PAPER • OPEN ACCESS

Towards the use of artificial intelligence deep learning networks for detection of archaeological sites

View the article online for updates and enhancements.

Surface Topography: Metrology and Properties

PAPER

Towards the use of artificial intelligence deep learning networks for detection of archaeological sites

Alexandra Karamitrou[1,*] , Fraser Sturt[1] , Petros Bogiatzis[2] and David Beresford-Jones[3]

1   University of Southampton, Department of Archaeology, United Kingdom
2   Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, United Kingdom
3   University of Cambridge, Department of Archaeology, United Kingdom
*   Author to whom any correspondence should be addressed.

E-mail: a.karamitrou@soton.ac.uk

Abstract

While remote sensing data have long been widely used in archaeological prospection over large areas, the task of examining such data is time consuming and requires experienced and specialist analysts. However, recent technological advances in the field of artificial intelligence (AI), and in particular deep learning methods, open possibilities for the automated analysis of large areas of remote sensing data. This paper examines the applicability and potential of supervised deep learning methods for the detection and mapping of different kinds of archaeological sites comprising features such as walls and linear or curvilinear structures of different dimensions, spectral and geometrical properties. Our work deliberately uses open-source imagery to demonstrate the accessibility of these tools. One of the main challenges facing AI approaches has been that they require large amounts of labeled data to achieve high levels of accuracy so that the training stage requires significant computational resources. Our results show, however, that even with relatively limited amounts of data, simple eight-layer, fully convolutional network can be trained efficiently using minimal computational resources, to identify and classify archaeological sites and successfully distinguish them from features with similar characteristics. By increasing the number of training sets and switching to the use of high-performance computing the accuracy of the identified areas increases. We conclude by discussing the future directions and potential of such methods in archaeological research.

1. Introduction

Analysis of aerial imagery revolutionized archaeology in the early twentieth century, exponentially increasing the number of known sites, allowing large areas to be rapidly surveyed and giving access to remote regions (Reeves 1936, Bewley and Raczkowski 2002; Mossun *et al* 2013; Lambers 2018). For example, a search for scientific publications related with Archaeology and Remote Sensing using the Dimensions scientific research database returns 2,732 articles on 2013, 5,172 on 2018 and 14,323 in 2021 (https://app.dimensions.ai; accessed in May 2022).

With the introduction of a wider range of airborne (i.e., manned aircraft and drones) and space-based data, including passive high spatial resolution optical sensors, multispectral and hyperspectral sensors, light detection and ranging (LIDAR), Synthetic aperture radar (SAR), thermal sensors and geophysical images, the amount of data available to archaeologists has also increased exponentially in recent years (e.g., Chi *et al* 2016; Tamiminia *et al* 2020). These data hold significant potential to transform our understanding of the archaeological record, but also pose a significant challenge with regards to the amount of time analysis would take using traditional human-led image analysis methods.

Artificial Intelligence (AI) offers a potential bypass to this bottle neck and therefore substantially reduce the required human labor. AI describes the ability of computers to perform tasks and reaching decisions through learning either directly from the data (unsupervised methods) or from past experience where the correct outcome is provided (supervised methods),

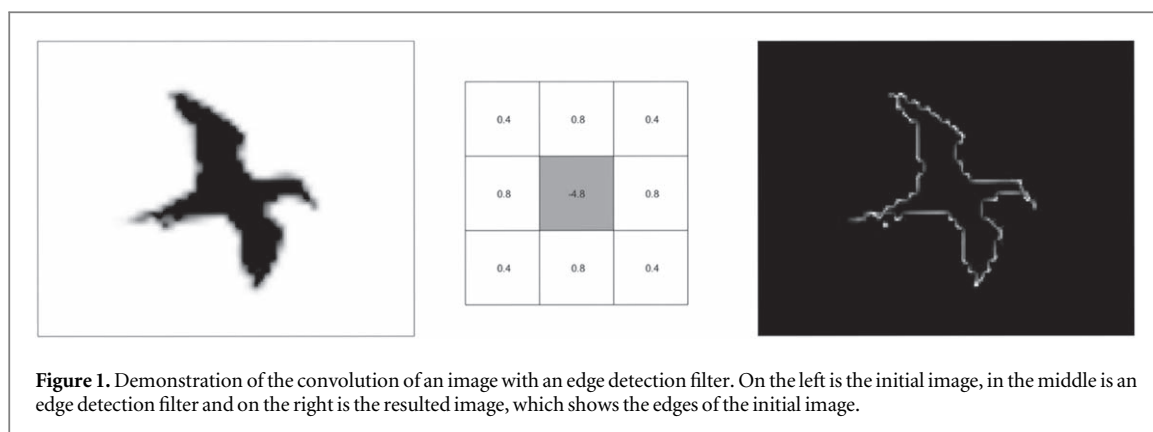imitating human intelligence (e.g., Dey, 2016; Copeland 2020).

Over the past three decades, applications of machine learning (ML) methods have seen significant increase in Archaeology. ML algorithms such as support vector machine (Cortes and Vapnik 1995; Kao *et al* 2004) random forests (Ho 1995; Ho 1998), K-means (Cao *et al* 2009; Jin and Han, 2011; Qi *et al* 2017) and other similar approaches have been widely adopted with considerable success in detecting or classifying archaeological sites, and artifacts (e.g., Kintigh and Ammerman 1982; Baxter 2009; Menze and Ur 2012; Flores *et al* 2019; Orengo *et al* 2020). These methods, often referred to as traditional ML algorithms, require the careful selection of input features (e.g., various spectral indices in satellite imaging) by human-experts, that are important for the outcome. Then through an iterative optimization process by the input of exemplar data the algorithm is trained based upon multivariate statistics and progressively improves its performance. Since it requires the determination and the prior calculation of a range of possible statistically significant input features, it inevitably suffers from a level of bias as although the training procedure can point out which from the features are statistically insignificant, it cannot suggest, or extract features different than the provided ones. Also, the relatively limited number of the features in most applications often cannot fully describe the targets at different situations or environmental conditions. Therefore, the applicability of these algorithms is often limited to specific cases and restricts the identification to features with limited spectral and geometric variations.

In the early 2000s a new machine learning technology emerged known as Deep Learning (DL) based on Artificial Neural Networks (ANN), and in the case of image applications, Convolutional Neural Networks (CNNs). This new technology was largely based on the seminal work of Fukushima (1980) as well as Hubel and Wiesel (1959) that introduced the 'neocognitron' (Fukushima 1980; 1983; 2003) and established the use of convolutional and down-sampling layers. In 1986, Rina Decher was one of the first to use the term 'deep learning' to the machine learning community, in which 'deep' was used to describe the use of multiple layers in a network. Later, Waibel (1987) proposed the time delay neural network (TDNN), one of the first convolutional networks followed by LeCun *et al* (1989) who applied that in a handwritten character recognition problem using a 7-level Convolutional Neural Netowork (CNN), called LeNet-5 (LeCun *et al* 1998). A significant advantage of deep learning methods is that the feature extraction and selection stage is performed by the learning algorithm automatically and not by a person. Yet, this usually requires significant amounts of labeled data and considerable computational resources for the training process. The utilization of GPUs in the training process was the

turning point for using CNNs in image recognition. In the 2012 ImageNet competition, the first CNN ever submitted, named AlexNet (Krizhevsky *et al* 2012), won the competition. The training of AlexNet used over one million labeled images about ~1000 object categories and took ~6 days using 2 GPUs (Krizhevsky *et al* 2012). Since then, deep neural networks have won many international pattern recognition competitions and have attracted broad attention, by outperforming legacy machine learning methods and handling better large amounts of data with minimum user intervention (Schmidhuber 2015). As such, they offer considerable potential for archaeology.

Among the common tasks assigned to deep learning CNN networks are image classification, object detection, and semantic segmentation. Classification is a basic process routinely performed in archaeology with the objective of classifying groups of images that share some common features, or objects into one of a number of predefined classes. For example, AI methods have been used to analyze use-wear on lithic tools (e.g., Van den Dries 1998) and to classify and identify types of pottery (e.g., Hörr *et al* 2008; Anichini *et al*, 2021; Pawlowicz and Downum 2021). Caspari and Crespo (2019), used an object-detection based method to identify Iron Age burial mounds in aerial imagery. More recently, Agapiou *et al* (2021) applied the object detection method to detect surface ceramics in drone images. Finally, semantic segmentation algorithms attempt to analyze images further, by partitioning them into semantically meaningful parts and afterwards by classifying each part into one of the 'X' predetermined classes i.e., interpretable image regions for instance, archaeological sites, regions of vegetation, modern structures and others (e.g., Garcia-Garcia *et al* 2018; Minaee *et al* 2020). Semantic segmentation operates at pixel-level in the sense that each pixel of an image is labeled according to the class it belongs to. This makes semantic segmentation a much more complicated and computationally intensive task, yet it can produce more informative and detailed results compared to classification and object identification (e.g., Kendall *et al* 2015; Garcia-Garcia *et al* 2018; Minaee *et al* 2020). The value of this approach for geophysical analysis has been demonstrated in the work of Küçükdemirci and Sarris's (2020) using ground-penetrating radar images.

For all this success, only recently there have been limited yet increasing work adopting CNN approaches for the automated detection of archaeological sites (Trier *et al* 2018; Caspari and Crespo, 2019; Kazimi *et al* 2019; Lambers *et al* 2019; Rayne *et al* 2020; Somrak *et al* 2020; Soroush *et al* 2020; Bonhage *et al* 2021; Verschoof-van der Vaart and Landauer 2021) from Earth observation (EO) data. In part, this is due to the need for an abundance of labeled data to enable the CNN to accurately identify different signatures. For example, ImageNet, an openly available visual database designed for use in everyday contemporary

Figure 1. Demonstration of the convolution of an image with an edge detection filter. On the left is the initial image, in the middle is an edge detection filter and on the right is the resulted image, which shows the edges of the initial image.

object recognition comprises 14,197,122 images (Russakovsky *et al* 2015). It is this volume of labelled data, which has enabled rapid advances in the use of CNN in day-to-day tasks. In archaeology however, similarly to other fields, the amount of freely available, properly labeled data is currently limited. Furthermore, online sharing of such data is often restricted by confidentiality issues that arise often from local legislation, related with the effort to protect these sites from looting.

