pca

european journal of
postclassicalarchaeologies

volume 12/2022

SAP Società Archeologica s.r.l.

Mantova 2022

pca

Post-Classical Archaeologies (PCA) is an independent, international, peer-reviewed journal devoted to the communication of post-classical research. PCA publishes a variety of manuscript types, including original research, discussions and review articles. Topics of interest include all subjects that relate to the science and practice of archaeology, particularly multidisciplinary research which use specialist methodologies, such as zooarchaeology, paleobotany, archaeometallurgy, archaeometry, spatial analysis, as well as other experimental methodologies applied to the archaeology of post-classical Europe.

Submission of a manuscript implies that the work has not been published before, that it is not under consideration for publication elsewhere and that it has been approved by all co-authors. Authors must clear reproduction rights for any photos or illustration, credited to a third party that they wishe to use (including content found on the Internet). For more information about **ethics** (including plagiarism), copyright practices and guidelines please visit the website www.postclassical.it.

PCA is published once a year in May. Manuscripts should be submitted to **editor@postclassical.it** in accordance to the guidelines for contributors in the webpage http://www.postclassical.it.

Post-Classical Archaeologies' manuscript **review process** is rigorous and is intended to identify the strengths and weaknesses in each submitted manuscript, to determine which manuscripts are suitable for publication, and to work with the authors to improve their manuscript prior to publication.

This journal has the option to publish in **open access**. For more information on our open access policy please visit the website www.postclassical.it.

How to **quote**: please use "PCA" as abbreviation and "European Journal of Post-Classical Archaeologies" as full title.

Cover image: © Michel Royon / Wikimedia Commons.

"Post-Classical Archaeologies" is indexed in Scopus. It was approved on 2015-05-13 according to ERIH PLUS criteria for inclusion and indexed in Carhus+2018. Classified A by ANVUR (Agenzia Nazionale di Valutazione del sistema Universitario e della Ricerca).

DESIGN:
Paolo Vedovetto

PUBLISHER:
SAP Società Archeologica s.r.l.
Strada Fienili 39/a, 46020 Quingentole, Mantua, Italy
www.saplibri.it

Authorised by Mantua court no. 4/2011 of April 8, 2011

ISSN 2039-7895

pca
european journal of
postclassicalarchaeologies

volume 12/2022

CONTENTS **PAGES**

Juan A. Barceló, Florencia Del Castillo,
Deniz Kayikci, Borja Urbistondo*

Neural networks for archaeological classification and typology: an overview

1. Introduction

Archaeologists see in the present objects, whose shape, size, texture, material and mechanical properties can be described and even quantified. They do not 'see', however, bowls, amphorae, knives or swords. Archaeology is a problem solving discipline trying to solve the "what is this?" problem by logical inference.

Archaeological identification problems can be solved when we learn a rule for grouping the geometric, material or mechanical features of the observed objects by virtue of which they belong to sets of material effects of the same social action. We will understand what a pot, a house, a castle, a burial, a tool were (in the past) when we learn how to link the observed element to a *prototypical* pot, house, castle, burial, or tool. If such association is statistically reliable then all we know about the prototype – how it was produced or used in the past – can be transferred to the observed element from the archaeological excavation. The prototype is just a "virtual" – ideal – instance explicitly built from a theoretical model, which predicts features that can be perceived in the archaeological record. We may not be able to identify the causal process completely, but we can construct a good and useful approximation (Barker 2020; Longo 2021).

Consequently, one of the most fundamental notions for archaeological inference is that of similarity: the solutions to an archaeological problem group together things that are similar. Two entities are *similar* because they have many *properties* in common. According to this view (Medin 1989):

1. similarity between two entities increases as a function of the number of properties they share;

* J.A. Barceló, D. Kayikci, B. Urbistondo: Prehistoria Department, Quantitive Archaeology Lab, Universitat Autonoma de Barcelona. Corresponding author: *Juanantonio.barcelo@uab.cat*. F. Del Castillo: CENPAT-CONICET, Universidad Nacional de la Patagonia SJB, Argentina. On September 6th we received the sad news of the sudden and unexpected death of Florencia del Castillo. Among her last works is this paper, which she corrected on her last days. We offer our joint work in tribute to her memory.

2. properties can be treated as independent and additive;
3. the properties determining similarity are all roughly the same level of abstractness;
4. these similarities are sufficient to describe a conceptual structure: a concept would be then equivalent to a list of the properties shared by most of its instances.

Ideally, objects are similar *because* they were manufactured and/or used in the same way, in the same place and at roughly the same time for the same intended goal or function.

2. From clustering to learning

Clustering is the process of grouping input samples into similarity classes. Each object is represented as an ordered set (vector) of features. 'Similar' objects are those that have nearly the same values for different features (Rousseeuw, Kaufman 1990; Engel, van der Broeck 2001). *Clustering* is usually distinguished from *classification* on the basis that the first implies a mere objective partition of a set of data according to some rule – objects are enough similar between them – whereas the second is a form of categorization by which ideas and objects are recognized, differentiated and understood. Statistical classification is the set of methods used to identify new observation as belonging to a predefined category, on the basis of a training set of instances of that category (Hoehne 1980; Niknazar, Bourgault 2017). In archaeology, the term *typology* should be related with this last form of categorization, in the sense that *prototypes* should be ideal instances of explanatory categories (Adams 1988). Unfortunately, too many archaeological 'typologies' merely divide space and time into disjunctive classes – objects that have the same form because they were made in the same historical place and time, and lack more explanatory categories: functionality, manufacturing, etc.

The archaeological problems – "what is this", "how it was made", "why did they made such objects in that way?" – cannot be reduced to mere clustering. We need additional information, beyond the resemblance in shape, visual, material, mechanical properties and spatial and temporal provenance. We need to test whether the objects have these particular regularities *because* they were produced and/or used in the *same way*. We should move beyond statistical clustering into *concept learning.* Objects with the same triangular shape, for instance, cannot be functionally explained until we learn that 'arrow points' are triangular in shape to stick more efficiently into the animal's muscle and been able to kill the deer. We need to compare objects that prove to be similar enough to some *objects of reference*, that is, *prototypes* or known instances of a given concept.

What we need is a form of *supervised* learning, on the grounds that some known instances of a particular cause-effect relationship are used. In this paradigm, an agent learns to classify stimuli as members of contrastive categories through trial and error with corrective feedback (prior knowledge) (Kotsiantis *et al.* 2007; Cunningham *et al.* 2008; Jiang *et al.* 2020). Known examples of a particular *input-output mapping* may be experimental replications and/or ethno-historical data. In other words, the idea is to look for common features between positive examples of the causal relationship to be predicted, and common differences between its negative examples. This task is exactly like an example of a truth-function learning problem

1	1	0	1	1	→	1
1	0	0	0	0	→	0
0	1	1	1	0	→	1
1	1	0	0	1	→	0
0	0	0	0	0	→	?

Concept learning problems have the same form, except that target outputs are either 'yes' or 'no' (or 'true'=1 and 'false'=0). Inputs that map onto 'yes' are treated as positive examples of a particular concept. Inputs that map onto 'no' are treated as negative examples (i.e. counterexamples). The process of finding a solution to such a problem is naturally viewed as the process of calculating the *communalities* among positive examples. As such, it is a variation of the philosophical theories seeing *induction* as a process involving the exploitation of similarity. Positive instances of some predefined concept can be used as a reward when trying to *learn* the concept from which they are instances. Negative instances punish the learning process. The aim is to learn the action that will achieve the highest reward and the lowest punishment (feedback).

For best generalization we need an algorithm able to match the complexity of the hypothesis with the complexity of the function underlying the data. If the hypothesis is less complex than the function, the resulting model will be underfitted. If the hypothesis is too complex, or the data is not enough to constrain it, we will end up with a bad hypothesis. If there is noise, the resulting model will be overfitted because it is not only based on the underlying function but also on the noise in the data. In such a case, having more examples, or known instances helps but only to a certain point (Alpaydin 2004).

3. The advantage of neural networks

Since the early days of Numerical Taxonomy, mathematicians have produced hundreds of algorithms to 'measure' similarity and to build clusters of similar objects. Clarke (1968), Doran and Hodson (1975) and Baxter (1994) exemplify the

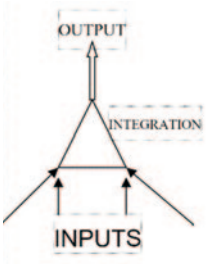

Fig. 1. Mathematical model of neuron.

first essays in archaeology. Since that time archaeologists have been powered by advances in statistical software packages to build 'classes' of similar objects (Read 2007; Hammer, Harper 2008; Van Pool, Leonard 2011; Barceló, Bogdanovic 2015; Carlson 2017; McCall 2018). The problem is that we cannot assume that clusters of *similar* objects are meaningful in themselves because they are not generalizable out of the specific limits of the input data used.

