization ability at a larger geographical scale (spatial generalization) and for more types of archaeological structures (typological generalization) would require further experiments. First, to assess spatial generalization, a pre-trained model could be used to identify topographical anomalies that have characteristics similar to those on the coast of Morbihan using the LiDAR dataset of relevant regions in the world. Second, to assess typological generalization, the model could be retrained to include new types of structures to increase the diversity of archaeological contexts assimilated by the deep CNN. These strategies would benefit from public benchmark dataset targeted to detect archaeological sites from remotely sensed data.

4.2. EVALUATION METRICS FOR AMBIGUOUS REFERENCE DATA

The results indicate that statistical assessment of the models provided an objective metric of the quality of predictions, but it did not completely capture the approach's performance because 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 were based on an overlap measurement (i.e., IoU) that was 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.

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. This study's examples of false-positive detections ([Figure 7d](#) and [7l](#)) are representative of this intentionally liberal strategy, with topographical structures detected (i) correctly because they share characteristics with known archaeological structures and (ii) incorrectly because they are ultimately interpreted as contemporary human earthworks that are not considered of archaeological importance. 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.

4.3. ONE-CLASS APPROACH AND POST-DETECTION CHARACTERIZATION: POTENTIAL FOR A NEW PARADIGM FOR (SEMI-) AUTOMATIC MAPPING IN ARCHAEOLOGY

Most approaches in machine learning-based archaeological mapping use a pre-defined nomenclature (e.g. barrows, charcoal kilns, Celtic fields, burial mounds, mining pits) to consider local archaeological characteristics (e.g. site morphology, chrono-typological relation, spatial relationship). However, a standard and consensual typology appropriate for remotely sensed archaeological structures that span time and space remains a concern (Tarolli et al. 2019). Moreover, classes are often distributed unequally (i.e. datasets of archaeological structures with a lack of samples for certain classes).

We used a one-class rather than multi-class approach to address these two issues because we assumed that the deep CNN would have higher generalization abilities (i.e. depend less on target type and variety) with a one-class approach. This was confirmed by the results obtained for the *Er Castell* motte, whose

structure type was not included in the training dataset. Although this artificially elevated terrain monument was the only example of its type in the study area, it was sufficiently similar to a tumulus for the model to detect it as an object of interest. These topographical and morphological similarities with certain tumulus were mentioned in an archaeological prospection report (Brochard 1994) and reinforced our assumption. Indeed, from a LiDAR perspective, archaeological sites of different chronologies and typologies share patterns that the deep CNN can discover and extract.

The characterization phase, based on the object-segmented mask and data-driven description, provides information that can help to identify the nature of the archaeological structures. For example, characterization of the detected objects and segmented at the *Park Er Guren* site made it possible to identify a tumulus and related dolmens. Although more examples are required to confirm this assumption, this approach provides new perspectives by inverting the common conceptual model in remote-sensing archaeological mapping in which a typology of target options must be defined before (semi-) automatic detection.

5. CONCLUSION AND OUTLOOK

We demonstrated potential methods that can detect and characterize archaeological structures by performing object segmentation using a deep CNN approach combined with transfer learning. Our study reveals that the approach developed can be used to (semi-)automatically detect, delineate and characterize topographic anomalies. The results, compared to archaeological reference data collected from archaeological documentation, showed detection accuracy (mAP@IoU.5) up to 0.77 and provided new perspectives for archaeological documentation and interpretation through morphometric and contextual characterization via object segmentation. The one-class detection method combined with a characterization-interpretation strategy provides a new paradigm for prospecting archaeological structures in varying states of conservation or with conflicting typologies. The application of such a deep CNN approach to large scale archaeological mapping in wider geographical and archaeological contexts still needs to be extended and assessed. Beside the necessary addition of a new set of reference data covering various geo-archaeological situations, this would also involve the development of methods for the optimal selection of training samples. It would also involve further investigation on the effectiveness of the LiDAR-derived VT as input to the automatic detection and segmentation processes. In this regards, the objective evaluation metrics provided by the deep CNN approach could be used for the benchmarking of new and existing VTs.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendix A.1.** Experimental configurations for 10 datasets randomly drawn (A–J), with 11 training datasets (10–110), 1 validation (val.) dataset and 1 test dataset per draw. DOI: <https://doi.org/10.5334/jcaa.64.s1>