In this paper, we offer a route forward by using openly available satellite high spatial resolution imagery and through examining two neural network architectures: The SegNet (Kendall *et al* 2015), a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labeling; and a custom 8-layer CNN designed for this research (SimpleNet). We also open-up access to these tools through providing a packaged application (supplementary information) allowing readers to run their own analysis, helping them to evaluate the strengths and weaknesses of this network and begin a more open and inclusive conversation about their use in archaeology.

2. Convolutional neural networks (CNN)

In this section we briefly introduce the fundamental concepts of CNNs. Although a more extensive presentation of CNNs is beyond the scope of this work, the interested reader can find detailed introductions focusing on various aspects of CNNs in several works including, Nielsen (2015); Wu (2017); Alzubaidi *et al* (2021); Li *et al* (2021); and Ulku and Akagündüz (2022).

Deep learning algorithms are a type of machine learning technique that uses ANN of several layers in a hierarchical architecture to enable machines to process data in a nonlinear manner. Artificial neural networks consist of circuits of simple, yet highly interconnected, nodes to selectively transmit signals in a process that mimics the biological neurons (Hopfield 1982), thereby simulating the way biological neural networks work. These nodes are organized in

layers which process information by outputting dynamic state responses to external inputs (commonly a response from a previous layer). Data are introduced to the ANN through an input layer and results delivered with a final output layer. All intermediate layers are termed hidden layers, which carry out all the processing. The larger the number of hidden layers, the 'deeper' the network, enabling the identification progressively of more complex patterns and details. For example, the first layer may learn recognizing edges in an image, the second shapes, the third objects and so on.

Information is passed between layers through connections that are characterized by weights and biases, so that the received total output corresponds to a weighted sum of individual node-inputs, plus some bias. The result output may or may not exceed a threshold defined by a pre-set activation function such as a sigmoid or most commonly a rectified linear activation function (ReLU; see below), essentially deciding if this information should be transmitted to the next layer (forward passed), as it is or in a modulated form, or rather filtered out. The optimal values of each weight and bias are defined by the training of the network: a non-linear optimization process whereby a cost function representing the distance between training labeled data and that predicted by network results is minimized.

The number of required deep layers within the network, and therefore indirectly the number of unknowns (i.e., parameters that are to be tunned through the training), depends on the complexity of the patterns to be identified and the amount of labeled data. At present, a limited number of labelled images for archaeology imposes a requirement for careful design of learning networks, keeping the number of layers and connections low enough to ensure that the optimization problem of network training is not under-determined i.e., the number of unknown parameters exceed the number of data and prior constraints that are used to regularize/stabilize the training and reduce the generalization error (overfitting) (e.g., Goodfellow *et al* 2016).
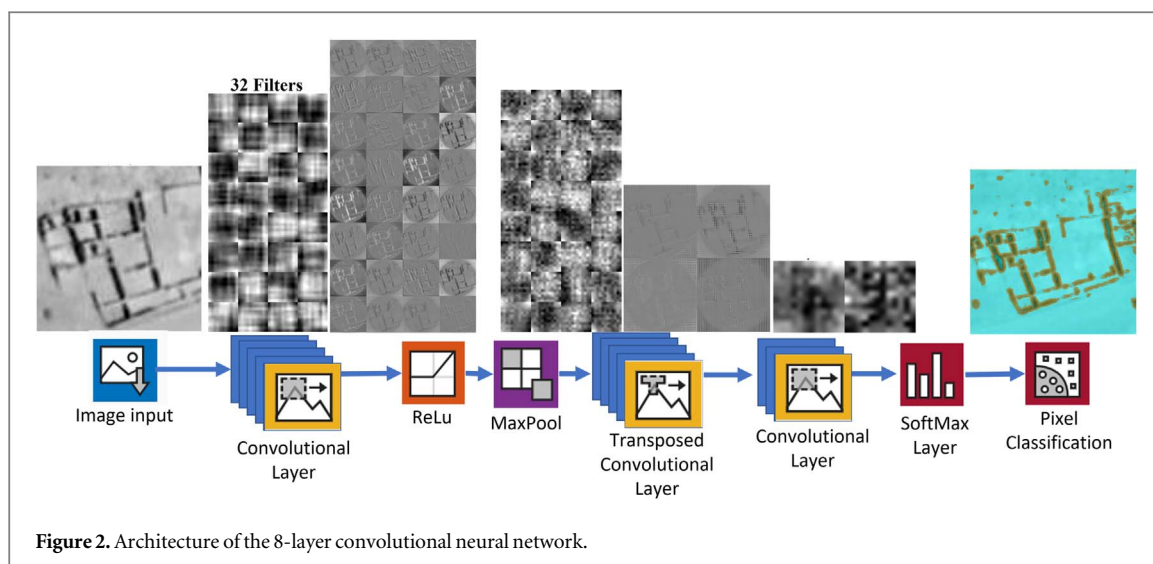
Figure 2. Architecture of the 8-layer convolutional neural network.

Table 1. Archaeological sites in Peru used to train the algorithm.

Archaeological areas & sites	Coordinates WGS84 (centre point)	Period
La Centinela (Chincha Valley)	−13.450385, −76.171092	Inca (AD 1476–1532) Late Intermediate (AD 1000–1476 AD)
Cahuachi (Nazca Valley)	−14.818241, −75.117462	Early Intermediate (c. 200 BC–AD 600)
Caral (Supe Valley)	−10.890938, −77.521858	Late Preceramic (c. 3000–1800 BC)
Tambo Colorado (Pisco Valley)	−13.704619, −75.829431	Inca (AD 1476–1532)

Table 2. Additional archaeological sites (areas) in Peru to further train the algorithm.

Archaeological areas & sites	Coordinates WGS84 (centre point)	Periods
Various (Lower Ica Valley)	−14.614319, −75.614994	Various (1800 BC–AD 1534)
Nazca Geoglyphs (Pampa de San José, Nasca Valley)	−14.696486, −75.178422	Early Intermediate (200 BC–AD 600 AD)
Cerro Sechín (Casma Valley)	−9.480703, −78.258997	Initial Period (1600 BC)

Here we examine two different, supervised, fully convolutional, neural networks: one based on the architecture of Semantic Segmentation called SegNet (Kendall *et al* 2015; Badrinarayanan *et al* 2017); and the other a custom 8-layer network developed by the authors called SimpleNet. Both are fully convolutional neural networks, a category of network consisting of locally connected layers so that each neuron only receives input from a small local subgroup of the pixels in the input image. Such Layers can perform convolution/deconvolution, pooling (i.e., a sample-based discretization process that effectively down-samples the image) and up-sampling, but not containing fully connected layers, and thus requiring significantly less memory and computational power (Long *et al* 2015). Semantic segmentation algorithms have been used widely in classifying features in various remote sensing images including high resolution Google Earth images (Yu *et al* 2021). Additionally, the custom 8-layer network was designed to be implemented for the low number of labeled data used in this work. In the following sections, we describe the architecture and functionality of these two networks.

2.1. SegNet

SegNet is a deep fully convolutional neural network that segments the image by classifying each pixel independently. It consists of an encoder network with 13 layers, each designed for object classification. Each layer is convolved using a set of 2D filters to produce a set of feature maps of increasing complexity as described previously. These maps are later batch normalized i.e., to have a mean output close to 0 and the output standard deviation close to 1. Next, a ReLU activation function is applied followed by down-sampling using a max pooling layer with a 2 × 2 non overlapping window (Kendall *et al* 2015; Badrinarayanan *et al* 2017). The ReLU activation function is a linear function that outputs the input if it is positive, or else, outputs zero (Hara *et al* 2015). The max pooling function calculates the maximum, in each patch of each feature map (Chollet 2017). In the final layer the resulting output, from the previous step, is subsampled by a factor of 2 while the boundary information is also stored. This is crucial as during the successive down-sampling operations the high frequency details of the image are lessened resulting in

Figure 3. A sample of the 2000 Training images, of size 256 × 256 × 3 pixels (Google Earth imagery), from various archaeological areas around Peru. The top row shows the initial images and the bottom row the labeled images.

blurry and inaccurate boundaries. However, boundaries are important in small objects and structures such as buildings, cropmarks etc and by storing this information it can be retrieved during the decoding stage.
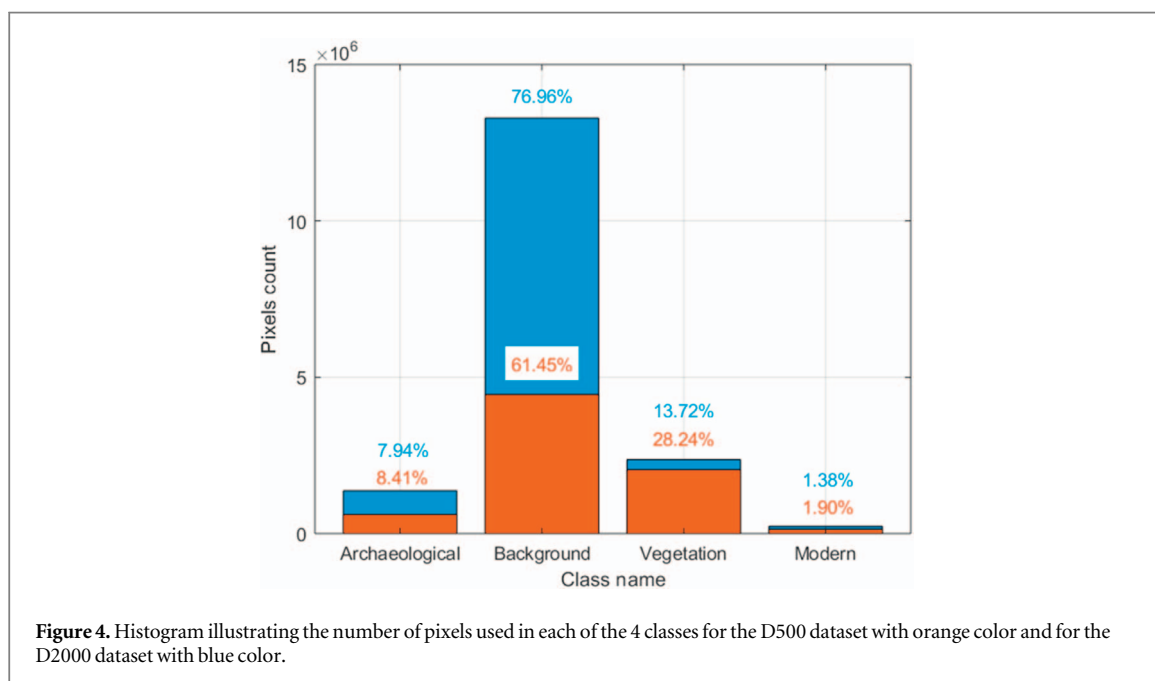
The network consists of 13 decoder layers each one corresponding to its respective encoder layer. The role of the encoder layers is to semantically project the lower resolution features extracted (learnt) by the encoder, onto the higher resolution image space to get a dense classification, i.e., a classification for each pixel in the original sized image. Each decoder layer produces dense feature maps (images) by up-sampling its input feature maps (the output of the previous layer) using the memorized max-pooling indices produced on the previous stage. Then convolution is applied using a trained dictionary of filters to produce dense feature maps. The final decoder output is fed into a SoftMax classifier, i.e., a layer that assigns each pixel independently to a class according to a probability score among the candidate classes (e.g., Nielsen 2015; Alzubaidi *et al* 2021).