Neural networks are a special kind of algorithm able to learn non-linear and non-monotone input-output relationships. They have three main characteristics that have contributed to the wrong idea that they mimic the way the human brain operates: they work in a distributed and parallel way; they are also the result of adaptive process of learning. 'Distributed' means that calculations are decomposed into thousands of basic calculations between some basic computational units. 'Parallel' means that all those calculations are made simultaneously and all of them contribute to the final solution. 'Adaptive' means that they learn through reinforcement of rewards in successive 'evolutive' steps.

What is commonly refereed as an 'artificial neuron' is a simple input-output non-linear calculation, in which a number of discrete numeric inputs are summed and when this sum exceeds a predefined threshold, a numeric output is produced (fig. 1).

By linking many of these artificial neurons, we build a neural network. Not any linking topology is efficient, however. The most popular is a multilayer topology in which a subset of input units receive external numerical information, which is sent to a different subset of hidden layers of artificial neurons through weighted links. The input units only read external information and send it forward. All input units have links with all artificial neurons in the hidden layer, although not between them, in the sense that the individual contribution of each input unit is assumed to be independent. Numerical inputs arrive to each hidden unit 'transformed' – multiplied – by the particular weight each link has. Each hidden neuron receives information from all inputs, integrates it and execute a simple non-linear function: it sends a numerical signal (1 or 0) to a third layer of artificial neurons,

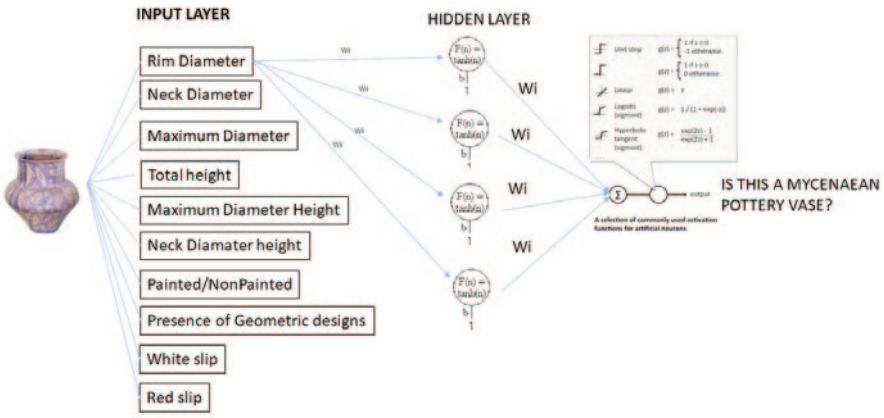

Fig. 2. A feed-forward multilayered neural network.

the output layer when the weighted integrated input signal exceeds a threshold. Each artificial neuron in the output layer receives as many inputs as units in the hidden layer, meaning that all hidden neurons are connected to all output neurons. This input is also transformed – multiplied – by the particular weight each link has. Output neurons activate at different degrees, depending on the way they integrate the numeric information they receive from the hidden layer.

Imagine we have some prehistoric pottery vases. Input information – shape, color, decoration, material properties – has been decomposed into single input units: rim, neck and maximum diameters, the height of those diameters, the presence/absence of painted decoration, the presence/absence of geometric designs, the presence/absence of white or red slip, etc. There is a single output: whether the vase is of Mycenaean origin or not. The idea is that when a vase is described using the input units, the algorithm will answer with the probability of whether can be of Mycenaean origin or not (fig. 2).

To be able to link the input with the output, link weights have to be 'learnt' from some training data. We need a big enough collection of pots with known provenience, some of them are known to be Mycenaean, and others are known to be from alternative cultural proveniences. We need more or less the same amount of Mycenaean and Non-Mycenaean pots to learn this particular relationship. The most known learning algorithm to solve this kind of problem was invented in 1986 by David Rummelhart, Geoffrey E. Hinton and others, and it is called Backpropagation (Rumelhart *et al.* 1995; Wythoff 1993; Hegazy *et al.* 1994; Kishore, Kaur

2012; Islam *et al.* 2019). At first, link weights are randomly initialized using random and very small numbers. When the first training instance is presented, its numerical input is transformed by random weights between the input and the hidden layers, this information arrives to hidden artificial neurons, which compute their non-linear summation and send a numeric output to the last layer. The output from hidden layers is also transformed by randomly initiated weights of links between the hidden and the output layers. This transformed signal arrives to the output layer and its artificial neurons are activated according to the intensity of the signal arrived and the particular activation function implemented. Obviously, because weights have been determined randomly during start-up, the final output is also the result of random decisions. The algorithm then compares what has been randomly calculated with what is known from the database of training exemplars, and calculates the difference. This error is sent backwards to modify weighted links, and the process begins with a new exemplar. The full training set is presented many times until the error between what is calculated and what is already known minimizes enough.

Once learned, a neural network can be used as associative memory, and therefore it assigns to new unseen inputs, the output that probably corresponds. It is a distributed representation of scientific knowledge because causal and other explanatory associations are stored throughout all the connections in the network, and because one set of connections can store several different associations. After learning, when using the network to categorize a new input, if the associative mechanism runs properly, then the pattern of activation in the output neurons will be the pattern that was originally associated with the cue pattern (Dawson 2004).

The learning process allows the model to determine, completely independently, which combinations of input nodes best predict the already known output values. Each hidden node in the artificial neural network serves to define specific connections within the input. The extent to which each input contributes information can vary between each hidden node, resulting in a model that can simultaneously consider input information in any combination to define different and complex relationships between input and output values. Similarly, an artificial neural network for multi-label classification predicts the probabilities of each output node independently of all other nodes, allowing information to be simultaneously classified into multiple categories, which distinguishes this classification algorithm from others, such as discriminant analysis, because the neural network does not compute the probabilities of outputs in conjunction with or in opposition to any other output (Lin 2021).

The advantages of this way of learning what archaeological elements may be – the concept to which they probably belong – are obvious. First of all, the relationship between input and output, between description and explanation can be

non-linear, when classical statistical classifers are limited by the intrinsic linearity of distinctions they can reproduce. Second, input information is distributed among minimal processing units, and there is no limit in the number, diversity and nature of them. Qualitative attributes can be added to quantitative ones. Parametric assumptions, normality or symmetry are not required. Any way to describe shape, size, visual appearance, materiality, mechanical properties, spatio-temporal coordinates can be integrated. Independence of inputs is assumed, but this is not an important requirement, because the learning algorithms will find a solution even in the case of strong dependency among inputs. The search for the best solution will be less efficient, but provided we have a big enough training set, adequately distributed between positive and negative instances of output categories, a solution will be found. The same is true for output units, that is, explanatory categories to be assigned to input information. In archaeology, explanatory categories are usually reduced to verbal labels: bowl, type 2A, knife, house, late Bronze Age, social elite, and the like. Using a neural network approach, explanations can be defined both by extension – the set of their positive instances in the training set – and by intention – in terms of the non-linear rule associating inputs with outputs. They also have a fuzzy nature. When actual input values enter the system, the entire memory fires at once, producing multiple outputs, and each output has an activation degree that can be expressed as a fuzzy number or even as a probabilistic value by including some constraining links between output neurons (Principe *et al.* 2000).

There is a drawback, however, the relationship cannot be expressed in simpler terms: it is a computational system that works well, but we do not 'see' how it works, because the association is distributed in the simultaneous activation of many minimal computing units. Due to this particularity, the processes, calculations and intrincaces that occur among the neurons which can't be seen, are often referred as the "blackbox" of neural networks. In any case, what in some sense can be considered as a limitation, in the other sense it is an advantage: because of the distributed and parallel nature of neural computing, even partial inputs can produce the activation of some outputs. If the object we want to categorize is broken, incomplete or post-depositionally altered, some input units will not be activated because of absence of information, or they will activate at a minor degree, depending on the ambiguity, presence of noise or uncertainty in descriptive information. Even in those circumstances, the remaining input signal flows through the weighted links and hidden neurons until the output. Obviously, this is not the same activation as when input is clear and complete, but we have some *potential* explanations, expressed in fuzzy terms, or even as a probability confidence level.

Neural networks began to be applied to solve archaeological classifications problems very early after the first publication of the classical backprop algo-

rithm. J.A. Barceló (1993; Barceló *et al.* 1996), J. Fulcher (1997), Gibson (1992, 1995), Bell and Croson 1998, Reeler (1999) are among the very first applications of this procedure in archaeology. Rock art, lithic tools, animal bones and provenience of archaeological objects based on composition were the applied domains. At that time, software implementation was still poorly efficient, and input information should be entered in form of discrete numeric variables or binary attributes. The difference with classical clustering and statistical classification and regression methods was not evident at the beginning, and archaeological applications were not very diverse nor abundant (Al-Nuaimy *et al.* 2000; Zupanek, Mlekuz 2001; Bell, Jantz 2002; Ducke 2003; Barceló, Pijoan-Lopez 2004; Bescoby *et al.* 2004; Lohse *et al.* 2004; Deravignone, Macchi 2006; Maaten 2006; Barceló 2009, 2010; Deravignone 2009; Ramazotti 2013; Spars 2013; Banerjee, Srivastava 2014; Alunni *et al.* 2015; Deravignone *et al.* 2015; Sharafi *et al.* 2016). Nowadays, building three layered feed forward neural networks and learning them using backpropagation algorithms and their variants are implemented in general use statistical software (SPSS, STATISTICA) and there is a wide variability of OpenSource and freeware implementations (NeuralNet in R, SimBrain.net, Aispace.org's neural networks package in java web, justNN, Sharky-neural networks, memBrain, Neocognitron, neuralDesigner, among many others). Users should only create a spreadsheet file in which columns define different inputs, and rows represent the value each training sample has on each input. In the same way, known outputs are necessary for the correct running of the backpropagation algorithm. Archaeological possible applications range from classification and typology, functional analysis and use-wear, spatial analysis in cemeteries, site prediction modelling, etc.

