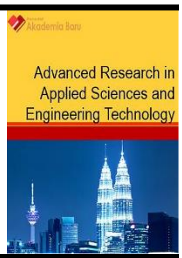
Penerbit Akademia Baru

Journal of Advanced Research in Applied Sciences and Engineering Technology

www.akademiabaru.com/ submit/index.php/araset/index
ISSN: 2462-1943




# A Review on Deep Learning Application for Detection of Archaeological Structures

Amirah Hanani Jamil[1,*], Fitri Yakub[1], Azizul Azizan[2], Shairatul Akma Roslan[1], Sheikh Ahmad Zaki[1], Syafiq Asyraff Ahmad[1]

[1] Malaysia-Japan International Institute Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[2] Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia Kuala Lumpur, Malaysia

**ABSTRACT**

Over the last few years, archaeologist have started to look at automated object detection for searching of potential historical sites, using object identification methods that includes neural network-based and non-neural network-based approaches. However, there is a scarcity of reviews on Convolutional Neural Networks (CNN) based Deep Learning (DL) models for object detection in the archaeological field. The purpose of this review is to examine existing research that has been implemented in the area of ancient structures object detection using Convolutional Neural Networks. Notably, CNN based object detection has the difficulty to draw a boundary box around the object and was implemented mainly for object classification. Various algorithms such as, the Region-based Convolutional Neural Network (R-CNN) and Mask Region-based Convolutional Neural Network (MR-CNN) was developed to solve this problem, yielding a more accurate, time-efficient, and bias-free deep learning model. This paper intends to provide a technical reference highlighting articles from Scopus, Web of Science, and IEEE Xplore databases pertaining to the usage of Convolutional Neural Network based techniques to detect structures and objects in the archaeological field.

***Keywords:***
Deep learning; convolutional neural network; archaeology; structure detection aids

**Received: 3 October 2021** **Revised: 15 January 2022** **Accepted: 20 January 2022** **Published: 25 January 2022**

## 1. Introduction

Archaeology is the study of history through tangible remnants. These relics contain objects that humanity created, altered, or utilised. There are two types of objects in archaeology: artifacts and features [1]. Artifacts are movable remnants like coins, vases, and clothes, whereas features are non-portable structures like pyramids and post-holes. Archaeologists investigate features and artifacts to learn about how people lived in various times and locales. They examine how people live their daily routine, how they were governed, how they talked with one another, and what they believed and valued.

Many archaeological sites are discovered by accident, generally while working on construction projects. Some archaeologists refer to what they do as "running in front of bulldozers to seize

* *Corresponding author.*
*E-mail address: amirahhanani@graduate.utm.my*

https://doi.org/10.37934/araset.26.1.714

artifacts before excavation at a site begins." Others are discovered by looking for mounds and abandoned homes, or by looking for indications in historical texts [2,3].

The approach for uncovering archaeology has undergone a revolution in the previous decade, allowing archaeologists to almost see through the soil without having to dig. Geophysics, soil chemistry, and remote sensing advances are speeding up the study of ancient sites and aiding archaeologists in acquiring latent knowledge of archaeological remnants [4]. Using conventional method, archaeologists have to crawl, kneel, laboriously scrub, and brushing mud off by filtering through a screen in order to retrieve tiny artifacts, soil, sand, and excavated debris [2].

Even if employing technology to detect archaeological features or artifacts does not result in the discovery of the actual relics, such approaches are far more sustainable and ecologically friendly than excavating and potentially damaging artifacts. Google Earth, Light Detection and Ranging (LiDAR), drones, and ground penetrating radar are some of the remote sensing technologies utilised by archaeologists and researchers. Since the advent of aerial photography and the widespread use of remote sensing in archaeology, archaeologists have recognised the necessity of utilising machine learning to identify new sites [5].

Even though many researchers and archaeologist has started to use machine learning (ML) for new site identification and classification, it is uncommon for researchers to employ computer algorithms in archaeology due to the complexity of the ML method which requires computer science specialist [6]. This paper will be organized according to the following sections; section 2 will explain on the basic concept of deep learning, section 3 is the methodology of how the literature review is chosen and reviewed, section 4 is the literature review made, section 5 is the research gap, section 6 is the conclusion and future work and lastly is the acknowledgement.