2.2. A custom 8-layer convolutional neural network (SimpleNet)

Since the amount of labeled data available for archaeology is limited, we constructed a custom 8-layer convolutional neural network (SimpleNet), based on the SegNet architecture with the aim of keeping the number of layers and trainable parameters as low as possible while achieving adequately accurate results. The first layer is an image input layer that receives RGB images. The next layer is a convolutional layer with 32 trainable filters applied in a non-overlapping moving window of size 5 × 5 and with stride 1. Stride shows how much the filter shifts around the input volume (in our case it shifts by one unit) while the filter approximates the Laplacian (i.e., a 2D second spatial derivative) of the Gaussian operator and essentially when convolved with an image derives as an output an approximation of its second spatial derivative. This means that in regions where the image has constant intensity the filter's response will be zero. In regions where the intensity (i.e., pixel brightness) changes rapidly, however, such as at the edges of an object, the filter's response yields high amplitudes (figure 1).

The filters can be conceived of as 2D images whose shape and color are adjusted through the training process to optimally express different features of the data (e.g., figure 2). Next, a rectified linear unit (ReLU) is applied followed by a max-pooling with a 2 × 2 non overlapping window with stride 2 and a padding with 0's. This is the most common configuration as it discards the 75% of the activations in an input image due to down-sampling by 2 in both width and height. Following this, a transpose convolution is applied with the same number of filters and a window with 4 × 4 size and stride 2. Likewise, this is a common configuration, as the divisibility of the window size by the stride mitigates the problem of checkerboard artifacts in the up-sampled image (e.g., Odena *et al* 2016). The sixth layer is another convolutional layer of 1 × 1 window size and stride 1. Then, a SoftMax classifier is applied, to the final output from the previous layer, to assign each pixel into a class. Finally, the image is segmented into the assigned classes by a classification layer that calculates the class weighed cross-entropy loss (e.g., Bishop 2006). The 8-layer convolutional neural network technique is illustrated in figure 2.

Figure 4. Histogram illustrating the number of pixels used in each of the 4 classes for the D500 dataset with orange color and for the D2000 dataset with blue color.

3. Training and optimisation

3.1. Data

We used openly available high-resolution images from Google Earth of archaeological sites in Peru as a training set for both networks. This geographical region was chosen for its continued discovery of new sites using remotely sensed data (Ruggles 2015; Bikoulis *et al* 2018; Cigna and Tapete 2018) and the availability of data from previous large-scale archaeological terrestrial surveys for evaluation purposes.

Initially, we labeled 500 images from 4 different archaeological sites, (table 1). A small part of the Tambo Colorado archaeological site was then used to train the algorithm and a larger area of the same site for testing.

Later, we augmented the original 500 images with a further 1500 from wider archaeological areas and sites across Peru to further train the algorithm (table 2) and check its performance as the number of labeled data increase. These additional images consist mostly of geoglyphs (usually linear features) marked in open desert pampa environments. Figure 3 shows some samples of labeled images used in this work.
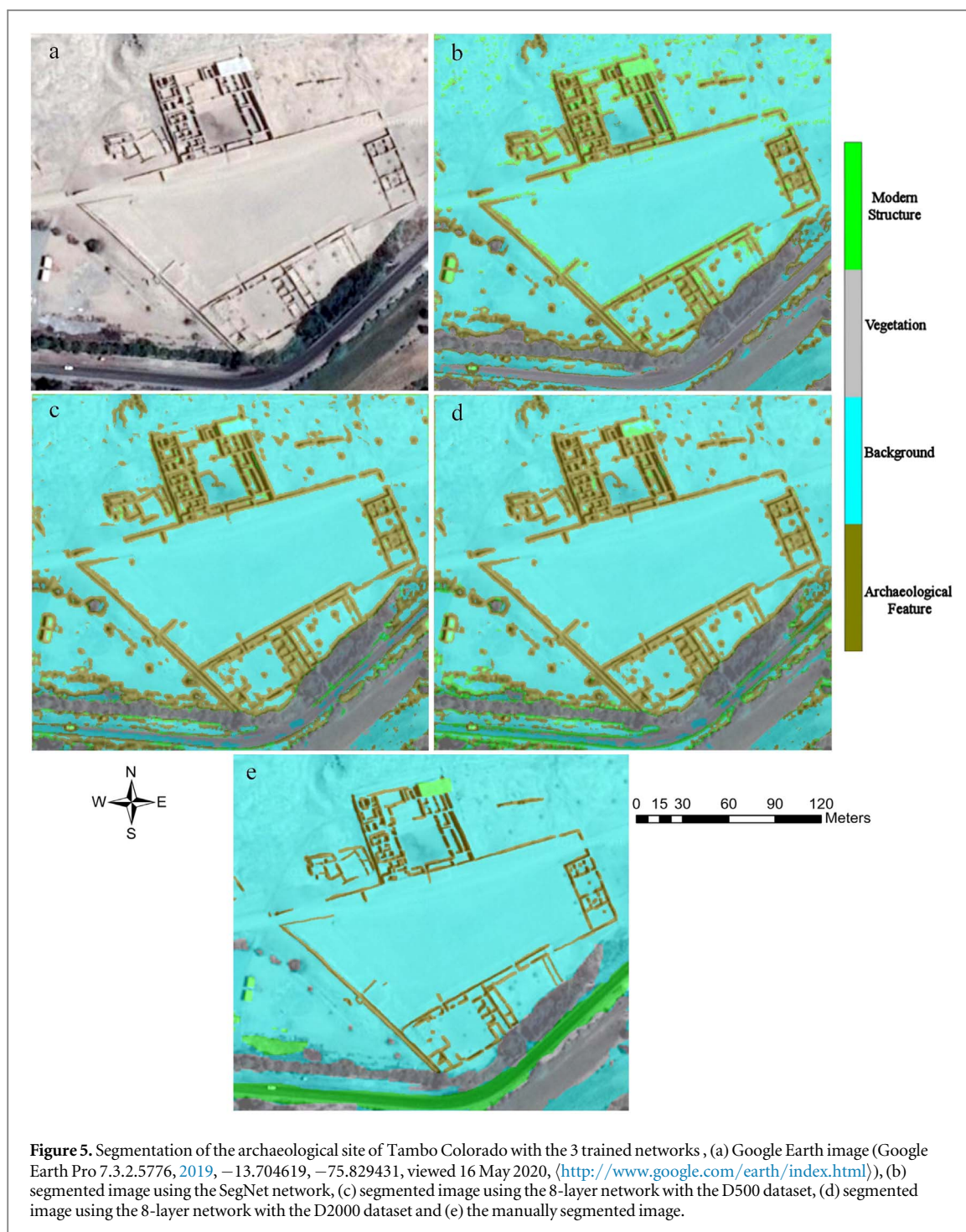
3.2. Optimisation process

The data were labelled with the ImageLabeler program in Matlab 9.6 using four different classes: 'archaeological', 'modern', 'vegetation' and 'background'. As 'archaeological' we included every target of archaeological interest, regardless of shape, condition, color, period etc We labeled linear, rectilinear, and circular features that were clearly visible in the Google Earth imagery, corresponding to a large variety of archaeological features. We used such broad terminology because the target of this work was to further increase the number of training images available to users in

Table 3. Optimal set of parameters for SegNet, 8-layer D500 and D2000 networks.

Parameter name	Value
Gradient decay factor	0.9000
Squared gradient decay factor	0.9990
Epsilon	1e-08
Initial learning rate	1e-04 (D500) and 1e-03 (D2000)
Drop rate factor	0.4
Drop period	5
L2- regularization parameter	1e-09
Gradient threshold method	Using the L2-norm
Max epochs	30
Mini batch size	5 (D500) and 15 (D2000)
Shuffle	At every epoch

future, with sub-classification open as an option to those who wish to make use of the dataset. As 'modern' we labeled modern structures such as modern buildings and vehicles. 'Vegetation' incorporates areas of grass, plants, and trees. Finally, as 'background' we classified everything else, such as soil, non-paved roads, and fields without vegetation. Images in the initial set of 500 were denoted as D500, and in the larger 2000 set as D2000. Images for the sites/areas of interest were extracted from Google Earth in RGB (Red, Green, Blue) as jpg files. Our goal, is to train an algorithm to use high resolution, freely available Google Earth images. Unfortunately, Google Earth does not provide raw images therefore we have to rely on the already processed images that are made available through the Google Earth application. It should be noted here that at present, the high-resolution images in Google Earth application are not available in Google Earth Engine and therefore is not possible to use this environment to train dataset.