The advantages of neural networks over other methods of statistical classification and machine learning lie on their capacity to deal with huge quantities of data and with heavy distributed representations. The dream of archaeologists has always been to process images directly. This can be realized by using pictures as the network input (a lithic tool, a pottery, an engraving, an animal or human bone, etc.) and the output its best explanation (a chronology, a function, a provenience, etc.). Given the complexity of images, composed of pixels, we need to adjust the general structure of the neural network. It is an this point that 'deep' learning begins.

4. 'Deep' learning

The word 'deep' in 'deep learning' refers to the number of layers through which input data is transformed. When including extra layers, we can enable composite features from lower layers, potentially modeling complex data with

fewer units than a similarly performing shallow network. All those additional layers are also organized into levels, and each level learns to transform its input data into a slightly more abstract and composite representation. In an image recognition application, the raw input may be a matrix of pixels; the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode arrangements of edges; the third layer may encode an identified geometric decorative motive; and the fourth layer may recognize that the image represents a pottery vase of a particular kind, characteristic of a limited historical period in some geographical area. By using multiple layers it is more efficient to progressively extract higher-level features from the raw input. Layers are also permitted to be heterogeneous, for the sake of efficiency, trainability and understandability, whence the 'structured' part. This distributed form of knowledge representation eliminates feature engineering, by translating the data into compact intermediate representations akin to principal components, and derive layered structures that remove redundancy in representation.

Convolutional Neural Networks (CNN) are a special kind of 'deep' multilayer feed forward neural networks with additional layers ('convolutional') to extract features from input images. It also does so in such a way that position information of pixels is retained. Each neuron in the convolution layers receives input from only a restricted area of the previous layer called the neuron's *receptive field*. After passing thorugh the convolutional, the transformed input information pass thorugh consecutive pooling layers to reduce the dimensions of incoming feature maps. After passing through the convolutional layers, the image becomes abstracted to a feature map, also called an activation map, and pass its result to the next layer. The last layers provide the probabilities of a given class (Guo *et al.* 2016). This network also learn from given examples (generally a very large set of labelled images). During forward-propagation the input image is fed through the different layers with the current parameters (weights and bias) fixed. The output is compared to the ground truth labels (the same manually labelled image) and used to calculate the loss cost. Based on the loss cost, the gradients of each parameter are computed and used to update all parameters during back-propagation. All layers are then prepared for the next forward-propagation.

In 2012, Krizhevsky *et al.* (2012) trained a large, deep convolutional neural network (AlexNet) to classify 1.2 million high-resolution images into 1000 different classes. They used a global image database organized according to a conceptual hierarchy in which each node of the hierarchy is depicted by hundreds and thousands of images (ImageNet: https://image-net.org/update-mar-11-2021.php. In its time, it was the largest image database ever built). The first neural network able to learn this very complex image recognition task had 60 million parameters and 650,000 neurons, and was built in terms of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connect-

ed layers, with a final 1000 neurons in the output layer where classification was produced. The network obtained a huge classification success, and opened the current popularity of convolutional neural network models, which are now ubiquitous for computer vision tasks like image classification, object detection, image recognition, etc.

After AlexNet, CNN architectures have increased its depth – number of hidden layers – to increase image recognition success. Hence, more layers seem to promise better performance. This is the case of VGG Net, with 16 to 19 layers (Vedaldi, Zisserman 2016) and GoogLeNet, a 22-layer deep convolutional neural network, using 224 x 224 input images (Szegedy *et al.* 2014).

Nevertheless, increasing network depth does not work by simply stacking layers together. Deep networks are hard to train because of the notorious vanishing gradient problem – as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient infinitively small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly. To solve these problems, the Residual Neural Network model (ResNet) was introduced in 2015, with the core idea that some neurons should not always connect with neurons contiguous layers. This network topology implements *shortcuts* to jump over some layers based on skip weights learnt using an additional weight matrix (He *et al.* 2016).

All these architectures of deep neural networks are downloadable by any user, and ready to be used for any complex classification task. Some models are implemented in MatLab platform[1], for Keras[2], or Torch[3] among many others. In addition to use an already tested architecture, we can 'pre-train' the network with some universal image repository like ImageNet. The idea is to begin learning not with random weights between layers, but with a network able to carry out a global task. If you want the network be able to classify between kinds of lithic knives, it would be very positive to begin with a network able to distinguish knives from arrow points and scrapers. Then, our network will begin learning a more specialized classificatory task once the weights necessary for distinguishing basic classes have already been learnt. In a way, it is a form of fine-tuning a pre-existing classificatory system (also known as *transfer learning*). By importing a pre-trained convolutional neural network as AlexNet or VGG19, the feature recognition part of the network can be shortcutted effectively, allowing most of

[1] https://es.mathworks.com/help/deeplearning/ref/resnet50.html or https://es.mathworks.com/help/deeplearning/ref/googlenet.html).

[2] https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c

[3] https://github.com/Lornatang/VGGNet-PyTorch, https://blog.paperspace.com/how-to-implement-a-yolo-object-detector-in-pytorch/

the effort done by the researchers to be focused on the design of the fully-connected part specifically programed to fulfill its task without even having to modify the previous convolutional layers if needed (Zabir *et al.* 2018, Conde, Turgutlu 2021).

5. Deep learning in archaeology

Quantitative based classification of pottery vases has been traditionally based on mineralogical and petrographical data. Bell and Crosson (1998), Baxter (2006), Ramil *et al.* (2008), Kvascev *et al.* (2012), Barone *et al.* (2019) have explored the use of simple three layer neural network applications for classifying pottery vases based on its mineralogical and petrological composition. Aprile *et al.* (2014) have followed a more complex approach using thin slide microscopic images instead of compositional data vectors. These authors used both plane and cross polarised light images acquired via a digital camera connected to optical microscopy in transmitted light to classify mineral inclusions (quartz, calcareous aggregates and secondary porosity) from Holocene potsherds (8900-4200 years BP) from Central Sahara.

Classifying pottery vases in terms of their visual appearance – shape, form, surface finish, decorative patterns – has been a classical expert task, relying on the particular and subjective experience of well trained individuals. The low efficiency of early neural networks, and the classical success of multidimensional statistics (Read 2007) prevented the application of machine learning procedures. Nowadays, the idea is to use images as input to distinguish between functionally, stylistically or historically different pottery productions using some kind of deep neural network. According to Tyukin *et al.* (2018) there are two possible levels on which automated computer recognition can be performed. One approach would be based on building detectors which will identify the object of interest – e.g. a complete vessel – in a 2D picture. In this case, we should create a separate detector for each artefact. If one of the detectors finds something in a picture of more than one vessel, we assume that this picture contains an artefact of a certain previously identified type. Since we know which detector exactly fired, we know the type of the artefact and the recognition is completed by this step. This would be a low-level recognition. The opposite approach is to build a detector that can 'find' all types of objects of interest in a picture of more than one artefact. In this case, when the detector fires, we still do not know exactly which artefact was found. We should send each candidate to a second computational stage, where a specially trained neural network would identify which type of object it is, according to existing classifications or typologies. In contrast to the low-level approach, this one would be a high-level one.

The ArchAIDE system (Gualandi *et al.* 2021) is one of the most advanced applications of deep neural networks to classify pottery vases based on shape geometric factors as: Outer profile, Inner profile, Handle outer profile (if present), Handle inner profile (if present), Handle section (if present), Rim point: the top point in the profile, Base point: the bottom point in the profile, Scale factor: the scaling value to bring all features to real scale.

In the same way, Cintas *el al.* (2020) have trained a convolutional neural network using 1133 binarized images of pottery profiles, already classified by domain experts into 11 different classes of wheel made pottery from the upper valley of the Guadalquivir River (Spain) and dated to Iberian period; nine of them correspond to closed shapes, and the two remaining correspond to open shapes. From a digitized profile, the system can determine with high accuracy and precision the category to which the ceramic vessel belongs. This representation allows the identification of which shapes of a given vessel is nearer or farther. Images of 64 × 64 pixels were feed to the input layer, whereas the output layer is composed of 11 output units for the classes and types to be predicted. The resulting model is able to provide classification on profile images automatically, with an accuracy mean score of 0.9013 when compared to classificatory decisions by domain experts based on morphological criteria, taking into account the presence or absence of certain parts, such as lip, neck, body, base and handles, and the ratios between their corresponding sizes. The same authors have expanded the project using an alternative neural network architecture (Navarro *et al.* 2021). They have used a residual neural network for automatically extracting learned features and enhance the previous classifier of Iberian ceramic pottery.