## 2. Basic Concept of Deep Learning

Artificial Intelligence (AI) is an information investigation machine that uses ML models and approaches that computerizes the decision making and analysis of scientific models. This is the component of man-made consciousness that is focused at intuition with minimal human interaction; computing frameworks can learn from information, perceive patterns, and finally decide. Over the years, the application of ML in archaeology has grown [7], by utilizing fundamental classical methodologies, such as Linear Regression, to cutting edge Deep Learning (DL) models, built with neural networks [7]. However, predictive ML model implementations for archaeological are often inadequate due to limited training data sets [8]. In general, ML can be defined as the capacity of intelligent systems to learn and develop based on prior data [9]. In ML, there are three types of learning which are Supervised learning, Unsupervised learning, and Semi-Supervised learning.

Deep learning models are also known as neural networks with deep structures. The origins of neural networks can be traced back to the 1940s [10,11], where the initial goal was to replicate the human brain system to address generic learning issues in a systematic manner. Rumelhart's proposal of the back-propagation algorithm made it popular again in the 1980s and 1990s [11,12]. However, neural networks fell out of favour in ML research in the early 2000s, due to overfitting in training, lack of large training datasets, limited hardware processing capacity, and insignificant performance improvement compared to other machine learning techniques. Later, deep learning has grown in popularity when a breakthrough in speech recognition was achieved in 2006 [11,13,14].

Nowadays, the most common deep neural network are based on three neural network architecture i.e., Artificial Neural Network (ANN), Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN), where CNN is the most well researched about [11,15]. These CNN models are used across numerous application domains and are particularly prevalent in projects involving image and video processing. Deep learning methods delivers a broad range of possibilities to tackle classical imaging tasks and are now applied in the spatial-spectral domain as well [16,17].

The benefit of using DL is that it can learn the features on its own which is time saving compared to classical ML algorithms. Another feature of deep learning is its infinite accuracy which means greater training or more data input may result in higher accuracy than the previous models. Conversely, DL models requires longer time for training but takes a shorter time for inferencing compared to conventional ML algorithms. For better accuracy, DL needs large amounts of datasets for input during training. The method learns features by passing input through multiple blocks comprising Convolution, Rectified linear units (ReLU), and Pooling layers. After feature learning, classification is performed by simplifying matrices, transferring vectored input from many fully connected layers, and then using the SoftMax algorithm to make decisions based on the model output [18].

## 3. Methodology

For this review of the deep learning application in archaeological structures two main focus area were predetermined, which are deep learning application and application in detection or classification of archaeological structures. The articles are searched using Scopus, Web of Science, and IEEE Xplore with the keywords of “deep learning”, “archaeology”, and “structure”. There are only 27 papers in Scopus, however, among these 27 papers, only six are relevant to this review. The other 19 is related to shipwrecked, potential and limitation of DL in archaeology, proposal of workflow, framework, identifying geo-spatial areas with archaeological artifacts, bronze inscription, tracing of hollow roads, identifying of construction era, processing of visual sensing, artifacts identification and classification, porosity types of structures in photomicrographs from archaeological soil, measuring spectral index of turbulent gas, speech synthesis, apperceptive patterning, recognition of ancient Tamil inscription, topography mapping, and brick segmentation. In IEEE Xplore, only four papers were relevant by using similar keywords. Among these four only two are significant and both of them is published in Scopus as well. The irrelevant articles were looking into archaeological structure as it uses DL for characteristic of extinct species and detection of ruin livestock enclosures. In Web of Science, only one paper is available based on the keywords and the paper has also been published in Scopus. Overall, only six papers are available for our review applying variety of deep learning application to detect archaeological structures.

## 4. Literature Review

In archaeology field, most of the ML and DL algorithm were used for classification and identification of artifacts. There is still insufficient research in detection of archaeological structure using DL algorithm especially using aerial imaging. This section briefly summarizes the research applications of DL for archaeological structure detection based on the six papers that were surveyed.