Figure 5. Segmentation of the archaeological site of Tambo Colorado with the 3 trained networks , (a) Google Earth image (Google Earth Pro 7.3.2.5776, 2019, −13.704619, −75.829431, viewed 16 May 2020, ⟨http://www.google.com/earth/index.html⟩), (b) segmented image using the SegNet network, (c) segmented image using the 8-layer network with the D500 dataset, (d) segmented image using the 8-layer network with the D2000 dataset and (e) the manually segmented image.

Although the native resolution is not provided, we estimate it at a ground resolution of 0.5m (Airbus imagery). Next, 256 × 256 × 3pixel tiles were generated from the original images. The labels for each image were stored in the form of an 8-bit monochromatic copy of the image, with 4 distinct intensity values reserved to represent each one of the 4 different classes, in addition to the null-intensity that represents the unlabeled regions of the image. Labeled pixels in each dataset corresponded to the classes of background (61.45% and 79.96% respectively); vegetation (28.24% and 13.72%, respectively); archaeological (8.41% and 7.94%); and modern (1.90% and 1.38%),

(figure 4, histogram). To avoid bias in the learning process due to this imbalance between classes (Verschoof-van der Vaart *et al* 2020), we weighted each class by the inverse of its frequency. The weighting was applied at the last layer of the classification networks, i.e., the Pixel Classification Layer.

For the training of both networks the Adaptive Moment Estimation (Adam) optimizer (Kingma and Ba 2015) was used. After multiple tests over both training datasets (D500 and D2000) and both networks, we concluded the quasi-optimal set of parameters shown in table 3. These were kept unchanged for both networks in subsequent work.
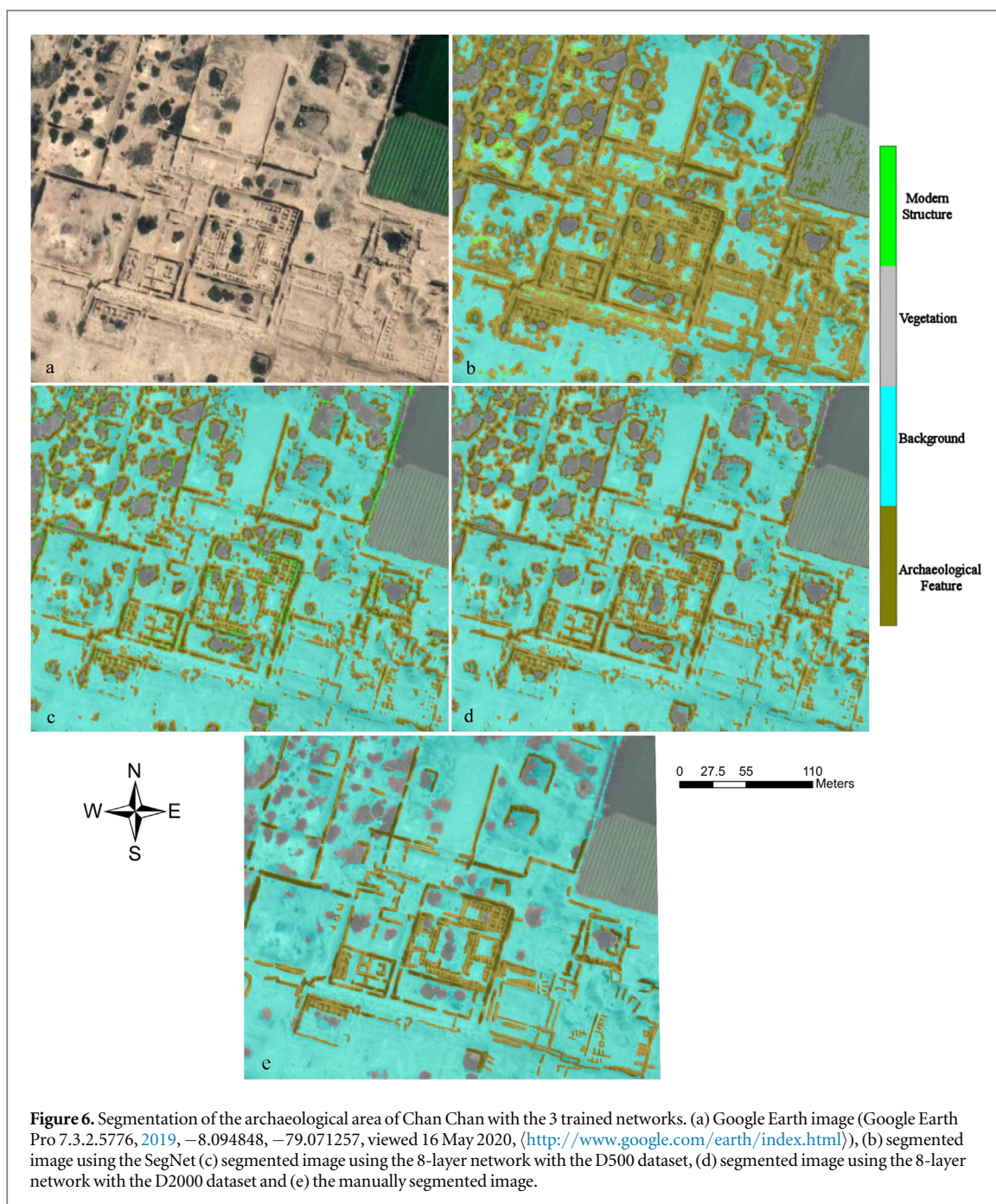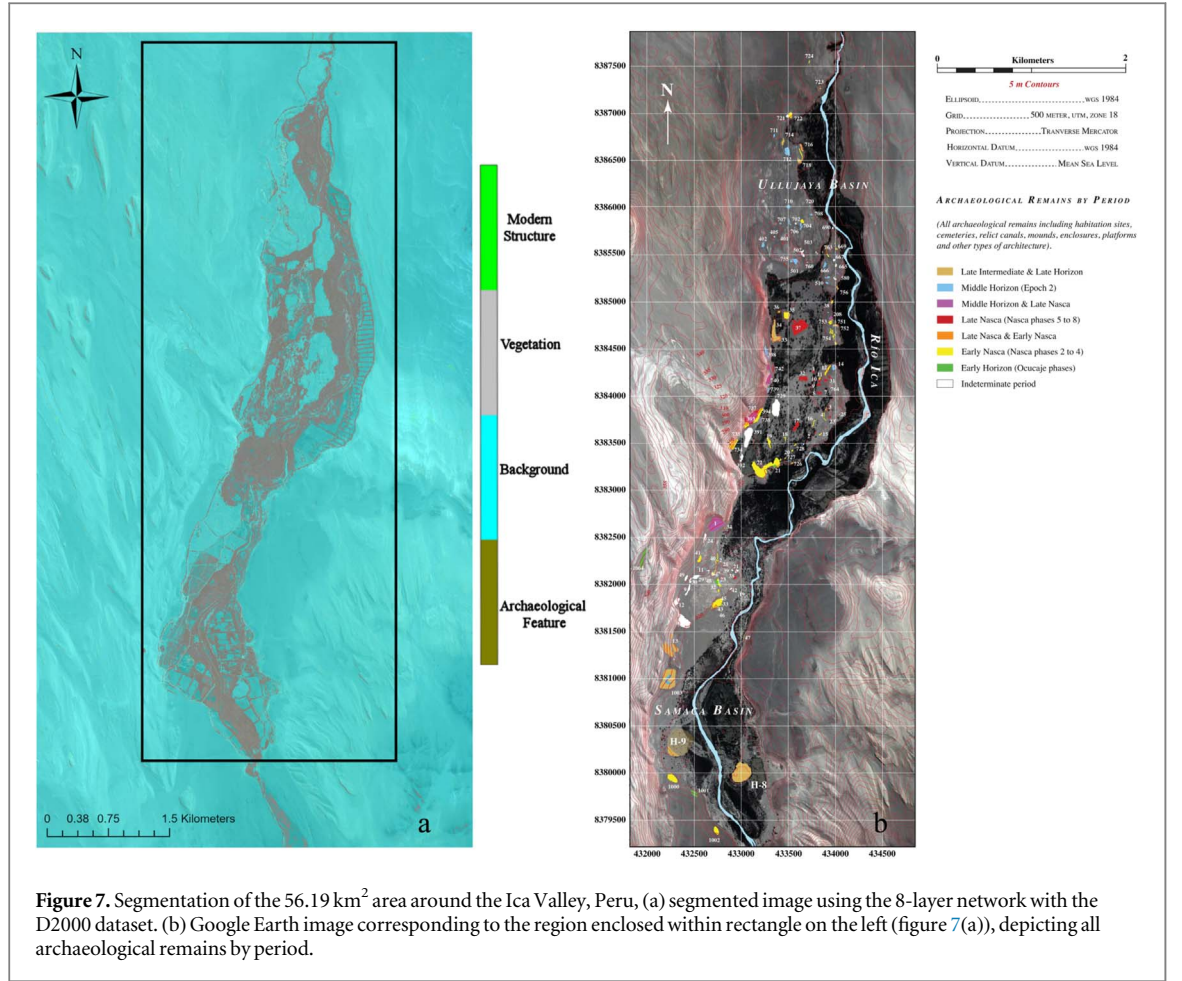
Figure 6. Segmentation of the archaeological area of Chan Chan with the 3 trained networks. (a) Google Earth image (Google Earth Pro 7.3.2.5776, 2019, −8.094848, −79.071257, viewed 16 May 2020, ⟨http://www.google.com/earth/index.html⟩), (b) segmented image using the SegNet (c) segmented image using the 8-layer network with the D500 dataset, (d) segmented image using the 8-layer network with the D2000 dataset and (e) the manually segmented image.

Table 4. Comparison between the manually labeled image and the segmented images of each trained network for the archaeological area of Tambo Colorado.