Alternatively, Dia *et al.* (2021) perform analysis and classification of tomographic images of ceramic fragments.

Pottery decorative patterns can be classified in a similar way, using pictures and/or surface models of the vessel. Chetouani *et al.* (2020) have studied early medieval pottery sherds engraved with repeated decorative patterns using a carved wooden wheel, about 1 mm deep and 1.5 to 3 cm wide, depending on the dimensions of the wooden cylinders used by the potters. The most common patterns notched by potters on wooden cylinders included sticks, squares, chevrons and diamonds in one or several lines. The inputs used in this work were greyscale relief maps instead of binary images. The authors experimented with pre-trained AlexNet, VGG16, VGG19 and ResNet50 to fine tuning the outputs with their dataset. The typical output layer of those universal image recognizers was modified and adapted to the number of solutions for the case at hand (four). The system obtained a classification successful rate of 95.23%.

Moscoso Thompson *et al.* (2021) is another example of pottery decoration classification using convolutional neural networks. A ResNet pre-trained convo-

lutional neural network was re-trained using images of pottery fragments decorated with relief patterns. The system has allowed segmenting reliefs of repeated motifs and decorations on the surface of an archaeological artefact using supervised learning to obtain the classification of the individual pixels of images extrapolated from 3D models.

Pawlowicz and Downum (2021) study a specific kind of ancient painted pottery from the American Southwest. They have collected an initial set of 3064 photographs of a particular type of sherds, which were stylistically classified by four experts in seven major types, as well as two less common types, and categories of 'Other' and 'Indeterminate'. The system also gives the option to select intermediate/mixed types comprised of temporally and stylistically adjacent types. A slightly modified versions of the VGG16 and ResNet50 convolutional neural networks were chosen given the availability of pre-trained model weights. Input images were standardized and resized to 224 × 224 pixels, and converted to grayscale, to remove any effects of sherd discoloration from soiling or weathering, or variations in white balance when photographing the sherds. Results showed far geater agreement than human experts.

Before leaving deep learning applications in the pottery analysis domain, we should mention Reese's (2021) predictions of site occupation time spans based on the identified type of dominant pottery. The analytical focus is put explicitly on the temporal relationship between ceramic assemblages and dates of occupation. The training dataset includes a total of 118 sites for which both ceramic type frequencies and tree-ring dates are available. Each row in the input table represents one site, and each column represents the total frequencies of each ceramic type identified at the corresponding site. Normalized ceramic data are used instead of raw counts because these re-scaled values are directly comparable between sites regardless of the total number of ceramics identified. The number of nodes in the input layer of the neural network is 14 – each representing a diagnostic ceramic type – and the number of nodes in the output layer is 851 – one for each year from AD 450-1300. After some experimentation, the optimal number of nodes within the hidden layer is 18, and the optimal smoothing window is 21 – meaning the model produced the most accurate results when the smoothing window applied to the predictions was similar to the average use-life of a small site residential structure across the entire study period (average use-life from AD 450–1300 was approximately 19 years). This combination creates an artificial neural network with a minimum 72.4% accuracy, and an average annual predictive accuracy of 93.5% for the AD 450-1300 study period. Once the optimal parameters were identified, the trained artificial neural network was applied to sites across the central Mesa Verde region with a residential component. Of the 7600 recorded sites with a residential component, approximately 60% have a corresponding ceramic tally reporting frequencies of ceramic types, approximately

20% have a ceramic tally that does not report any diagnostic wares or only reports the presence/absence of various ceramic types, and approximately 20% have no reported ceramic information. This means occupation can be predicted for 60% of sites in the central Mesa Verde region using the trained artificial neural networks, but other steps must to be taken to account for the remaining 40% of sites. Results from the artificial neural network analysis, extrapolated to unsurveyed areas and smoothed by life-expectancy, are used to estimate annual population of the central Mesa Verde region from AD 450-1300.

The shape of lithic tools can also be classified using supervised neural networks, as shown by Nash and Prewitt (2016) and Grove and Blinkhorn (2020). These authors have used a quantitative analytical framework based on the use of neural networks to distinguish the changing technologies between the Middle and Late Stone Ages in East Africa. The trained network correctly classified more than 94% of them and identified 7 basic technologies that significantly discriminated between classes (see also Resler *et al.* 2021).

The analysis of cut marks on the surface of bones and prehistoric or ancient tools is another relevant image-based archaeological task where neural networs can be applied for classificatory purposes. Byeon *et al.* (2019) and Domínguez-Rodrigo et al. (2021) have defined deep convolutional neural networks (DCNN) to recognize marks with accuracy that far excedes that of human experts (91% over 63%). Zotkina and Kovalev (2019) have explored traces and marks left on the surface of prehistoric rock-art panels, trying to distinguish two types of tool marks: metal and lithic, and a natural surface as well. Input data adopts the form of 3D local descriptors using 3D key points – points in a point cloud that are stable, distinctive, and can be identified using a well defined detection criterion. In a preliminary step, key points with rich information contents are first identified and their associated scales (spatial extents) are then determined. In a second, feature description phase, local geometric information around a key point is extracted and stored in a high dimensional vector (i.e., feature descriptor). Neural network training is based on a set of descriptors of key points taken from a point cloud of a certain shape. Then, upon receipt of a test set of descriptors obtained from the test point cloud, the neural network can determine the presence of points forming such 3D compositions.

Remote sensing is another quintesentially visual analysis task. Computer driven image analysis offers a way to rapidly extract archaeological information from such very high resolution images, and can even be used to detect objects of interest that are not visible to the naked eye. Pixel-based classification approaches use the smallest entity within an image, the picture element (or pixel), in order to extract the feature information in relation to one or more predefined classes. Each pixel can be addressed by the *x* and *y* coordinates of the two dimensional image space, and can be attributed with one or more values derived

from the *z* axis, which is formed by the multiple layers of the dataset. In all cases the attributes of each pixel (in most cases, one or more spectral values) function as the basis for the classification, in which an assembly of pixels with similar attribute values form an arrangement describing, for example, an archaeological element to be classified. The investigator uses known simple areas based on prior knowledge to 'train' the classifier to recognize areas in the image with similar valued pixels corresponding to particular elements – burials, sites, agricultural structures, defensive structures, etc. – differentiated from their neighborhood. In contrast to the pixel-based approach, which starts from the basic entities of an image, an object-based analysis uses the entire image or data set and breaks it down into meaningful segments. One essential factor that distinguishes object-based approaches from the more traditional pixel oriented approaches is the integration of the image content into the classification procedure. The aim of this approach is to subdivide the image into homogeneous segments that describe the target features (pit, burial mound, etc.) as correctly as possible. The advantage of this approach lies on the superior reliability of image segments to closely represent real-world objects (Sevara *et al.* 2016; Ball *et al.* 2017; Fiorucci *et al.* 2022).

Sharafi *et al.* (2016), Toumazet *et al.* (2017), Pasquet *et al.* (2017), Traviglia and Torsello (2017), Trier *et al.* (2018), Engel *et al.* (2019), Monna *et al.* (2020), Soroush *et al.* (2020), Berganzo-Besga *et al.* (2021), Guyot *et al.* (2021), Suh *et al.* (2021), are among the best examples of using neural networks and supervised learning algorithms to detect archaeological elements on aerial and satellite images. Nowadays, Light Detection And Ranging (LIDAR) facilitates the identification of spatial features categorized as archaeological elements that were hitherto difficult to investigate, due to forest and other vegetation cover. LIDAR data can be presented to the convolutional Neural Network input layer in form of Shaded Relief Maps images. After extracting geometric and spatial features, archaeological entities like ancient field delimitation or ancient roads can be identified (Verschoof, Lambers 2019; Verschoof *et al.* 2020; Olivier, Verschoof 2021).

Exemples discussed so far refer to images, that is 2D arrays of numbers. Deep learning methods can also be used to classify 3D geometric models of objects or buildings. The problem is much more complex than the already difficult problem of classifying high resolution images, and require more complex architectures (Maturana, Scherer 2015; Zhi *et al.* 2017; Qi *et al.* 2021). To pre-train such networks, and similar to ImageNet for 2D pictures of objects, the Princeton's ModelNet project provides a comprehensive collection of 3D CAD models of the most common object categories in the world (https://modelnet.cs.princeton.edu/). ShapeNet is another large scale repository for 3D CAD models (https://shapenet.org/).

3D object identification is still underdeveloped in archaeology. However, some innovative work has been published on the use of neural network supervised learning for reconstructing objects from fragments. Hermoza and Sipiran (2018) show how these methods of machine learning can be applied to predict the geometry of missing parts in damaged objects. A neural network can be trained to complete what has been discovered incomplete. The goal of the method is to take a 3D scan of a fractured object as input and predict the complete object as output. Colmenero-Fernández and Feito (2021) also propose using neural networks for three-dimensional volumetric reconstruction. Their approach is based on differentiating the area of interest of the profile using object-oriented neural networks, graphic normalization, and the 3D volumetric model using object-oriented neural networks. It includes model computing analysis algorithms that are trained to prevent loss of information in only one stage of detection obtained from pooling and masking operations. As a result, huge graphs of the profile section and three-dimensional models of archaeological sketches are obtained.