### *4.1 Automated Qanat Shaft Detection*

In this paper, Mehrnoush Soroush [19] applied CNN based deep learning model to detect qanat shaft using Cold-War Era CORONA Satellite Images. Automated qanat shaft detection is the first attempt that utilized automated detection on historical satellite images. A typical machine learning pipeline includes data collection, data pre-processing, model construction, and model assessment.

During the collection and pre-processing of data, the annotated imaged data is pre-processed using real-world data or expert judgement.

A binary classification model based on CNN is given for the segmentation of qanat characteristics in CORONA pictures. The deep network design is made up of convolutionary layers that are created in a sequential sequence *($l \in [1, L]$).* A set of $K$ kernels $W_l = \{W^1, . . . , W^K\}$ and biases $b_l = \{b^1, . . . , b^K\}$ at each convolutional layer $l$ rotates the input feature map (image) to create a new feature map. After that, the feature map is subjected to a nonlinear activation function $f$, which produces the output $Y_l$, which serves as the input to the following layer. The $n^{th}$ feature map of the $l^{th}$ layer's output may be represented as follows

$$Y_l^n = f\left(\sum_{k=1}^{K} w_l^{n,k} * y_{l-1}^k + b_l^n\right) \tag{1}$$

The mix of feature maps in each layer provides a variety of patterns to the network, which becomes more complicated as the network deepens. CNN is trained using a predefined iteration of the Stochastic Gradient Descent (SGD) algorithm, in which the network analyses a sample of training data. To decrease losses, SGD adjusts network parameters (kernel weight and bias) at each iteration based on losses computed using a cost function. For segmentation issues, Fully Convolutional Neural Networks (FCNNs) are utilized. End-to-end learning is enabled through the use of FCNs for image segmentation, with FCNs mapping each pixel of the input picture to an output segmentation map.

This paper recommends CNN based on 2D FCN. The FCN segmentation consists of an encoder strip (contract) and a decoder strip (expand). The encoder path consists of a repeating convolution layer on the selected feature map, followed by an activation function with a maximum consolidation layer. The encoder path reduces the resolution of the feature map by specifying the maximum number of small patches of the feature map unit. During the training, the researchers used the hard -frame SGD method with the Adam update rule. The learning level was initially set at 0.001. If the mean validation of the Dice score did not increase 10–5 in five periods or epochs, the learning rate was lowered by a factor of 0.8. Five-fold cross-validation was used to train eleven 2000 x 2000 CORONA image patches (approximately 5.5 km x 5.5 km). The overall accuracy and recall of the model are 0.62 and 0.74, respectively. The percentage of positives successfully identified for all annotated positives was measured by recall. Accuracy is the positive part detected correctly for all expected positives.

### 4.2 *Detection of Ringforts using Aerial Photography*

Keith Phelan [20] uses both satellite and aerial photography to detect ringforts by using the data source from Republic of Ireland (RoI) and The Archaeological Survey of Ireland (ASI), Northern Ireland (NI). *fast.ai* library was used during the experiments and the dataset is trained using a CNN based Resnet34 model. The problem for detecting ringfort is that it has a similar shape and size with various features that can be found in Irish structure. ASI and RoI data sets are imported into Comma Separated Value (CSV) and Geographic JSON (GeoJSON) formats, analysed using the *Pandas* Python package, and then stored in the standard GeoJSON format. The data was produced in two steps: the first was to isolate the ring from the other monuments, and the second was to clean and format the data. The RoI Sites and Monuments Record (SMR) database has 154,274 objects categorised by class code. The database was cleaned up, leaving with 29,887 entries, and the NI SMR database, which contained 16,694 items categorised by class code, was cleaned up as well, leaving only 3,031 ringforts entries.