Class Name	True positive %			False positive %			False negative %		
	SegNet	SN-D500	SN-D2000	SegNet	SN-D500	SN-D2000	SegNet	SN-D500	SN-D2000
Archaeological Feature	87.41	97.41	97.39	13.97	18.62	16.00	12.58	2.58	2.60
Background	87.45	86.99	88.12	21.67	23.34	23.57	12.54	13.00	11.87
Vegetation	80.67	78.65	77.86	4.696	2.70	2.20	19.32	21.34	22.13
Modern Structure	6.61	9.85	32.74	5.591	1.68	2.39	93.38	90.14	67.25

We used the IRIDIS supercomputer of the University of Southampton to train the network. The training ran in parallel using 12 CPU cores and one Node with 264 GB of main memory and took approximately 3 h to complete. For the calculations we used MATLAB 9.6 (2019a).

Figure 7. Segmentation of the 56.19 km^2 area around the Ica Valley, Peru, (a) segmented image using the 8-layer network with the D2000 dataset. (b) Google Earth image corresponding to the region enclosed within rectangle on the left (figure 7(a)), depicting all archaeological remains by period.

3.3. Evaluation

Evaluation of how successful the training had been carried out through analysis of additional images from Peru which had not been included in the training datasets. We compared the results from the D500 and D2000 trained networks, as well as SegNet versus SimpleNet. The evaluation image covered two different areas:

3.3.1. Tambo colorado

Tambo Colorado (−13.704619, −75.829431) is a large Inca administrative and ceremonial center built in the last quarter of the 15th C in the Pisco Valley, south coast Peru alongside the imperial highway linking the coast with the southern highlands (Hyslop 1985; Protzen and Morris 2004). A small part of the archaeological site (not shown in figure 5) was used to train the algorithm, while the rest of the site and its surrounding area (figure 5) was used to test it.

To evaluate the performance, we manually labeled the image with the ImageLabeler program and then used this baseline to compare each segmented image independently. The labelling was completed according to what is clearly visible and can be confidently classified as archaeological. The results, expressed as percentage, are shown in table 4 There are three types of results, the True Positive (*R*) corresponds to the successfully identified pixels:

$$R|_c = \frac{\sum_{i=1}^{N} \delta_{c,G(i)} \cdot \delta_{c,S(i)}}{\sum_{i=1}^{N} \delta_{c,G(i)}} \qquad (1)$$

where *N* is the total number of pixels, *i* is the pixel index within the segmented image, δ is the Kronecker delta, *c* is the current class, *G* is the ground truth image in column indexing order and *S* is the segmented image of the current trained network.

False Positive (*FP*), represents a pixel falsely identified as a member of the class though, it belongs in another class.

$$FP|_c = \frac{\sum_{i=1}^{N} (1 - \delta_{c,G(i)}) \cdot \delta_{c,S(i)}}{\max\left(\sum_{i=1}^{N} (1 - \delta_{c,G(i)}), N\right)} \qquad (2)$$

False Negative (*FN*), corresponds to a pixel which although it belongs to the class it was falsely classified as not belonging to it.

$$FN|_c = \frac{\sum_{i=1}^{N} \delta_{c,G(i)} \cdot (1 - \delta_{c,S(i)})}{\max\left(\sum_{i=1}^{N} \delta_{c,G(i)}, N\right)} \qquad (3)$$

The data show in table 4 and the final segmented images (figure 5) show similar results between the 8-layer CNN with the D500 and D2000 datasets, with the exception of the modern structure class which is far better covered in the D2000 trained network. SegNet, on the other hand, gives less reliable results with
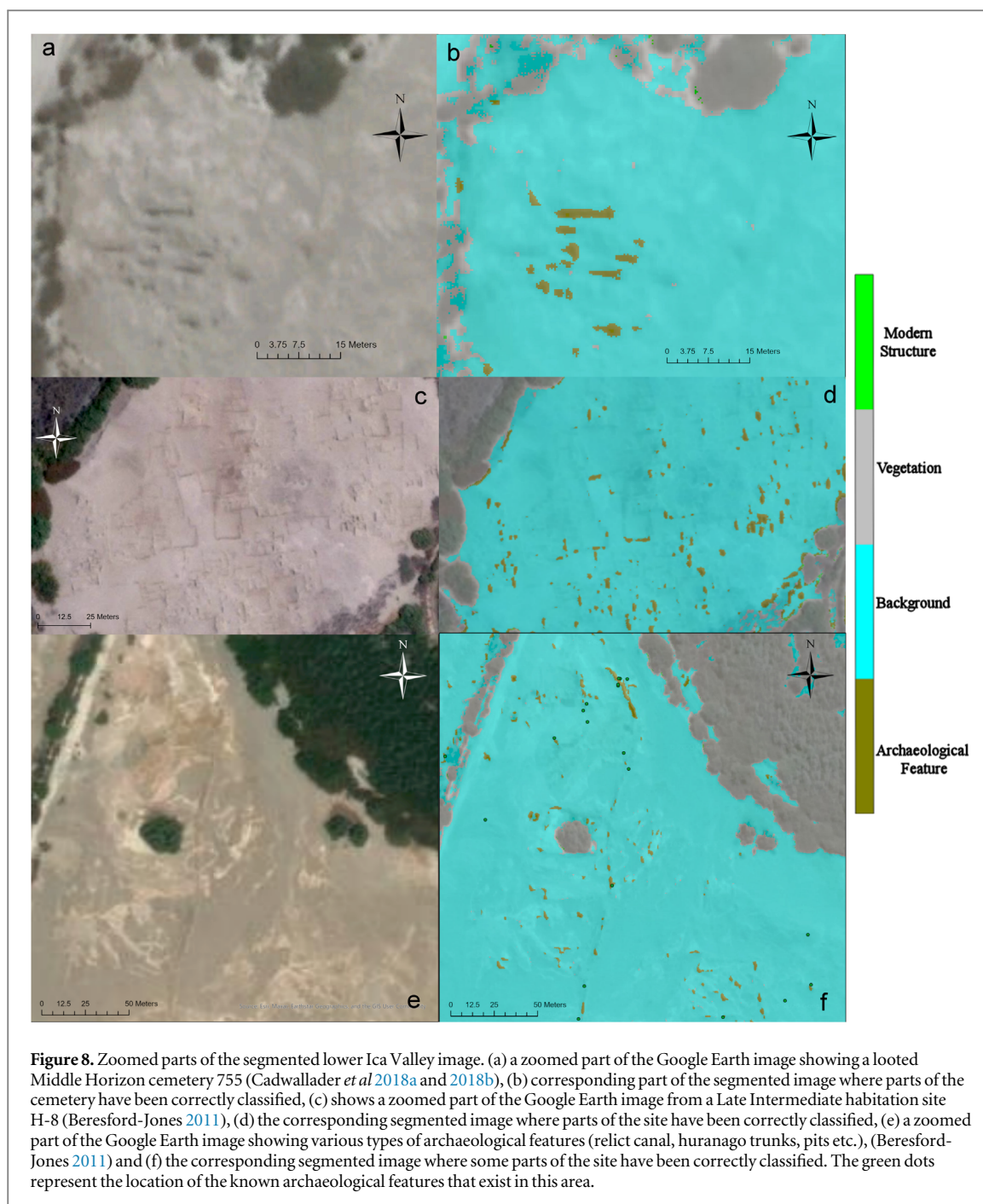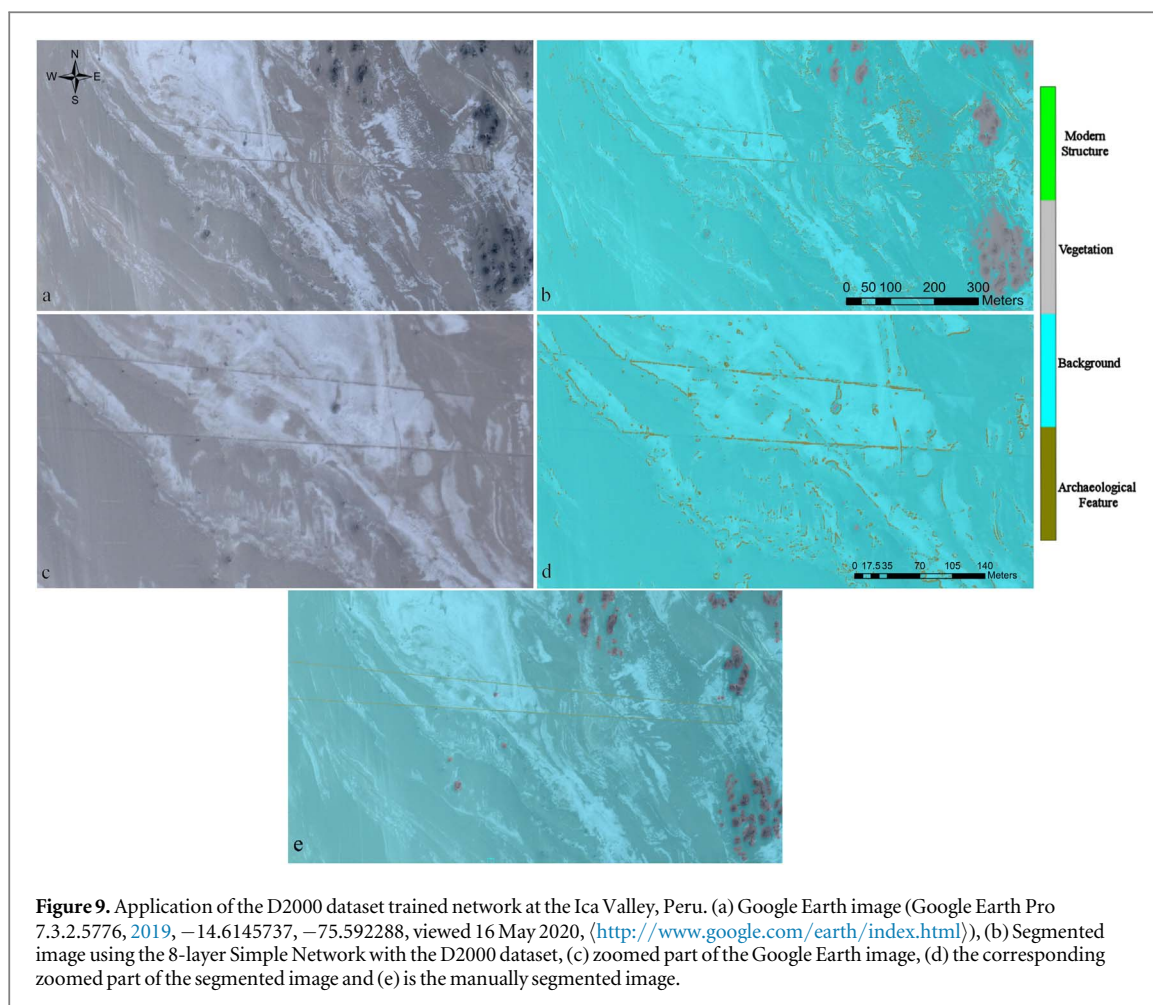
Figure 8. Zoomed parts of the segmented lower Ica Valley image. (a) a zoomed part of the Google Earth image showing a looted Middle Horizon cemetery 755 (Cadwallader *et al* 2018a and 2018b), (b) corresponding part of the segmented image where parts of the cemetery have been correctly classified, (c) shows a zoomed part of the Google Earth image from a Late Intermediate habitation site H-8 (Beresford-Jones 2011), (d) the corresponding segmented image where parts of the site have been correctly classified, (e) a zoomed part of the Google Earth image showing various types of archaeological features (relict canal, huranago trunks, pits etc.), (Beresford-Jones 2011) and (f) the corresponding segmented image where some parts of the site have been correctly classified. The green dots represent the location of the known archaeological features that exist in this area.