A similar but simpler approach has been proposed by Rasheed and Nordin, (2020). They have built a three layers feed forward backpropagation network to link broken pottery vases fragments with other fragments that may match. Similar approaches are those by Grilli *et al.* (2019) and Ostertag and Beurton-Aimar (2020).

6. Conclusion

Archaeological classification is not the only way of archaeological inference. In this paper we have not made any reference to interpolation – regression methods are an exemple of this kind of inference models –, which are based on abductive and deductive reasoning instead of the inductive basis of classification. Neural networks are not the only technology for induction and supervised learning. Beyond classical statistics, other inductive algorithms are being explored for archaeological classification, like Rule Induction, Genetic Algorithms, Bayesian networks, Support Vector Machines, Random Forests, among many others. In some cases, those algorithms offer better results that neural networks – i.e., more accurate results, with less classification result when compared with prior knowledge.

As discussed in this paper, neural networks, both simple and deep ones, have their proper role in archaeology, especially when input and output, when description and explanation, cannot be reduced to simple terms or discrete attributes. This is specifically the case for images. If archaeology is a quintessentially visual discipline (Shelley 1996; James 2015), then we need deep machine

learning methods to process complex and complicated visual information. Obviously, complex non-visual archaeological data and explanations can also be processed using neural networks – material, archaeometric, spatial data.

The suitability of neural networks is not only due to their efficiency and 'success' in the task of classifying complex inputs into pre-defined classes. In a neural network, information and knowledge, that is, observable features and explanatory concepts appear to be distributed across a very large population of individualized regions in a multidimensional space, rather than as sentential or propositional contents. Instead of having an enumerated set of responses particularizing the way in which an archaeologist can solve problems, a mathematical function will transform the sensory input into an explanatory inference, in such a way that inputs are mapped into a potentially very large set of responses (Churchland, Sejnowski 1990; Browne, Sun 2001; Levine, Aparicio 2013; Barceló 2009, 2015).

The approach exposed here challenges the received conception of an explanation as a basically invariant structure. The classical approach implied that the kind of knowledge we have when we grasp a concept such as 'bowl', 'amphorae', '*terra sigillata',* 'burial', 'activity area', 'Bronze Age', 'Late Republican Period', etc. is knowledge of some necessary and sufficient defining conditions allowing us to test for membership of the category by testing for the presence/absence of the features cited. The same is usual for any other archaeological concept. This classical view, however, looks to be undermined in some important respects (Barceló 2009). A connectionist *depictive* representation seems to be inherently more efficient at encoding contextual dependencies and at handling interactions among constituents (Markman 2013). Additionally, for many of our concepts there are not any necessary and sufficient conditions to be discovered. Finally, our judgments of category membership (whether or not a concept is applicable to a given case) are scalar.

The ability to represent both prototypical information and information about specific instances is the basis of the neurocomputing success. We can activate two properties, and discover which outputs are most likely to fit that scenario. The network will initially produce higher activations in the output units which posses any of these properties, with those sharing both properties getting the highest activations. The units for the most widely shared properties also become more active than others. Thus the network not only identifies which outputs shared the initial pair of properties, but what their other properties were likely to be, and which among those not possessing the initial pair show the best fit with those who did satisfy the initial pair of properties.

This is an important property of the parallel distributed models of computation, but the importance of this property increases when we realize that the model can average several patterns in the same composite memory trace. Thus, one

network can be trained to exhibit behavior appropriate to knowledge of a number of distinct prototypes, such as an arrow point, a stone knife, and an axe. Interestingly, if the input is indeterminate between a stone knife and a stone scraper, the neural network will generate an overall pattern, as if it had an idea not just of knives and scrapers but also on stone tools. We see then that the talent of the system is used to generate a *typical* set of properties associated with some description, even though all the system directly knows about are individuals, none of whom need to be a perfectly typical instantiation of the description in question (Barceló 2009, 2015).

We may call this approach to archaeological reasoning *category-based* because explanatory elements are accessed through a process of categorization. It implies that the input reaching the successive layers of a neural network is sorted out into discrete, distinct categories whose members somehow come to resemble one another more than they resemble members of other categories. We have already seen that the categorization (or pattern recognition) approach proposes that two operations are involved. First, the system classifies an object as being a member of a large number of known categories according to its input properties. Second, this identification allows access to a large body of stored information about this type of object, including its function and various forms of expectations about its future behaviors. This two-step schema has the advantage that any explanatory property can be associated with any object, because the relation between the form of an object and the information stored about its function, history, and use can be purely arbitrary, owing to its mediation by the process of categorization. That means that the responses of the neural network to the incoming input are not dependent of any particular attribute of the input. Rather, the solution to the archaeological problem will make sense only when considered as one component in a causal chain that generates responses entirely according to the probability distribution of the past significance of the same or related input. The answer provided by the intelligent machine exemplifies not the stimulus or its sources as such, but the accumulated interactions with all the possible sources of the same or similar stimuli in proportion to the frequency with which they have been experienced.

Aknowledgements

This paper has been possible thanks to funding from different Agencies: Spanish Ministry for Science and Innovation (Grant PID2019-109254GB-C21), European Research Executive Agency (Project 101066298), Agencia de Gestión de Ayudas Universitarias y de Investigación from Generalitat de Catalunya (2017SGR243), European Union Erasmus+ KA220-HED (-IT02-KA220-HED-000031190).

Abstract

Although some of the now popular deep learning techniques and technologies have a relatively long history (nearly 30 years), it has been in the last 5 years when these applications have reached mainstream Archaeology, cultural heritage and museum studies. There is a new conscience of data processing in Archaeology, although the nature of these data hardly arrives to the usual label 'big data', and it has opened the methodological toolbox at use, especially in domains like reconstruction, remote sensing, object recognition, typological analysis, and collection management and visitor studies. In this paper, the history and current applications of neural networks and related methods of machine learning in archaeology, cultural heritage and museum studies are investigated. The necessary theoretical background on induction and learning is provided to understand the possibilities and limitations of computational techniques.
Keywords: archaeology, cultural heritage, artificial intelligence, neural networks, deep learning.

Nonostante alcune tecniche e tecnologie di deep learning *abbiano una storia relativamente lunga (quasi 30 anni), solo negli ultimi 5 anni queste sono state applicate più direttamente all'archeologia, al patrimonio culturale e agli studi museali. C'è oggi una nuova coscienza del processamento dati in archeologia (anche se la natura di questi dati a stento raggiunge l'usuale etichetta "big data") che ha aperto una "scatola degli attrezzi" metodologica, specialmente nei campi come ricostruzione,* remote sensing*,* object recognition*, analisi tipologica, gestione delle collezioni, studi dei visitatori. In questo articolo vengono esplorate la storia e le attuali applicazioni delle reti neurali e degli associati metodi di* machine learning *in archeologia e studi museali. Il necessario* background *teoretico viene fornito per comprendere le possibilità e le limitazioni delle tecniche computazionali.*
Parole chiave: *archeologia, patrimonio culturale, intelligenza artificiale, reti neurali,* deep learning*.*

References

A. Aamodt, E. Plaza 1994, *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*, "Artificial Intelligence Communications", 7(1) pp. 39-52.

W.Y. Adams 1988, *Archaeological classification: theory versus practice,* "Antiquity", 62(234), pp. 40-56.

D.W. Aha 1998, *The omnipresence of case-based reasoning in science and application,* "Knowledge-based systems", 11(5-6), pp. 261-273.

I. Aizenberg, N.N. Aizenberg, J.P.L. Vandewalle 2000, *Multi-Valued and Universal Binary Neurons: Theory, Learning and Applications*, Boston.

W. Al-Nuaimy, Y. Huang, M. Nakhkash, M.T.C. Fang, V.T. Nguyen, A. Eriksen 2000*, Automatic Detection Of Buried Utilities And Solid Objects With GPR Using Neural Networks And Pattern Recognition.* "Journal of Applied Geophysics", 43, pp. 157-165.

E. Alpaydin 2004, *Introduction to Machine Learning*, Cambridge (MA).

V. Alunni, P. du Jardin, L. Nogueira, Buchet, G. Quatrehomme 2015, *Comparing discriminant analysis and neural network for the determination of sex using femur head measurements,* "Forensic Science International", 253, pp. 81-87.

R. Anand, T. Shanthi, M.S. Nithish, S. Lakshman 2020, *Face recognition and classification using GoogleNET architecture*, "Soft Computing for Problem Solving", pp. 261-269.

A. Aprile, G. Castellano, G. Eramo 2014, *Combining image analysis and modular neural networks for classification of mineral inclusions and pores in archaeological potsherds.* "Journal of Archaeological Science", 50, pp. 262-272.

J.E. Ball, D.T. Anderson, C. Seng Chan 2017, *Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community*, "Journal of Applied Remote Sensing", 11(4), pp. 1-54.