Data sets derived from maps frequently have two major issues: Registration Noise and Ommision Noise. Registration noise occurs when an object's location on the map is inaccurate, and omission noise happens when the object is missing. Several actions are done to address the issue of noisy data. The first was general analysis, which aimed to manage outliers and duplication in a Python script utilising *Pandas* and *GeoPandas* libraries. This includes screening any location that is outside the country's borders, a crude approach but enough for the initial run. Following filtering of both the NI SMR and RoI datasets, the residual data sets were 29,772 in RoI SMR and 3,019 in NI SMR.

The Resnet34 model was then trained using a one-round method, allowing researchers to train the model faster; after two rounds, the accuracy is 85%. Based on the initial findings, a learning rate of 0.0001 was chosen, and the model was trained across 10 cycles, providing 95% accuracy.

### *4.3 Barrow and Celtic Field Detection in LiDAR Data*

The Workflow for Object Detection of Archeology in the Netherlands (WODAN) was a newly created automated detection method that was utilised for archaeological practise by the researchers [21]. WODAN was used to look for Celtic barrows and farms in LiDAR data from the Dutch Midden-Limburg region, which differed from the Veluwe in archaeology, geomorphology, and land usage. WODAN was able to locate probable Celtic barrows and farms, including previously unknown instances, as well as offer historical landscape architectural information.

The object detection algorithm was trained using 1,152 LiDAR images (600 x 600 pixels). The data was a collection of pictures that included archaeological artifacts from diverse places, and images that did not contain archaeological items were removed from the training data set. A test data set of 4,405 LiDAR images was created using a 265 $km^2$ region in Midden-Limburg (600 x 600 pixels).

WODAN was utilised in two ways: one model was trained and assessed on non-visualized Digital Terrain Model (WODAN DTM) data, while the second model was trained and evaluated on data visualised with the Local Assistance Model (WODAN LRM). The same pre-processing approach was used in [22] to transform LiDAR data into input pictures for both versions. WODAN's performance is not evaluated using metrics such as F1-score or accuracy since the researcher do not have a validated baseline against which to compare performance. As a result, the ratio of overlap between the automatic detection results and the two reference data sets is reported. WODAN has successfully identified 40.9% of barrows recorded and visible in LiDAR data, and 30.7% of barrows annotated manually. These findings indicate that WODAN was able to identify Celtic barrows and farms in the Midden-Limburg area.

### *4.4 Ancient Maya Structure Semantic Segmentation*

In a research by Marek Bundzel [23], a data collection from the Pacunam LiDAR Initiative study of Guatemala's lowland Maya area, were used to identify ancient Maya buildings that have been hand identified and extensively ground checked. For semantic segmentation, two DL-based models were constructed and evaluated. The two DL algorithms are U-Net and Mask Region-Based Convolutional Neural Network (R-CNN).

The segmentation models were applied to two tasks which are identifying areas of historic construction activity and identifying the remains of ancient Maya structures. A variety of quantifiers were used to assess the quality of the final forecast. In structure segmentation, both U-Net and Mask R-CNN achieved the similar results, and in mound segmentation, U-Net algorithm outperform Mask R-CNN. There is no clear reason on why Mask R-CNN is outperformed by U-Net on mound segmentation. In overall, the U-Net-based model outperformed Mask R-CNN in both tests, successfully recognizing 60% of all items and 74% of medium-sized objects.

### *4.5 Structure Classification in the Maya Settlement using Airborne Laser Scanning*

Manual inspection of Airborne Laser Scanning (ALS) data takes a long time and poses a considerable challenge in the data analysis method. As a result, researchers [24] attempted to create and test deep neural network models to identify ancient Maya architecture previously explained manually at the Chactn archaeological site in Campeche, Mexico. Using previously published ALS pictures of man-made aguadas, buildings, and platforms, as well as photographs of the surrounding environment, several CNN VGG-19 variants were investigated to tackle the challenge of detecting the structure of the visible sample (four classes and approximately 12,000 anthropogenic structures). The researchers used six different combinations of alternate visualisations, two different edge buffer measurements, extra data, and an architecture with many layers of frozen layers that could not be trained at network starting to explore how many factors influenced model performance. A vast number of models evaluated under different situations achieved an overall classification accuracy of 95%.