Table 5. Comparison between the manually labeled image and the segmented images of each trained network for the archaeological area of Chan-Chan.

Class name	True positive %			False positive %			False negative %		
	SegNet	SN-D500	SN-D2000	SegNet	SN-D500	SN-D2000	SegNet	SN-D500	SN-D2000
Archaeological Feature	95.16	64.76	73.94	45.58	18.45	20.69	4.83	35.23	26.05
Background	48.85	77.95	78.15	1.12	12.86	12.25	51.54	22.04	21.84
Vegetation	69.81	78.48	77.78	0.83	2.61	1.41	30.18	21.51	22.21
Modern Structure	N/A	N/A	N/A	1.26	2.28	0.00	N/A	N/A	N/A

Figure 9. Application of the D2000 dataset trained network at the Ica Valley, Peru. (a) Google Earth image (Google Earth Pro 7.3.2.5776, 2019, −14.6145737, −75.592288, viewed 16 May 2020, ⟨http://www.google.com/earth/index.html⟩), (b) Segmented image using the 8-layer Simple Network with the D2000 dataset, (c) zoomed part of the Google Earth image, (d) the corresponding zoomed part of the segmented image and (e) is the manually segmented image.

lower values for the classes of archaeological features, background, and modern structures.

3.3.2. Chan chan
Chan Chan (−8.094848, −79.071257) was the urban capital of the Chimú Empire (c. AD 900–1475) built in the Moche Valley near the modern city of Trujillo, north coast Peru. Covering over 20 km^2 it comprises nine palace-mausoleum enclosures (cuidadelas) of the Chimú ruling dynasty some with extant adobe walls up to 10m high, surrounded by dense agglomerations of thousands of humble dwellings and workshops (Holstein 1927; Moseley 1975; Smailes 2011).

The data shown in table 5 and the final segmented images (figure 6) show that for Chan-Chan the SegNet algorithm segmented archaeological features better than the other two methods, although segmentation of background and vegetation classes is poor compared with the other two algorithms. Moreover, both SegNet and the 8-layer D500 networks appear to have falsely classified pixels as modern structure when they should belong to another class. In table 5, N/A represents the lack of modern structures in the image. Generally, the 8-layer D2000 network seems to give the most balanced and reliable results.

4. Application to larger and more complex landscape

The evaluation described above focused on the ability of the network to identify linear features in areas where there was a strong and distinctive archaeological signature. While a critical first step, this testbed is hardly representative of more complex and extensive areas typically of interest to archaeologists. To investigate the performance and efficiency of the trained network in such circumstances a 56.19 Km^2 area of the lower Ica Valley (Peru) was analyzed with the D2000 8-layer network. This area was chosen because we have carried out extensive pedestrian and aerial archaeological survey there since 2002 (Beresford-Jones 2011; Haburaj *et al* 2017; Cadwallader *et al* 2015, 2018a and 2018b) thereby allowing for direct testing of AI generated results.

4.1. Lower ica valley, peru
The archaeological record of the lower Ica Valley comprises looted cemeteries, eroded mounds and fragmentary adobe architecture, dispersed pottery scatters, middens and linear relict canal features amidst a largely wind-deflated arid landscape, patchy riparian vegetation, and some modern agriculture

Table 6. Types of archaeological sites around the Ica Valley and the number of correctly classified features with the total features of the same type.

Archaeological features	Correctly classified/total
Looted tombs	**5/5**
Tombs	**22/22**
Grey layer	**2/2**
Cemetery	**11/16**
Canal & Canal Fragments	**11/24**
Post holes	**11/12**
Habitation site	**8/9**
unknown	**7/16**
Midden	**1/16**
Mound	**1/32**
Fired Earth pit	**4/14**
Pit	**2/6**
Looters pit	**1/3**
Ceramic scatter	**2/6**
Adobe structures	**3/13**
Round stone wall	**1/2**
Plant Macro	**8/20**
Huarango trunk	**3/24**
Platform	**5/6**
Camelid bones	**0/1**
Datum	**0/2**
Geoglyphs	**0/9**
Duracrete	**0/1**
Nazca	**0/6**
Ocucaje	**0/3**
Feature	**0/7**
Fluvial deposit	**0/1**

(Beresford-Jones 2011, Cadwallader *et al* 2018a and 2018b). Such remains provide highly variable archaeological signatures, mostly far more diffuse and contrasting significantly with the well-defined rectilinear architecture in the labelled image dataset used to train the network.

The large size of the study area corresponds to an image of 17853 × 34429 x 3 pixels. This meant that segmentation was applied patch-wise, with each patch measuring 1024 × 3072 × 3. Similarly, the resulted segmented image was constructed as a mosaic, patch by patch. These calculations took about 1 and a half hours on a mid-range laptop with intel i7, 7500 CPU and 16 GB of memory, demonstrating the computational efficiency and accessibility of this approach, even when applied to large geographic areas.

The trained algorithm classified a significant number of features as archaeological in a region in which we had previously surveyed 278 diverse archaeological features of 27 different types, including habitation sites, mounds and cemeteries, but also artifacts well below the resolution of the satellite imagery, such as ceramic scatters. Nevertheless, we superimposed the known site locations over the segmented images for evaluation. After segmentation, at least 102 archaeological sites were correctly identified at the pixel level (figure 7). The areas of greatest success related to larger and linear features (figure 8) which are the types of

archaeological features that the algorithm was trained to identify. As was expected given the nature and resolution of the input labelled images, the algorithm was less successful at identifying more diffuse features such as eroded mounds and middens (table 6).

The desert landscape of the lower Ica Valley includes many linear geological exposures which, in some cases, were falsely classified as archaeological by the algorithm because of their similarity to the training data. Although the lower Ica Valley image is not labeled at the pixel level, and thus direct quantified comparisons are not possible, the broader site location-to-classified-pixel assessment given in table 6 and figure 8, when tested against archaeological survey data, shows good agreement across all four classes.

The pampa alongside the lower Ica Valley also manifests a significant number of trapezoidal geoglyphs (Cadwallader *et al* 2018b), although they can be challenging to identify and map by aerial image analysis or pedestrian survey because of their often-remote locations and relative invisibility at ground level. We selected an area that included a previously surveyed trapezoidal geoglyph (−14.61457, −75.59229) with which we could test the trained algorithm (figure 9(a)).

The resulting segmented image (figure 9(b)) shows that while the algorithm has in a few places falsely identified some geological exposures as archaeology (2.4%) it has also successfully classified 29.8% of the geoglyph.