P. Ballester, R.M. Araujo 2016 (February), *On the performance of GoogLeNet and AlexNet applied to sketches*, in *Thirtieth AAAI conference on artificial intelligence*, https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12278/11712.

R. Banerjee, P.K. Srivastava 2014, *Remote sensing based identification of painted rock shelter sites: appraisal using advanced wide field sensor, neural network and field observations*, in P.K. Srivastava (ed), *Remote sensing applications in environmental research*, Cham, pp. 195-212.

J.A. Barceló 1993, *Back-propagation algorithms to compute similarity relationships among archaelogical artifacts*, in J.W. Wilcock, K. Lockyear (eds), *Computer Applications and Quantitative Methods in Archaeology 1993*, Oxford, pp. 165-176.

J.A. Barceló 2009, *Computational Intelligence in Archaeology*, Hershey-NY.

J.A. Barceló 2010, *Visual analysis in archaeology. An artificial intelligence approach*, in J.A. Barceló, A. Elwa (eds), *Morphometrics for Nonmorphometricians*, Berlin-Heidelberg, pp. 93-156.

J.A. Barceló 2015, *Measuring, Counting and Explaining: An Introduction to Mathematics in Archaeology*, in Barceló, Bogdanovic 2015, pp. 3-64.

J.A. Barceló, I. Bogdanovic (eds) 2015, *Mathematics and Archaeology*, Boca Raton.

J.A. Barceló, J. Pijoan-Lopez 2004, *Cutting Or Scrapping? Using Neural Networks To Distinguish Kinematics In Use Wear Analysis,* in *Enter the Past. The E-way into the Four Dimensions of Culture Heritage,* Oxford, pp. 427-431.

J.A. Barceló, A. Vila, J. Gibaja 1996, *An Application of Neural Networks to Use-Wear Analysis. Some Preliminary Results,* in K. Lockyear, T.J.T. Sly, V. Mihailescu-Birliba (eds), *Computer Applications and Quantitative Methods in Archaeology,* Oxford, pp. 63-70.

S.F. Barker 2020, *Induction and Hypothesis: A Study of the Logic of Confirmation,* Ithaca-New York.

G. Barone, P. Mazzoleni, G.V. Spagnolo, S. Raneri 2019, *Artificial neural network for the provenance study of archaeological ceramics using clay sediment database*, "Journal of Cultural Heritage", 38, pp. 147-157.

M.J. Baxter 2006, *A review of supervised and unsupervised pattern recognition in archaeometry*, "Archaeometry", 48(4), pp. 671-694.

M. Baxter 1994, *Exploratory Multivariate Analysis in Archaeology*, Edinburgh.

S. Bell, C. Croson 1998, *Artificial Neural Networks as a tool for Archaeological Data Analysis*, "Archeometry", 40(1), pp. 139-151.

S. Bell, R. Jantz 2002, *Neural Network Classification of Skeletal Remains*, in G. Burenhult (ed), *Archeological Informatics: Pushing the Envelope*, Oxford, pp. 205-212.

Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle 2007, *Greedy layer-wise training of deep networks*, "Advances in neural information processing systems", 19, pp. 153-160.

H. Benhabiles, H. Tabia 2016, *Convolutional neural network for pottery retrieval*, "Journal of Electronic Imaging", 26(1). DOI: https://doi.org/10.1117/1.JEI.26.1.011005.

I. Berganzo-Besga, H.A. Orengo, F. Lumbreras, M. Carrero-Pazos, J. Fonte, B. Vilas-Estévez 2021, *Hybrid MSRM-based deep learning and multitemporal sentinel 2-based machine learning algorithm detects near 10k archaeological tumuli in North-Western Iberia*, "Remote Sensing", 13(20). DOI: https://doi.org/10.3390/rs13204181.

D.C. Bescoby, G.C. Cawley, P.N. Chroston 2004, *Enhanced Interpretation Of Magnetic Survey Data Using Artificial Neural Networks: A Case Study From Butrint, Southern Albania*, "Archaeological Prospection", 11(4), pp. 189-199.

A. Browne, R. Sun 2001, *Connectionist inference models*, "Neural Networks", 14(10), pp. 1331-1355.

W. Byeon, M. Domínguez-Rodrigo, G. Arampatzis, E. Baquedano, J. Yravedra, M.A. Maté-González, P. Koumoutsakos 2019, *Automated identification and deep classification of cut marks on bones and its paleoanthropological implications,* "Journal of Computational Science", 32, pp. 36-43.

D.L. Carlson 2017, *Quantitative Methods in Archaeology Using R,* Cambridge.

G. Caspari, P. Crespo 2019, *Convolutional neural networks for archaeological site detection – Finding "princely" tombs,* "Journal of Archaeological Science", 110, 104998.

R. Cheng 2020, *A survey: Comparison between Convolutional Neural Network and YOLO in image identification*, "Journal of Physics: Conference Series", 1453(1), 012139.

A. Chetouani, S. Treuillet, M. Exbrayat, S. Jesset 2020, *Classification of engraved pottery sherds mixing deep-learning features by compact bilinear pooling*, "Pattern Recognition Letters", 131, pp. 1-7.

P.S. Churchland, T.J. Sejnowski 1990, *Neural representation and neural computation*, "Philosophical Perspectives", 4, pp. 343-382.

C. Cintas, M. Lucena, J.M. Fuertes, C. Delrieux, P. Navarro, R. González-José, M. Molinos 2020, *Automatic feature extraction and classification of Iberian ceramics based on deep convolutional networks*, "Journal of Cultural Heritage", 41, pp. 106-112.

D.L. Clarke 1968, *Analytical Archaeology*, London.

A. Colmenero-Fernández, F. Feito 2021, *Image processing for graphic normalization of the ceramic profile in archaeological sketches making use of deep neuronal net (DNN),* "Digital Applications in Archaeology and Cultural Heritage", 22, e00196.

M.V. Conde, K. Turgutlu 2021, *CLIP-Art: contrastive pre-training for fine-grained art classification*, "Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition", pp. 3956-3960.

P. Cunningham, M. Cord, S.J. Delany 2008, *Supervised learning*, in M. Cord, P. Cunningham (eds), *Machine learning techniques for multimedia*, Berlin, pp. 21-49.

M.R.W. Dawson 2004, *Minds and Machines. Connectionism and Psychological Modeling*, London.

R. Dechter 1986, *Learning While Searching in Constraint-Satisfaction-Problems*, in T. Kehler (ed), *Proceedings AAAI*, Philadelphia-PA, pp. 178-185. Online in: https://www.aaai.org/Papers/AAAI/1986/AAAI86-029.pdf?spm=5176.100239.blogcont215226.26.uwOrvt&file=AAAI86-029.pdf&source=post_page.

L. Deravignone 2009, *Intelligenza artificiale: sviluppo metodologico e applicativo nell'archeologia del territorio*, PhD Thesis, Università degli Studi di Siena.

L. Deravignone, G. Macchi 2006, *Artificial Neural Networks in Archaeology*, "Archeologia e Calcolatori", 17, pp. 121-136.

L. Deravignone, H.P. Blankholm, G. Pizziolo 2015, *Predictive modeling and artificial neural networks: from model to survey,* in Barceló, Bogdanovic 2015, pp. 335-351.

K. Dia, V. Coli, L. Blanc-Féraud, J. Leblond, L. Gomart, D. Binder 2021, *Applications of Learning Methods to Imaging Issues in Archaeology, Regarding Ancient Ceramic Manufacturing*, in *2nd International Conference on Deep Learning Theory and Applications-DeLTA* (Online streaming), pp. 109-116, DOI: https://dx.doi.org/10.5220/0010519101090116.

M. Domínguez-Rodrigo, E. Baquedano, L. Varela, P.S. Tambusso, M.J. Melián, R.A. Fariña 2021, *Deep classification of cutmarks on bones from Arroyo del Vizcaíno (Uruguay),* "Proceedings of the Royal Society B", 288, DOI: https://doi.org/10.1098/rspb.2021.0711.

J.E. Doran, F.R. Hodson 1975, *Mathematics and Computers in Archaeology*, Harvard.

B. Ducke 2003, *Archaeological Predictive Modelling In Intelligent Network Structures*, in M. Doerr, A. Sarris (eds), *The Digital Heritage of Archaeology. Computer Applications and Quantitative Methods in Archaeology*, Heraklion, pp. 267-273.

A. Engel, C. Van den Broeck 2001, *Statistical mechanics of learning*, Cambridge.

C. Engel, P. Mangiafico, J. Issavi, D. Lukas 2019, *Computer vision and image recognition in archaeology,* in *Proceedings of the Conference on Artificial Intelligence for Data Discovery and Reuse*, Pittsburgh-PA, pp. 1-4. DOI: https://doi.org/10.1145/3359115.3359117.

B.C. Falkenheimer 1990, *A Unified Approach to Explanation and Theory Formation*, in J. Shrager, P. Langley, M. Kaufmann (eds), *Computational Models of Scientific Discovery and Theory Formation*, San Francisco, pp. 156-196.