### *4.6 Convolutional Neural Network with Modified Mask Region Based for Automated Detection of Archaeological Site*

With the complicated backdrop and uncertain target orientation, automatic object detection in archaeology is challenging. The two-stage Mask R-CNN method has lately yielded impressive results in object recognition and sample segmentation issues, and it has been successfully applied to the analysis of archaeological ALS data. The researchers reported a modified Mask R-CNN technique for detecting the location of a relic charcoal fireplace using a digital elevation model (DEM) based on LiDAR data in [25].

The team was able to enhance the model's accuracy and minimise training time by combining picture magnification and image preprocessing with an adaptive gradient approach based on deep learning and a dynamic strategy on learning rate optimization. To produce high-contrast pictures of the landscapes and contours of locations of interest in the Northern German Lowlands, a DEM based on high-resolution LiDAR data and visualisations for the archaeological topography method was utilised. As a consequence, the algorithm correctly recognised relic charcoal fireplace sites with an average retraction of 83% and an accuracy of 87%. Summary of overall literature review is summarized in Table 1.

**Table 1**
Summary of overall literature review

| Article | Deep Learning Type | Object Detection | Classification | Satellite Imagery | Aerial Photography | LiDAR | ALS | Accuracy |
|---|---|---|---|---|---|---|---|---|
| [19] | CNN | X | | X | | | | 62% |
| [20] | CNN | X | | X | X | | | 95% |
| [21] | WODAN | X | | | | X | | Not available |
| [23] | U-Net & Mask R-CNN | X | | | | X | | 60% & 55% |
| [24] | CNN | | X | | | | X | 95% |
| [25] | Modified Mask R-CNN | X | | | | | X | 87% |

## 5. Research Gap

DL consumes large amounts of data in order to achieve greater accuracy. There are not many large data sets available for the archaeological field and from this survey there is just one article for object detection applied to aerial photography. For object detection, the majority of the research was conducted utilising LiDAR data and satellite images. In comparison to aerial images, satellite imagery is more difficult for the general public to access. As a result, there is a large research gap that needs to be filled in order to implement DL based object detection for aerial photography of ancient structures.

## 6. Conclusion and Future Works

DL models are gaining importance in the realm of archaeology application, offering higher accuracy than any previous non-intrusive approach. There is a potential research gap in which deep learning may be utilised for archaeological application in diverse structures classification. DL methods can assist archaeologists in detecting objects faster than standard conventional methods whilst saving money and time as shown in the articles that has been reviewed. To date, there is no open and public access standard data, that are utilised by archaeologists worldwide; therefore, a standard dataset highly can assist archaeologists to evaluate ML and DL models in object detection and classification.

## Acknowledgement
A research grant from Universiti Teknologi Malaysia was used to fund this work as part of the Industry-International Incentive Movement, Vot. Number 02M92. This work was partly supported by a research grant from Universiti Teknologi Malaysia under the UTMShine Grant Vot. Number 09G38.

## References

[1] Diane Boudrea, Melissa McDaniel, Erin Sprout, Turgeon A. *National Geographic - Encyclopedia - Archaeology* 2011.

[2] *Archaeology, Tools, Methods and Analysis* 2018.

[3] Bahn P. *Written in Bones: How Human Remains Unlock the Secret of the Dead*: Firefly Books; 2012. 224 p.

[4] Horton, M., and V. Heyd. "Six tools that are revolutionising archaeology by helping us find sites without digging." *Retrieved from theconversation. com: https://theconversation. com/six-tools-thatare-revolutionising-archaeology-by-helping-us-find-sites-without-digging-51826* (2015).

[5] Monna, Fabrice, Jérôme Magail, Tanguy Rolland, Nicolas Navarro, Josef Wilczek, Jamiyan-Ombo Gantulga, Yury Esin, Ludovic Granjon, Anne-Caroline Allard, and Carmela Chateau-Smith. "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 (2020): 118-128.

[6] Caspari G, Crespo P. *Convolutional Neural Networks for Archaeological Site Detection -Finding "Princely" Tombs*. Archaeological Science. 2019;**110**:26.