ACKNOWLEDGEMENTS

The authors thank members of DRAC Bretagne, Service régional de l'archéologie for providing access to the LiDAR raw data, the archaeological documentation and reference dataset as well as for the field verifications. OSUR/OSUNA and GeoFIT are also acknowledged for the LiDAR acquisition. The Nantes Rennes LiDAR platform was funded by the Region Bretagne and the Region Pays de la Loire with European Regional Development Fund (ERDF).

FUNDING INFORMATION

This research was co-funded by the Région Bretagne (project “Patrimoine – Mégalithes de Bretagne”) and DRAC Bretagne, Service régional de l'archéologie and Hytech-Imaging.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Alexandre Guyot  orcid.org/0000-0002-8415-9240
Laboratoire LETG – UMR 6554, Université Rennes 2, FR

Marc Lennon
Hytech-Imaging, FR

Thierry Lorho
DRAC Centre-Val de Loire, Service régional de l'archéologie/UMR 6566 CReAAH, FR

Laurence Hubert-Moy
Laboratoire LETG – UMR 6554, Université Rennes 2, FR

REFERENCES

- Abdulla, W.** 2017. *matterport/Mask_RCNN: Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow*. 2017. Available at https://github.com/matterport/Mask_RCNN [Last accessed 5 November 2020].
- Bennett, R, Welham, K, Hill, RA and Ford, A.** 2012. A Comparison of Visualization Techniques for Models Created from Airborne Laser Scanned Data: A Comparison of Visualization Techniques for ALS Data. *Archaeological Prospection*, 19(1): 41–48. DOI: <https://doi.org/10.1002/arp.1414>
- Bewley, RH, Crutchley, SP and Shell, CA.** 2005. New light on an ancient landscape: lidar survey in the Stonehenge World Heritage Site. *Antiquity*, 79(305): 636–647. DOI: <https://doi.org/10.1017/S0003598X00114577>
- Brochard, E.** 1994. *Les Fortifications de terre du Comté de Vannes*. Rennes, France: Drac SRA Bretagne.
- Caspari, G and Crespo, P.** 2019. Convolutional neural networks for archaeological site detection – Finding “princely” tombs. *Journal of Archaeological Science*, 110: 104998. DOI: <https://doi.org/10.1016/j.jas.2019.104998>
- Cayot-Délandre, M.** 1847. *Le Morbihan, son histoire et ses monuments*. Vannes: éditions Cauderan.
- Chase, AF, Chase, DZ, Weishampel, JF, Drake, JB, Shrestha, RL, Slatton, KC, Awe, JJ and Carter, WE.** 2011. Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol, Belize. *Journal of Archaeological Science*, 38(2): 387–398. DOI: <https://doi.org/10.1016/j.jas.2010.09.018>
- Chiba, T, Kaneta, S and Suzuki, Y.** 2008. Red relief image map: new visualization method for three dimensional data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B2): 1071–1076.
- Devereux, BJ, Amable, GS and Crow, P.** 2008. Visualisation of LiDAR terrain models for archaeological feature detection. *Antiquity*, 82(316): 470–479. DOI: <https://doi.org/10.1017/S0003598X00096952>
- Devereux, BJ, Amable, GS, Crow, P and Cliff, AD.** 2005. The potential of airborne lidar for detection of archaeological features under woodland canopies. *Antiquity*, 79(305): 648–660. DOI: <https://doi.org/10.1017/S0003598X00114589>
- Doneus, M.** 2013. Openness as Visualization Technique for Interpretative Mapping of Airborne Lidar Derived Digital Terrain Models. *Remote Sensing*, 5(12): 6427–6442. DOI: <https://doi.org/10.3390/rs5126427>
- Evans, D.** 2016 Airborne laser scanning as a method for exploring long-term socio-ecological dynamics in Cambodia. *Journal of Archaeological Science*, 74: 164–175. DOI: <https://doi.org/10.1016/j.jas.2016.05.009>
- Evans, DH, Fletcher, RJ, Pottier, C, Chevance, J-B, Soutif, D, Tan, BS, Im, S, Ea, D, Tin, T, Kim, S, et al.** 2013. Uncovering archaeological landscapes at Angkor using lidar. *Proceedings of the National Academy of Sciences*, 110(31): 12595–12600. DOI: <https://doi.org/10.1073/pnas.1306539110>
- Freeland, T, Heung, B, Burley, DV, Clark, G and Knudby, A.** 2016. Automated feature extraction for prospection and analysis of monumental earthworks from aerial LiDAR in the Kingdom of Tonga. *Journal of Archaeological Science*, 69: 64–74. DOI: <https://doi.org/10.1016/j.jas.2016.04.011>
- Gallwey, J, Eyre, M, Tonkins, M and Coggan, J.** 2019. Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial

- Heritage through Deep Transfer Learning. *Remote Sensing*, 11(17): 1994. DOI: <https://doi.org/10.3390/rs11171994>
- Girshick, R.** 2015. Fast r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. 2015 Santiago, Chile: IEEE. pp. 1440–1448. DOI: <https://doi.org/10.1109/ICCV.2015.169>
- Gouezin, P.** 2017. *Structures funéraires et pierres dressées. Analyses architecturales et spatiales. Mégalithes du département du Morbihan*. Unpublished thesis (PhD). Université Rennes 1.
- Guyot, A, Hubert-Moy, L and Lorho, T.** 2018. Detecting Neolithic Burial Mounds from LiDAR-Derived Elevation Data Using a Multi-Scale Approach and Machine Learning Techniques. *Remote Sensing*, 10(2): 225. DOI: <https://doi.org/10.3390/rs10020225>
- He, K, Gkioxari, G, Dollár, P and Girshick, R.** 2017. Mask R-CNN, arXiv:1703.06870. In: *Proceedings of the IEEE International Conference on Computer Vision*, 22–29 October 2017. Venice: IEEE. pp. 2961–2969. DOI: <https://doi.org/10.1109/ICCV.2017.322>
- Hesse, R.** 2010. LiDAR-derived Local Relief Models – a new tool for archaeological prospection. *Archaeological Prospection*, 17(2): 67–72. DOI: <https://doi.org/10.1002/arp.374>
- Inomata, T, Triadan, D, Vázquez López, VA, Fernandez-Díaz, JC, Omori, T, Méndez Bauer, MB, García Hernández, M, Beach, T, Cagnato, C, Aoyama, K and Nasu, H.** 2020. Monumental architecture at Aguada Fénix and the rise of Maya civilization. *Nature*, 582(7813): 530–533. DOI: <https://doi.org/10.1038/s41586-020-2343-4>
- Jung, AB, Wada, K, Crall, J, Tanaka, S, Graving, J, Reinders, C, Yadav, S, Banerjee, J, Vecsei, G, Kraft, A, Rui, Z, Borovec, J, Vallentin, C, Zhydenko, S, Pfeiffer, K, Cook, B, Fernández, I, De Rainville, F-M, Weng, C-H, Ayala-Acevedo, A, Meudec, R, Laporte, M, et al.** 2020. *imgaug*. 2020. Available at <https://github.com/aleju/imgaug> [Last accessed 5 November 2020].
- Kazimi, B, Thiemann, F, Malek, K, Sester, M and Khoshelham, K.** 2018. Deep Learning for Archaeological Object Detection in Airborne Laser Scanning Data. In: *Proceedings of the 2nd Workshop On Computing Techniques For Spatio-Temporal Data in Archaeology And Cultural Heritage*. 2018 Melbourne, Australia: CEUR Workshop Proceedings. p. 15. DOI: <https://doi.org/10.4230/LIPICs.COARCH.2018>
- Kazimi, B, Thiemann, F and Sester, M.** 2019. Object Instance Segmentation in Digital Terrain Models. In: Vento, M and Percannella, G (eds.). *Proceedings of the 18th International Conference on Computer Analysis of Images and Patterns*, 3–5 September 2019. Cham: Springer International Publishing. pp. 488–495. DOI: https://doi.org/10.1007/978-3-030-29891-3_43
- Kazimi, B, Thiemann, F and Sester, M.** 2020. Detection of Terrain Structures in Airborne Laser Scanning Data Using Deep Learning. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-2–2020: 493–500. DOI: <https://doi.org/10.5194/isprs-annals-V-2-2020-493-2020>