5. Discussion

In this work we examined two different deep learning networks, SegNet and an 8-layer network designed for the project (SimpleNet). Initial evaluation suggested that the 8-layer network performed better at identifying features of interest, and that increasing the amount of labelled data available improved the performance (albeit not in dramatic fashion between 500 to 2000 images). Our results show that there is not systematic behavior in the occurrence false-positives/negatives break down, and this largely depends upon the properties of each case/tested image. For example, false positive instances of archaeological features are dominated in regions that should be classified as background, in the presence of rough topography. This is because at its heart the network focuses on picking out linear and slightly rounded features, and currently performs less well on more circular to curvilinear forms such as eroded mounds and middens. Rather than a reflection of conceptual limits, this is an outcome of the labelled images used to train the network. This, in turn, means that there are a large number of false positives, mostly of geological or rough topographic features in the desert landscape picked out as of archaeological interest. This is illustrated in figure 10 which shows an area of the Nazca valley that has been processed with geological features classified as archaeology. Yet, within
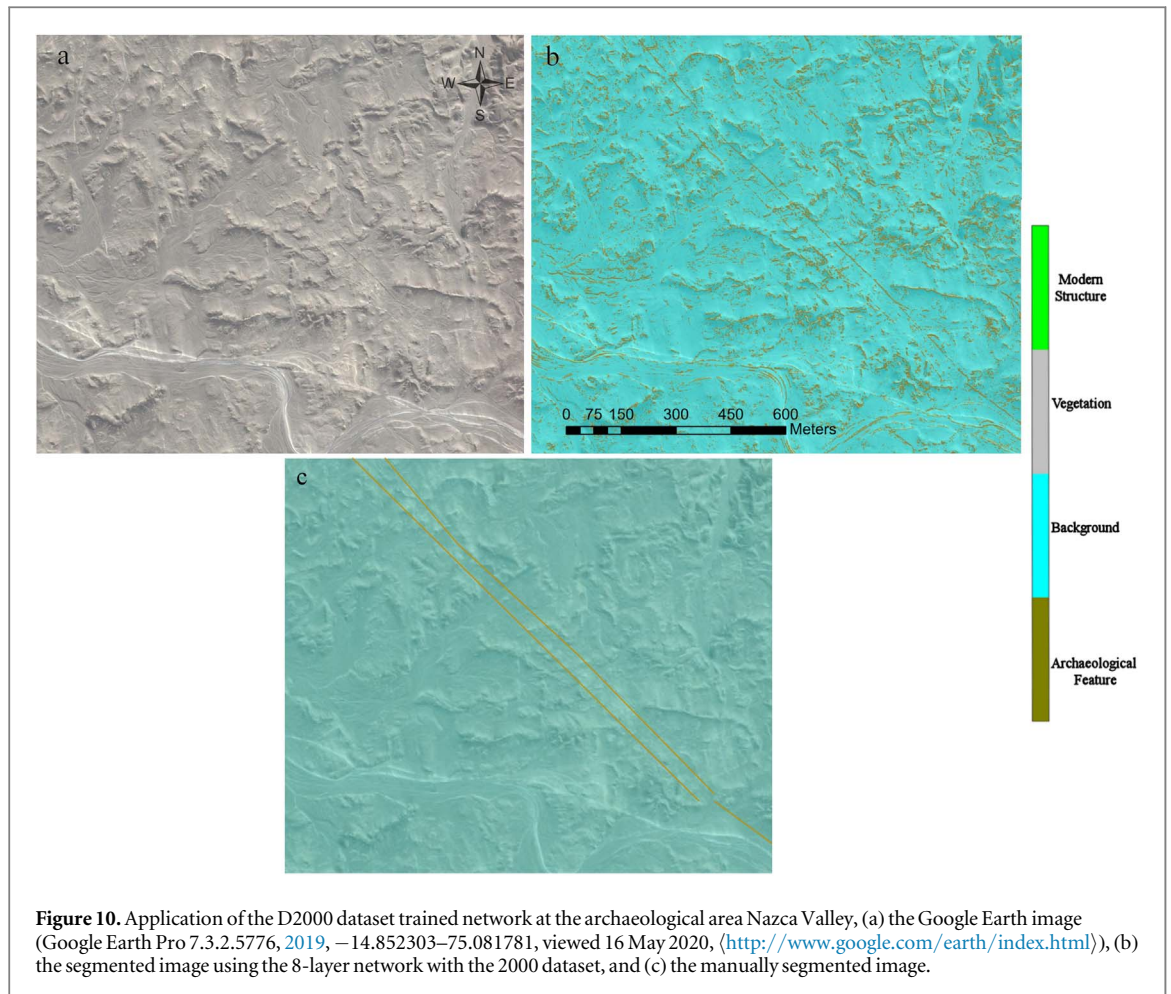
Figure 10. Application of the D2000 dataset trained network at the archaeological area Nazca Valley, (a) the Google Earth image (Google Earth Pro 7.3.2.5776, 2019, −14.852303–75.081781, viewed 16 May 2020, ⟨http://www.google.com/earth/index.html⟩), (b) the segmented image using the 8-layer network with the 2000 dataset, and (c) the manually segmented image.

figure 10(b), it is now also possible to pick out much more easily archaeological geoglyph features, which otherwise would not have been so easily identifiable.

This, together with the comparison between the network's analysis of the lower Ica Valley and the results of pedestrian archaeological survey, shows the significant contribution such approaches can play in large area analysis. There is no doubt that the network has misidentified some features, and only partially recognized or failed to acknowledge others.

In other cases, false negative pixel classifications occur because the delineation of the archaeological features in the segmented images is wider than what it has been manually marked but it is certainly associated with true features (e.g., figure 6). Possibly the application of the method in larger amounts of data with wider variety in image, background and target properties, where ground truth is available, will allow for more insightful statistical analysis recognizing possible systematic errors and/or improving its accuracy when these data are included in training.

Nonetheless, this was a network trained on only 2000 images, carried out in rapid assessment. Even at this level of accuracy as a tool to aid and accelerate large area survey these are significant results. Clearly, deep learning methods have significant potential to identify archaeological targets of different dimensions, spectral and geometrical properties.

Moreover, the shortcomings of the analysis presented here can be easily overcome: most straightforwardly by radically increasing the amount of labelled data available for the network to learn. It is this task, of increasing the shared resource of open access labelled data that other disciplines such as marine biology are now rapidly working to achieve (Gregory *et al* 2019; Sagi *et al* 2020). Deep learning techniques work best when this pool reaches the order of millions, or more. What we have demonstrated here is that even in the initial stages of building these resources there are significant benefits to be had. A possible future direction is to automatically label images using georeferenced mapping data of archaeological sites. This can include features that are not directly perceptible in satellite images and remove the human bias during the labeling. However, this task requires tracing the perimeter of the features to provide adequate accuracy in the labeling process. However, it should be noted that there is always the possibility that certain confirmed archaeological features may not be physically visible from the satellite image for various reasons (e.g., covered by vegetation, resolution, etc). Therefore, including such data could be challenging.

Another direction is to explore the benefits of transfer learning in improving the effectiveness of the segmentation. In transfer learning a pretrained network focused on a different application can be used as starting point, and then replace the very last layers which map the recognized features to the specific classes of the present application. Then a fine-tune training stage is required to optimize only the weights and the parameters of these final new layers. This allows to take advantage of the ability of the trained network to extract/recognize complicated features from its previous application, without requiring large amounts of data. Furthermore, the computational resources for the training is kept at minimum as only a small subset of parameters are to be optimized while the others of the previous layers are kept fixed.

6. Conclusions

In this work the data we used to train the algorithm are from various archaeological areas in Peru. The main target of this project is to test the performance of two different CNN algorithms and their applicability in limited number of training images and create a trained deep learning algorithm that will be able to identify different types of archaeological structures from around the world. We have shown that even with limited data, we can train a network to identify various kinds of archaeological features. Using openly available images like those provided by Google Earth, the trained network can successfully identify regions of archaeological interest, helping to guide analysts to locations for further investigation. Our future target is to continue building this library by adding more datasets from different types of sensors from other locations worldwide. In turn, these datasets become the foundations on which larger, more variable libraries of labelled images can be built, and shared with the community. It is only through doing this, through acting collectively, that the greatest gains will be made and the wider potential of CNNs realized within archaeology.

Acknowledgments

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

Alexandra Karamitrou ⓘ https://orcid.org/0000-0002-4142-1958
Fraser Sturt ⓘ https://orcid.org/0000-0002-3010-990X
Petros Bogiatzis ⓘ https://orcid.org/0000-0003-1902-7476
David Beresford-Jones ⓘ https://orcid.org/0000-0003-2427-7007

References

Anichini F, Dershowitz N, Dubbini N, Gattiglia G, Itkin B and Wolf L 2021 The automatic recognition of ceramics from only one photo: The ArchAIDE app *Journal of Archaeological Science: Reports.* **36** 102788

Agapiou A, Vionis A and Papantoniou G 2021 Detection of Archaeological surface ceramics using deep learning image-based methods and very high-resolution UAV imageries *Land* **10** 1365

Alzubaidi L *et al* 2021 Review of deep learning: concepts, CNN architectures, challenges, applications, future directions *J Big Data* **8** 53

Badrinarayanan V, Kendall A and Cipolla R 2017 SegNet: a deep convolutional encoder-decoder architecture for image segmentation *IEEE Trans. Pattern Anal. Mach. Intell.* **39** 2481–95

Baxter M J 2009 Archaeological data analysis and fuzzy clustering *Archaeometry* **51** 1035–54

Beresford-Jones D G 2013 *The Lost Woodlands of Ancient Nasca: A Case-study in Ecological and Cultural Collapse* (London: British Academy Scholarship Online) (2011; Online edn) (https://doi.org/10.5871/bacad/9780197264768.001.0001)

Bewley R H and Raczkowski W 2002 Aerial archaeology: Developing Future Practice *IOS Press, NATO Science Series, I: Life and Behavioural Sciences* **337** 311–27

Bikoulis P, Gonzalez-Macqueen F, Spence-Morrow G, Bautista S, Alvarez W and Jennings J 2018 Ancient pathways and geoglyphs in the sihuas valley of southern peru *Antiquity* **92** 1377–91

Bishop C M 2006 *Pattern Recognition and Machine Learning.* (New York: Springer) XX, 738

Bonhage A, Eltaher M, Raab T, Breuß M, Raab A and Schneider A 2021 A modified Mask region-based convolutional neural network approach for the automated detection of archaeological sites on high-resolution light detection and ranging-derived digital elevation models in the north german lowland *Archaeological Prospection* **28** 177–86

Cadwallader L, Torres S A, Pullen A G and Beresford-Jones D G 2018b Algunas contribuciones a la comprensión de la transición Paracas-Nasca, Samaca, valle bajo de Ica *Boletín de Arqueología de la Pontificia Universidad Católica del Perú* **25** 147–71

Cadwallader L, Arce Torres S, O'Connell T, Pullen A and Beresford-Jones D 2015 Dating the dead: new radiocarbon dates from the lower ica valley, south coast peru *Radiocarbon* **57** 765–73

Cadwallader L, Beresford-Jones D G, Sturt F C, Pullen A G and Arce S 2018a Doubts about how the middle horizon collapsed (c. AD 1000) and other insights from the looted cemeteries of the lower ica valley on the south coast of peru *Journal of Field Archaeology* **43** 316–31

Cao F, Liang J and Jiang G 2009 An initialization method for the K-Means algorithm using neighborhood model *Comput. Math. Appl.* **58** 474–83

Caspari G and Crespo P 2019 Convolutional neural networks for archaeological site detection—Finding 'princely' tombs *J. Archaeolog. Sci.* **110** 104998

Chi M, Plaza A, Benediktsson J A, Sun Z, Shen J and Zhu Y 2016 Big data for remote sensing: challenges and opportunities *Proc. IEEE* **104** 2207–19

Chollet F 2017 *Deep Learning with Python.* 1st (Manning Publications: Simon and Schuster) 384

Cigna F and Tapete D 2018 Tracking human-induced landscape disturbance at the nasca lines UNESCO world heritage site in peru with COSMO-SkyMed InSAR *Remote Sens.* **2018** 572

Copeland B J 2020 Artificial intelligence. website name: encyclopedia britannica, encyclopedia britannica, inc. URL: https://britannica.com/technology/artificial-intelligence. Access Date: March 29, 2020