[7] Fiorucci, Marco, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. "Machine learning for cultural heritage: A survey." *Pattern Recognition Letters* 133 (2020): 102-108.

[8] Yaworsky, Peter M., Kenneth B. Vernon, Jerry D. Spangler, Simon C. Brewer, and Brian F. Codding. "Advancing predictive modeling in archaeology: An evaluation of regression and machine learning methods on the Grand Staircase-Escalante National Monument." *PloS one* 15, no. 10 (2020): e0239424.

[9] Akinosho, Taofeek D., Lukumon O. Oyedele, Muhammad Bilal, Anuoluwapo O. Ajayi, Manuel Davila Delgado, Olugbenga O. Akinade, and Ashraf A. Ahmed. "Deep learning in the construction industry: A review of present status and future innovations." *Journal of Building Engineering* (2020): 101827.

[10] Pitts, Walter, and Warren S. McCulloch. "How we know universals the perception of auditory and visual forms." *The Bulletin of mathematical biophysics* 9, no. 3 (1947): 127-147.

[11] Zhao Z-Q, Zheng P, Xu S-T, Wu X. *Object Detection with Deep Learning: A Review*. IEEE Transactions on Neural Networks and Learning Systems. 2019;**PP**:1-21.

[12] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." *nature* 323, no. 6088 (1986): 533-536.

[13] Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *science* 313, no. 5786 (2006): 504-507.

[14] Hinton, Geoffrey, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior et al. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups." *IEEE Signal processing magazine* 29, no. 6 (2012): 82-97.

[15] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436-444.

[16] Signoroni, Alberto, Mattia Savardi, Annalisa Baronio, and Sergio Benini. "Deep learning meets hyperspectral image analysis: A multidisciplinary review." *Journal of Imaging* 5, no. 5 (2019): 52.

[17] Li, Shutao, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi, and Jon Atli Benediktsson. "Deep learning for hyperspectral image classification: An overview." *IEEE Transactions on Geoscience and Remote Sensing* 57, no. 9 (2019): 6690-6709.

[18] Moazzam, Syed I., Umar Shahbaz Khan, Mohsin Islam Tiwana, Javed Iqbal, Waqar S. Qureshi, and Syed Irfan Shah. "A review of application of deep learning for weeds and crops classification in agriculture." In *2019 International Conference on Robotics and Automation in Industry (ICRAI)*, pp. 1-6. IEEE, 2019.

[19] Soroush, Mehrnoush, Alireza Mehrtash, Emad Khazraee, and Jason A. Ur. "Deep learning in archaeological remote sensing: Automated qanat detection in the kurdistan region of Iraq." *Remote Sensing* 12, no. 3 (2020): 500.

[20] Phelan, Keith, and Daniel Riordan. "Detection of ringforts from aerial photography using machine learning." In *2020 31st Irish Signals and Systems Conference (ISSC)*, pp. 1-6. IEEE, 2020.

[21] Verschoof-van der Vaart, Wouter B., and Karsten Lambers. "Applying automated object detection in archaeological practice: A case study from the southern Netherlands." *Archaeological Prospection* (2021).

[22] Verschoof-van der Vaart, Wouter B., Karsten Lambers, Wojtek Kowalczyk, and Quentin PJ Bourgeois. "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, no. 5 (2020): 293.

[23] Bundzel, Marek, Miroslav Jaščur, Milan Kováč, Tibor Lieskovský, Peter Sinčák, and Tomáš Tkáčik. "Semantic segmentation of airborne lidar data in maya archaeology." *Remote Sensing* 12, no. 22 (2020): 3685.

[24] Somrak, Maja, Sašo Džeroski, and Žiga Kokalj. "Learning to classify structures in ALS-derived visualizations of ancient Maya settlements with CNN." *Remote Sensing* 12, no. 14 (2020): 2215.

[25] Bonhage, Alexander, Mahmoud Eltaher, Thomas Raab, Michael Breuß, Alexandra Raab, and Anna Schneider. "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* (2021).