- Kokalj, Ž and Hesse, R.** 2017. *Airborne laser scanning raster data visualization a guide to good practice*. Ljubljana: Založba ZRC. DOI: <https://doi.org/10.3986/9789612549848>
- Krizhevsky, A, Sutskever, I and Hinton, GE.** 2012. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira, F, Burges, CJC, Bottou, L and Weinberger, KQ (eds.), *Advances in Neural Information Processing Systems* 25. Curran Associates, Inc. pp. 1097–1105.
- Küçükdemirci, M and Sarris, A.** 2020. Deep learning based automated analysis of archaeogeophysical images, *Archaeological Prospection*, 27: 107–118. DOI: <https://doi.org/10.1002/arp.1763>
- Lambers, K, Verschoof-van der Vaart, W and Bourgeois, Q.** 2019. Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection. *Remote Sensing*, 11(7): 794. DOI: <https://doi.org/10.3390/rs11070794>
- Le Rouzic, Z.** 1933. Morphologie et chronologie des sépultures préhistoriques du Morbihan. *L'Anthropologie*, 3–4: 225–265.
- Le Rouzic, Z, Péquard, S-J and Péquard, M.** 1922. Carnac (Morbihan), Fouilles faites dans la région. Allée couverte du Net, dite « Er-Bé (la tombe), Commune de Saint-Gildas de Rhuys. *Revue Anthropologique*, 183–189.
- LeCun, Y, Bengio, Y and Hinton, G.** 2015. Deep learning. *Nature*, 521(7553): 436–444. DOI: <https://doi.org/10.1038/nature14539>
- Lin, T-Y, Maire, M, Belongie, S, Bourdev, L, Girshick, R, Hays, J, Perona, P, Ramanan, D, Zitnick, CL and Dollár, P.** 2014. Microsoft COCO: Common Objects in Context. arXiv:1405.0312 [cs]. In: Fleet, D, Pajdla, T, Schiele, B and Tuytelaars, T (eds.), *Computer Vision – ECCV 2014. Lecture Notes in Computer Science*, 2014. Cham: Springer International Publishing. pp. 740–755. DOI: https://doi.org/10.1007/978-3-319-10602-1_48
- Lindsay, JB, Cockburn, JMH and Russell, HAJ.** 2015. An integral image approach to performing multi-scale topographic position analysis. *Geomorphology*, 245: 51–61. DOI: <https://doi.org/10.1016/j.geomorph.2015.05.025>
- Niculiță, M.** 2020. Geomorphometric Methods for Burial Mound Recognition and Extraction from High-Resolution LiDAR DEMs. *Sensors*, 20(4): 1192. DOI: <https://doi.org/10.3390/s20041192>
- Opitz, R and Herrmann, J.** 2018. Recent Trends and Long-standing Problems in Archaeological Remote Sensing. *Journal of Computer Applications in Archaeology*, 1(1): 19–41. DOI: <https://doi.org/10.5334/jcaa.11>
- Sadr, K.** 2019. Kweneng: A Newly Discovered Pre-Colonial Capital Near Johannesburg. *Journal of African Archaeology*, 17(1): 1–22. DOI: <https://doi.org/10.1163/21915784-20190001>
- Selvaraju, RR, Cogswell, M, Das, A, Vedantam, R, Parikh, D and Batra, D.** 2020. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, *International Journal of Computer Vision*, 128(2): 336–359. DOI: <https://doi.org/10.1007/s11263-019-01228-7>