M. Fiorucci, W.B. Verschoof-van der Vaart, P. Soleni, B. Le Saux, A. Traviglia 2022, *Deep Learning for Archaeological Object Detection on LiDAR: New Evaluation Measures and Insights,* "Remote Sensing", 14(7), 1694.

J. Fulcher 1997, *Neural Networks for Archaeological Provenancing*, in E. Fiesler, R. Beale (eds), *Handbook of Neural Computation*, Oxford.

R. Garozzo, C. Santagati, C. Spampinato, G. Vecchio 2021, *Knowledge-based generative adversarial networks for scene understanding in Cultural Heritage,* "Journal of Archaeological Science: Reports", 35, 102736.

P. Gibson 1993, *The Potentials of Hybrid Neural Network Models for Archaeological Ageing and Interpretation*, in J. Andresen, T. Madsen, I. Scollar (eds), *Computing the Past. Computer Applications and Quantitative Methods in Archaeology (CAA92),* Aarhus, pp. 263-272.

P. Gibson 1995, *An Archaeofaunal Ageing Comparative Study into the Performance of Human Analysis Versus Hybrid Neural Network Analysis*, "Analecta Prehistorica Leidensia", 28(1), pp. 229-232.

E. Grilli, E. Özdemir, F. Remondino 2019, *Application of Machine and Deep Learning Strategies for the Classification of Heritage Point Clouds,* "International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences", XLII-4/W18.

M. Grove, J. Blinkhorn 2020, *Neural networks differentiate between Middle and Later Stone Age lithic assemblages in eastern Africa,* "PloS one", 15(8), e0237528.

M. Gualandi, G. Gattiglia, F. Anichini 2021, *An Open System for Collection and Automatic Recognition of Pottery through Neural Network Algorithms*, "Heritage", 4(1), pp. 140-159.

Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, M.S. Lew 2016, *Deep learning for visual understanding: A review,* "Neurocomputing", 187, pp. 27-48.

A. Guyot, M. Lennon, L. Hubert-Moy 2021, *Objective comparison of relief visualization techniques with deep CNN for archaeology,* "Journal of Archaeological Science: Reports", 38, 103027.

Ø. Hammer, D.A. Harper 2008, *Paleontological Data Analysis*, Oxford.

K. He, X. Zhang, S. Ren, J. Sun 2016, *Deep residual learning for image recognition*, "Proceedings of the IEEE conference on computer vision and pattern recognition", pp. 770-778.

T. Hegazy, P. Fazio, O. Moselhi 1994, *Developing practical neural network applications using back-propagation*, "Computer-Aided Civil and Infrastructure Engineering", 9(2), pp. 145-159.

I. Hein, A. Rojas-Domínguez, M. Ornelas, G. D'Ercole, L. Peloschek 2018, *Automated classification of archaeological ceramic materials by means of texture measures,* "Journal of Archaeological Science", 21, pp. 921-928.

R. Hermoza, I. Sipiran 2018, *3D reconstruction of incomplete archaeological objects using a generative adversarial network*, in *Proceedings of Computer Graphics International (June 2018)*, pp. 5-11. DOI: https://doi.org/10.1145/3208159.3208173.

K.A. Hoehne 1980, *Classification vs typology: A difference of practical importance*, "JAMA", 244(10), pp. 1099-1100.

M. Islam, G. Chen, S. Jin 2019, *An overview of neural network*, "American Journal of Neural Networks and Applications", 5(1), pp. 7-11.

S. James 2015, *Visual competence in archaeology: a problem hiding in plain sight*, "Antiquity", 89(347), pp. 1189-1202.

T. Jiang, J.L. Gradus, A.J. Rosellini 2020, *Supervised machine learning: a brief primer*, "Behavior Therapy", 51(5), pp. 675-687.

R. Kishore, T. Kaur 2012, *Backpropagation algorithm: an artificial neural network approach for pattern recognition,* "International Journal of Scientific & Engineering Research", 3(6), pp. 1-4.

T. Kohonen 2001, *Self-Organizing Maps*, Berlin.

S. Kotsiantis, I. Zaharakis, P. Pintelas 2007, *Supervised machine learning: A review of classification techniques,* "Emerging Artificial Intelligence Applications in Computer Engineering", 160(1), pp. 3-24.

A. Krizhevsky, I. Sutskever, G.E. Hinton 2012, *ImageNet Classification with Deep Convolutional Neural Networks*, "Neural Information Processing Systems", 12, pp. 1097-1105. Online in: https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf (accessed 4 july 2015).

G. Kvascev, M.G. Kvascev, Z. Djurovic 2012, *Radial basis function network based feature extraction for improvement the procedure of sourcing neolithic ceramics*, in *11th Symposium on Neural Network Applications in Electrical Engineering*, Belgrade, pp. 95-100. DOI: 10.1109/NEUREL.2012.6419973.

Q. Lei, B. Wen, Z. Ouyang, J. Gan, K. Wei 2020, *Research on image recognition method of ethnic costume based on VGG,* in X. Chen, H. Yan, Q. Yan, X. Zhang (eds), *Machine Learning for Cyber Security*, Cham, pp. 312-325. DOI: https://doi.org/10.1007/978-3-030-62463-7_29.

D.S. Levine, M. Aparicio IV 2013, *Neural Networks for Knowledge Representation and Inference,* Hillsdale.

J.W. Lin 2021, *Is the algorithm of artificial neural network a deduction or induction? Discussion between natural sciences, mathematics and philosophy*, "European Journal of Information Technologies and Computer Science", 1(4), pp. 6-8.

X. Liu, M. Chi, Y. Zhang, Y. Qin 2018, *Classifying high resolution remote sensing images by fine-tuned VGG deep networks,* in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 7137-7140. DOI: 10.1109/IGARSS.2018.8518078.

J. Llamas, M.P. Lerones, R. Medina, E. Zalama, J. Gómez-García-Bermejo 2020, *Classification of architectural heritage images using deep learning techniques,* "Applied Science", 7(10), 992. DOI: https://doi.org/10.3390/app7100992

E.S. Lohse, C. Schou, R. Schlader, D. Sammons 2004, *Automated Classification of Stone Projectile Points in a Neural Network,* in *The E-Way into the Four Dimensions of Culture Heritage*, Oxford, pp. 431-437.

A. Longo 2021, *The Automation of Philosophy or the Game of Induction,* "Philosophy Today", 65(2), pp. 289-303.

L. Maaten, P. Boon, G. Lange, H. Paijmans, E. Postma 2006, *Computer Vision and Machine Learning for Archaeology,* in J.T. Clark, E.M. Hagemeister (eds), *Digital Discovery. Exploring New Frontiers In Human Heritage,* Fargo, pp. 361-368.

A.B. Markman 2013, *Knowledge Representation*, New York.

D. Maturana, S. Scherer 2015, *Voxnet: A 3d convolutional neural network for real-time object recognition,* in *IEEE/RSJ international conference on intelligent robots and systems* (IROS), pp. 922-928. DOI: 10.1109/IROS.2015.7353481.

G.S. McCall 2018, *Strategies for quantitative research: Archaeology by numbers*, London.

D.L. Medin 1989, *Concepts and conceptual structure*, "American Psychologist", 44(12), pp. 1469-1481.

F. Monna, J. Magail, T. Rolland, N. Navarro, J. Wilczek, J.O. Gantulga, Y. Esin, L. Granjon, A.C. Allard, C. Chateau-Smith 2020, *Machine learning for rapid mapping of archaeological structures made of dry stones–Example of burial monuments from the Khirgisuur culture, Mongolia*, "Journal of Cultural Heritage", 43, pp. 118-128.

E. Moscoso Thompson, A. Ranieri, S. Biasotti 2021, *Automatic segmentation of archaeological fragments with relief patterns using convolutional neural networks*, in A. Chalmers, V. Hulusic (eds), *EUROGRAPHICS Workshop on Graphics and Cultural Heritage*. DOI: https://doi.org/10.2312/gch.20211411.

B.S. Nash, E.R. Prewitt 2016, *The use of artificial neural networks in projectile point typology*, "Lithic Technology", 41(3), pp. 194-211.

P. Navarro, C. Cintas, M. Lucena, J.M. Fuertes, C. Delrieux, M. Molinos 2021, *Learning feature representation of Iberian ceramics with automatic classification models,* "Journal of Cultural Heritage", 48, pp. 65-73.

P. Niknazar, M. Bourgault 2017, *Theories for classification vs. classification as theory: Implications of classification and typology for the development of project management theories*, "International Journal of Project Management", 35(2), 191-203.

M. Olivier, W.B. Verschoof 2021, *Implementing state-of-the-art deep learning approaches for archaeological object detection in remotely-sensed data: the results of cross-domain collaboration,* "Journal of Computer Applications in Archaeology", 4(1), pp. 274-289.

C. Ostertag, M. Beurton-Aimar 2020, *Matching ostraca fragments using a siamese neural network,* "Pattern Recognition Letters", 131, pp. 336-340.