Cortes C and Vapnik V 1995 Support-vector networks *Mach. Learn.* **20** 273–97

Dey A 2016 Machine learning algorithms: a review *International Journal of Computer Science and Information Technologies* **7** 1174–9

Flores F C, Ugalde F G, Díaz J L P, Navarro J Z, Gastelum-strozzi A, Angeles M D P and Miyatake M N 2019 Computer algorithm for archaeological projectile points automatic classification *Journal on Computing and Cultural Heritage* **12** 1–30

Fukushima K 1980 Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position *Biol. Cybern.* **36** 193–202

Fukushima K, Miyake S and Ito T 1983 Neocognitron: a neural network model for a mechanism of visual pattern recognition *IEEE Transactions on Systems, Man, and Cybernetics. SMC-* **13** 826–34

Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P and Garcia-Rodriguez J 2018 A survey on deep learning techniques for image and video semantic segmentation *Appl. Soft Comput.* **70** 41–65

Goodfellow I, Bengio Y and Courville A 2016 Chapter 7: regularization for deep learning *Deep Learning* (Cambridge, MA: MIT Press) 7

Google Earth Pro 7.3.2.5776 2019 (http://google.com/earth/index.html) last accessed 16 May 2020

Gregory K, Groth P, Cousijn H, Scharnhorst A and Wyatt S 2019 Searching data: a review of observational data retrieval practices in selected disciplines *J Assoc Inf Sci Tech* **70** 419–32

Haburaj V, Berking J, Beresford-Jones D, Knitter D, Zeki L, Sturt F, Pullen A, Huaman O, Lane K and French C 2017 Geo-statistical methods to analyse changes in pre-Hispanic settlement patterns in the Río Ica catchment, Peru *Journal of Archaeological Science: Reports* **12** 272–87

Hara K, Saito D and Shouno H 2015 Analysis of function of rectified linear unit used in deep learning *international Joint Conference on Neural Networks (IJCNN) (Killarney, 2015 Jul 12)* 1–8

Ho K T 1995 Random decision forests *Proc. of 3rd Int. Conf. on Document Analysis and Recognition, Montreal, Quebec, Canada* 1, 278–82

Ho K T 1998 The random subspace method for constructing decision forests *IEEE Trans. Pattern Anal. Mach. Intell.* **20** 832–44

Holstein O 1927 Chan-chan: capital of the great chimu *Geographical Review* **17** 36

Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8

Hörr C, Lindinger E and Brunnett G 2008 New paradigms for automated classification of pottery *Proc. of the 36th Computer Applications and Quantitative methods in Archaeology Conf., Budapest, 2–6 April 2008* 366–75 (CD-ROM 268-277)

Hubel D H and Wiesel T N 1959 Receptive fields of single neurones in the cat's striate cortex *J. Physiol.* **148** 574–91

Hyslop J 1985 *Incawasi: The New Cuzco.* (Oxford: British Archaeological Reports International Series 234)

Jin X and Han J 2011 K-Means clustering ed C Sammut and G I Webb *Encyclopedia of Machine Learning.* (Boston, MA: Springer)

Kao W-C, Chung K-M, Sun C-L and Lin C-J 2004 Decomposition methods for linear support vector machines *Neural Comput.* **16** 1689–704

Kazimi B, Thiemann F and Sester M 2019 Object instance segmentation in digital terrain models *Lecture Notes in Computer Science (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). Computer Analysis of Images and Patterns: 18th International Conference, CAIP 2019 (Salerno, Italy, September 3–5, 2019)* (https://doi.org/10.1007/978-3-030-29891-3_43)

Kendall A, Badrinarayanan V and Cipolla R 2015 Bayesian segnet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding arXiv:1511.02680

Kingma D P and Ba J L 2014 Adam: a method for stochastic optimization 1412.6980 arXiv preprint arXiv:1412.6980

Kintigh K W and Ammerman A J 1982 Heuristic approaches to spatial analysis in archaeology *American Antiquity* **47** 31–63

Krizhevsky A, Sutskever I and Hinton G E 2012 ImageNet classification with deep convolutional neural networks *NeurIPS Proc.. Curran Associates, Inc.* (https://proceedings.neurips.c.c./paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf)

Küçükdemirci M and Sarris A 2020 Deep learning based automated analysis of archaeo-geophysical images *Archaeological Prospection*

Lambers K 2018 Airborne and spaceborne remote sensing and digital image analysis in archaeology *Digital Geoarchaeology.* (Springer: Springer, Cham) 109–22

Lambers K, Verschoof-van der Vaart W B and Bourgeois Q P J 2019 Integrating remote sensing, machine learning, and citizen science in dutch archaeological prospection *Remote Sensing* **11** 794

LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W and Jackel L D 1989 Backpropagation applied to handwritten zip code recognition *Neural Comput.* **1** 541–51

LeCun Y, Bottou L, Bengio Y and Haffner P 1998 Gradient-based learning applied to document recognition *Proc. IEEE* **86** 2278–324

Li Z, Liu F, Yang W, Peng S and Zhou J 2021 A survey of convolutional neural networks: analysis, applications, and prospects IEEE Trans. Neural Networks Learn. Syst.1–21

Long J, Shelhamer E and Darrell T 2015 Fully convolutional networks for semantic segmentation *2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Boston, MA)* 3431–40

Menze B H and Ur J A 2012 Mapping patterns of long-term settlement in Northern Mes-opotamia at a large scale *Proc. Natl Acad. Sci. USA* **109** E778–87

Minaee S, Boykov Y, Porikli F M, Plaza A J, Kahtarnavaz N and Terzopoulos D 2021 Image Segmentation using deep learning: a survey *IEEE transactions on pattern analysis and machine intelligence* **44** 3523–42

Moseley M 1975 . Chan chan: andean alternative of the preindustrial city *Science.* **187** 219–25

Mossun C, Palmer R and Campana S 2013 Flights into the Past. Aerial photography, photo interpretation and mapping for archaeology *Aerial Archaeological Research Group*

This is an on-line book Nielsen M A 2015 *Neural Networks and Deep Learning* (on-line book: Determination Press) (online available at (http://neuralnetworksanddeeplearning.com) last accessed on 24 Aug 2022.)

Odena A, Dumoulin V and Olah C 2016 Deconvolution and Checkerboard Artifacts *Distill* (https://doi.org/10.23915/distill)

Orengo H A, Conesa F C, Garcia-Molsosa A, Lobo A, Green A S, Madella M and Petrie C A 2020 Automated detection of archaeological mounds using machine-learning classification of multisensor and multitemporal satellite data *Proc. of the National Academy of Sciences of the United States of America* (https://doi.org/10.1073/pnas.2005583117)

Pawlowicz M L and Downum C E 2021 Applications of deep learning to decorated ceramic typology and classification: A case study using Tusayan White Ware from Northeast Arizona *J. Archaeolog. Sci.* **130** 105375

Protzen J and Morris C 2004 Los colores de tambo colorado: una reevaluación *BOLETÍN DE ARQUEOLOGÍA PUCP* **8** 267–76

Qi J, Yu Y, Wang L, Liu J and Wang Y 2017 An effective and efficient hierarchical K-means clustering algorithm *Int. J. Distrib. Sens. Netw.* **13** 8

Rayne L, Gatto M C, Abdulaati L, Al-Haddad M, Sterry M, Sheldrick N and Mattingly D 2020 Detecting change at

archaeological sites in north africa using open-source satellite imagery *Remote Sensing* **12** 3694

Reeves D M 1936 Aerial photography and archaeology *American Antiquity* **2** 102–7

Ruggles C L N 2015 Geoglyphs of the peruvian coast ed C Ruggles *Handbook of Archaeoastronomy and Ethnoastronomy.* (New York: Springer)

Russakovsky O *et al* 2015 ImageNet large scale visual recognition challenge *Int. J. Comput. Vision* **115** 211–52

Sagi T, Lehahn Y and Bar K 2020 Artificial intelligence for ocean science data integration: Current state, gaps, and way forward *Elementa.* (https://doi.org/10.1525/ELEMENTA.418)

Schmidhuber J 2015 Deep learning in neural networks: an overview *Neural Netw.* **61** 85–117

Smailes R 2011 Building chan chan: a project management perspective *Latin American Antiquity* **22** 37–63

Somrak M, Džeroski S and Kokalj Ž 2020 Learning to classify structures in ALS-derived visualizations of ancient maya settlements with CNN *Remote Sensing* **12** 2215

Soroush M, Mehrtash A, Khazraee E and Ur J A 2020 Deep learning in archaeological remote sensing: automated qanat detection in the kurdistan region of Iraq *Remote Sensing* **12** 500

Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S and Brisco B 2020 Google earth engine for geo-big data applications: a meta-analysis and systematic review *ISPRS J. Photogramm. Remote Sens.* **164** 152–70

Trier Ø D, Cowley D C and Waldeland A U 2018 Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland *Archaeological Prospection* **26** 2

Ulku I and Akagündüz E 2022 A survey on deep learning-based architectures for semantic segmentation on 2D images *Applied Artificial Intelligence* **36** 2032924

Van den Dries , M H 1998 Archaeology and the application of artificial intelligence: case-studies on use-wear analysis of prehistoric flint tools *Doctoral thesis, Faculty of Archaeology* (University of Leiden: Faculty of Archaeology, University of Leiden) 207

Verschoof-van der Vaart W B, Lambers K, Kowalczyk W and Bourgeois Q P J 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** 293

Verschoof-van der Vaart W B and Landauer J 2021 Using carcassonnet to automatically detect and trace hollow roads in LiDAR data from the Netherlands *Journal of Cultural Heritage* **47** 143–54

Waibel A 1987 *Phoneme Recognition Using Time-Delay Neural Networks. Meeting of the Institute of Electrical* (Tokyo, Japan: Information and Communication Engineers (IEICE))

Wu J 2017 *Introduction to Convolutional Neural Networks* (online available at (https://jasoncantarella.com/downloads/CNN.pdf) last accessed on 24 August, 2022)

Yu B, Chen F, Xu C, Wang L and Wang N 2021 Matrix SegNet: a practical deep learning framework for landslide mapping from images of different areas with different spatial resolutions *Remote Sensing* **13** 3158