- Shorten, C** and **Khoshgoftaar, TM**. 2019. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1): 60. DOI: <https://doi.org/10.1186/s40537-019-0197-0>
- Sorosh, M, Mehrtash, A, Khazraee, E** and **Ur, JA**. 2020. Deep Learning in Archaeological Remote Sensing: Automated Qanat Detection in the Kurdistan Region of Iraq. *Remote Sensing*, 12(3): 500. DOI: <https://doi.org/10.3390/rs12030500>
- Štular, B, Kokalj, Ž, Oštir, K** and **Nuninger, L**. 2012. Visualization of lidar-derived relief models for detection of archaeological features. *Journal of Archaeological Science*, 39(11): 3354–3360. DOI: <https://doi.org/10.1016/j.jas.2012.05.029>
- Tarolli, P, Cao, W, Sofia, G, Evans, D** and **Ellis, EC**. 2019. From features to fingerprints: A general diagnostic framework for anthropogenic geomorphology. *Progress in Physical Geography: Earth and Environment*, 43(1): 95–128. DOI: <https://doi.org/10.1177/0309133318825284>
- Trier, ØD, Cowley, DC** and **Waldeland, AU**. 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*. DOI: <https://doi.org/10.1002/arp.1731>
- Trier, ØD, Salberg, AB** and **Holger Pilø, L**. 2016. Semi-automatic mapping of charcoal kilns from airborne laser scanning data using deep learning. In: Matsumoto, M and Uleberg, E (eds.). *CAA2016: Oceans of Data. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology*, 30 March–3 April 2016. Oxford: Archeopress. pp. 221–232.
- Trier, ØD** and **Pilø, LH**. 2012. Automatic Detection of Pit Structures in Airborne Laser Scanning Data: Automatic detection of pits in ALS data. *Archaeological Prospection*, 19(2): 103–121. DOI: <https://doi.org/10.1002/arp.1421>
- Verschoof-van der Vaart, WB** and **Lambers, K**. 2019. Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. *Journal of Computer Applications in Archaeology*, 2(1): 31–40. DOI: <https://doi.org/10.5334/jcaa.32>
- Verschoof-van der Vaart, WB, Lambers, K, Kowalczyk, W** and **Bourgeois, QPJ**. 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. DOI: <https://doi.org/10.3390/ijgi9050293>
- Weiss, K, Khoshgoftaar, TM** and **Wang, D**. 2016. A survey of transfer learning. *Journal of Big Data*, 3(1): 9. DOI: <https://doi.org/10.1186/s40537-016-0043-6>
- Wilson, JP** and **Gallant, JC**. 2000. *Terrain analysis: principles and applications*. New York: Wiley.
- Wu, Y, Kirillov, A, Massa, F, Lo, W-Y** and **Girshick, R**. 2019. *Detectron2*. 2019. Available at <https://github.com/facebookresearch/detectron2> [Last accessed 5 November 2020].

TO CITE THIS ARTICLE:

Guyot, A, Lennon, M, Lorho, T and Hubert-Moy, L. 2021. Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach. *Journal of Computer Applications in Archaeology*, 4(1), pp. 1–19. DOI: <https://doi.org/10.5334/jcaa.64>

Submitted: 14 September 2020 Accepted: 15 January 2021 Published: 03 February 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Journal of Computer Applications in Archaeology is a peer-reviewed open access journal published by Ubiquity Press.