J. Pasquet, S. Demesticha, D. Skarlatos, D. Merad, P. Drap 2017, *Amphora Detection Based on a Gradient Weighted Error in a Convolution Neuronal Network*, in *IMEKO International Conference on Metrology for Archaeology and Cultural Heritage* (Lecce, October 23-25, 2017), Lecce. Online: http://ktisis.cut.ac.cy/handle/10488/10646.

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

J.N. Pérez 2014, *Implementación de Redes Neuronales Artificiales en el diseño de modelos predictivos de expectativa arqueológica*, "Mapping", 165, pp. 4-17.

J.C. Principe, N.R. Euliano, W.C. Lefebvre 2000, *Neural and Adaptive Systems. Fundamentals through Simulations*, New York.

S. Qi, X. Ning, G. Yang, L. Zhang, P. Long, W. Cai, W. Li 2021, *Review of multi-view 3D object recognition methods based on deep learning,* "Displays", 69, 102053.

M. Ramazzotti 2013, *ARCHEOSEMA. Sistemi artificiali adattivi per un'archeologia analitica e cognitiva dei fenomeni complessi*, "Archeologia e Calcolatori", 24, pp. 283-303.

A. Ramil, A.J. López, A. Yáñez 2008, *Application of artificial neural networks for the rapid classification of archaeological ceramics by means of laser induced breakdown spectroscopy (LIBS),* "Applied Physics A", 92(1), pp. 197-202.

N.A. Rasheed, M.J. Nordin 2020, *Classification and reconstruction algorithms for the archaeological fragments,* "Journal of King Saud University-Computer and Information Sciences", 32(8), pp. 883-894.

D.W. Read 2007, *Artifact Classification: A Conceptual and Methodological Approach*, London.

C. Reeler 1999, *Neural Networks and Fuzzy Logic Analysis in Archaeology*, in L. Dingwall, S. Exon, V. Gaffney, S. Laflin, M. Van Leusen (eds), *Archaeology in the Age of the Internet (CAA 97),* Oxford (CD-ROM). Online in: https://bibliographie.uni-tuebingen.de/xmlui/bitstream/handle/10900/62053/CD14_Reeler_CAA_1997.pdf?sequence=2

K.M. Reese 2021, *Deep learning artificial neural networks for non-destructive archaeological site dating,* "Journal of Archaeological Science", 132, 105413.

A. Resler, R. Yeshurun, F. Natalio, R. Giryes 2021, *A deep-learning model for predictive archaeology and archaeological community detection*, "Humanities and Social Sciences Communications", 8(1), pp. 1-10.

P.J. Rousseeuw, L. Kaufman 1990, *Finding Groups in Data: An Introduction to Cluster Analysis*, Hoboken.

D.E. Rumelhart, R. Durbin, R. Golden, Y. Chauvin 1995, *Backpropagation: The basic theory*, in *Backpropagation: Theory, Architectures and Applications*, Mahwah (NJ), pp. 1-34.

W. Samek, G. Montavon, S. Lapuschkin, C.J. Anders, K.R. Müller 2021, *Explaining deep neural networks and beyond: A review of methods and applications*, "Proceedings of the IEEE", 109(3), pp. 247-278.

J. Schmidhuber 2015, *Deep learning in neural networks: An overview*, "Neural Networks", 61, pp. 85-117.

E. Schuettpelz, BB. Frandsen, R.B. Dikkow, A. Brown, S. Orli, P. Melinda, A. Metallo, V.A. Funk, L.J. Dorr 2017, *Applications of Deep Convolutional Neural Networks to Digitized Natural History Collections*. "Biodiversity Data Journal", C.5, e21139.

A. Sengupta, Y. Ye, R. Wang, C. Liu, K. Roy 2019, *Going deeper in spiking neural networks: VGG and residual architectures*, "Frontiers in Neuroscience", 13, DOI: https://doi.org/10.3389/fnins.2019.00095.

C. Sevara, M. Pregesbauer, M. Doneus, G. Verhoeven, I. Trinks 2016, *Pixel versus object – A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data,* "Journal of Archaeological Science: Reports", 5, pp. 485-498.

S. Sharafi, S. Fouladvand, I. Simpson, J.A. Barceló 2016, *Application of Pattern Recognition in Detection of Buried Archaeological Sites based on Analyzing Environmental Variables*, "Journal of Archaeological Science: Reports", 8, pp. 206-215.

C. Shelley 1996, *Visual abductive reasoning in archaeology*, "Philosophy of Science", 63(2), pp. 278-301.

M. Soroush, A. Mehrtash, E. Khazraee, J.A. Ur 2020, *Deep learning in archaeological remote sensing: Automated qanat detection in the kurdistan region of Iraq*, "Remote Sensing", 12(3). DOI: https://doi.org/10.3390/rs12030500.

A.S. Soto i Serrano 2017, *YOLO Object Detector for Onboard Driving Images,* End of Degree Project in Computer Science, Universitat Autonoma de Barcelona.

S. Spars 2013, *Exposing those who bury the dead: A new perspective in modern conflict archaeology,* in P. Cornish, N.J. Saunders (eds), *Bodies in Conflict: Corporeality, Materiality and Transformation,* London, pp. 103-122.

J.W. Suh, E. Anderson, W. Ouimet, K.M. Johnson, C. Witharana 2021, *Mapping Relict Charcoal Hearths in New England Using Deep Convolutional Neural Networks and LiDAR Data*, "Remote Sensing", 13(22). DOI: https://doi.org/10.3390/rs13224630.

C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich 2015, *Going deeper with convolutions*, in *Proceedings of the IEEE Conference on Computer Vision and Pat-*

tern Recognition, pp. 1-9. Online in: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html

S.M. Taheri, F.I. Ghadim, M. Kabirian 2019, *Application of fuzzy inference systems in archaeology*, in *7th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, pp. 1-4. DOI: doi: 10.1109/CFIS.2019.8692167.

J.P. Toumazet, F. Vautier, E. Roussel, B. Dousteyssier 2017, *Automatic detection of complex archaeological grazing structures using airborne laser scanning data*, "Journal of Archaeological Science: Reports", 12, pp. 569-579.

A. Traviglia, A. Torsello 2017, *Landscape pattern detection in archaeological remote sensing*, "Geosciences", 7(4), 128. DOI: https://doi.org/10.3390/geosciences7040128.

Ø.D. Trier, D.C. Cowley, A.U. Waldeland 2018, *Using deep neural networks on airborne laser scanning data: results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland,* "Archaeological Prospection", 26(2), pp. 165-175.

I. Tyukin, K. Sofeikov, J. Levesley, A.N. Gorban, P. Allison, N.J. Cooper 2018, *Exploring automated pottery identification [Arch-I-Scan],* "Internet Archaeology", 50, DOI: https://doi.org/10.11141/ia.50.11.

T.L. Van Pool, R.D. Leonard 2011, *Quantitative Analysis in Archaeology*, Chichester.

A. Vedaldi, A. Zisserman 2016, *Vgg convolutional neural networks practical*, Department of Engineering Science, Oxford.

W.B. Verschoof, K. Lambers 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), pp. 31-40.

W.B. Verschoof, K. Lambers, W. Kowalczyk, Q.P.J. Bourgeois 2020, *Combining deep learning and location-based ranking for large-scale archaeological prospection of LiDAR data from the Netherlands, "*ISPRS International Journal on Geo Information", 9, 293. DOI: https://doi.org/10.3390/ijgi9050293.

W.B. Verschoof-van der Vaart, J. Landauer 2021, *Using CarcassonNet to automatically detect and trace hollow roads in LiDAR data from the Netherlands,* "Journal of Cultural Heritage", 47, pp. 143-154.

B.J. Wythoff 1993. *Backpropagation neural networks: a tutorial*, "Chemometrics and Intelligent Laboratory Systems", 18(2), pp. 115-155.

M. Zabir, N. Fazira, Z. Ibrahim, N. Sabri 2018, *Evaluation of pre-trained convolutional neural network models for object recognition*, "International Journal of Engineering and Technology", 7(3-15), pp. 95-98.

R. Zhao, S. Liu, H.C. Ng, E. Wang, J.J. Davis, X Niu, X. Wang, H. Shi, G.A. Constantinides, P.Y.K. Cheung, W. Luk 2018, *Hardware compilation of deep neural networks: An overview*, in *29th international conference on application-specific systems, architectures and processors (ASAP)*, pp. 1-8. DOI: doi: 10.1109/ASAP.2018.8445088.

S. Zhi, Y. Liu, X. Li, Y. Guo 2017, *LightNet: A Lightweight 3D Convolutional Neural Network for Real-Time 3D Object Recognition.* Online in: Eurographics Association. DOI: 10.2312/3dor.20171046.

L.V. Zotkina, V.S. Kovalev 2019, *Lithic or metal tools: Techno-traceological and 3D analysis of rock art,* "Digital Applications in Archaeology and Cultural Heritage", 13, e00099.

B. Zupanek, D. Mlekuz 2001, *Counting the Uncountable: A Quantitative Approach to the Religious Differences Between the Roman Towns of Emona and Poetovio*, in Z. Stancic, T. Veljanovski (eds), *Computing Archaeology for Understanding the Past*, Oxford, pp. 263-268.