Survey of Intelligent Archaeology

Lin Yuhang, Peng Shubo

***Abstract*—AAA**

***Index Terms*—AAA**

I. Introduction

ARTIFICIAL Intelligence (AI for short) is a new study area rising in recent centuries, and can be divided into several directions, like Machine Learning (ML) or Neural Network (NN). Five to ten years ago, they were concepts unknown to archaeologists. But now, AI is widely used, there are even sessions dedicated to AI at archaeological conferences.[4]

A. Machine Learning

ML is programming allowing algorithm learning from data and adjust its parameters and then make predictions on new data. Objects must be quantify into digital data first and can be any type, like sonar data under water[3] or aerial laser scanning data [5]. ML uses mathematical techniques to analyze a set of already-classified objects and generate "classifiers" for each category. In theory, objects in every category is identified from other categories in mathematics. In short, ML use math to classify quantifiable objects into different groups.[2]

B. ML Application

ML application can be divided into two main types: to classify archaeological objects, and to detect objects in archaeological sites.

II. Application

Deep learning (DL) has been successfully used in many applications. Among the DL methods, recurrent neural networks (RNNs) are good at dealing with sequential data as they take into account temporal information. RNNs have been applied in speech recognition to map acoustic sequences to phonetic sequences. RNNs have also been used in natural language processing to translate text from one language to another. Another famous method in DL is convolutional neural networks (CNNs). CNNs take into account spatial correlation among data points and hence perform well in image-based data. CNNs have been used in image classification, face recognition, scene labelling, and so on. DL methods are also used in the field of remote sensing. For example, Hu and Yuan used CNNs to extract digital terrain models (DTMs) and filter out non-ground points from airborne laser scanning (ALS) data, which was claimed that this method performs better than previous filtering methods[6].

In the papers we've read, we've found that all of them, without exception, use CNNs to study a specific thing. This is because the things they studied were all based on image data, in which CNNs performs better than RNNs.

We divide the papers that use CNNs into two types: one is to classify archaeological objects like papyrus, potteries, and so on, the other is to detect objects in archaeological sites.

A. classify archaeological objects

- In the paper[1], a method is proposed for matching and assembling pairs of ancient papyrus fragments containing mostly unknown scriptures. This task, which is assembling fragments in a puzzle-like manner into a composite picture, plays an important role in the archaeology, because it can help historic artifacts to reconstruct archaeological objects for research. The proposed method is to use image processing and machine learning techniques to identify matching fragments, and then support the quick and automated classification of matching pairs of papyrus fragments as well as the geometric alignment of the pairs against each other. The algorithm was trained on a batch of fragments which was excavated from the Dead Sea caves and is dated circa the 1st century BCE. The algorithm shows excellent results on a validation set which is of a similar origin and conditions. Then the algorithm was used to against a real-life set of fragments for with no prior knowledge or labeling of matches. This test batch is considered extremely challenging due to its poor condition and the small size of its fragments. Evidently, numerous researchers have tried seeking matches within this batch with very little success. The algorithm performance on this batch was sub-optimal, returning a relatively large ratio of false positives. However, the results showed that this algorithm eliminated 98% of the possible matches thus reducing the amount of work needed for manual inspection, which means this algorithm was quite useful. Indeed, experts that reviewed the results have identified some positive matches as potentially true and referred them for further investigation.
- In the paper[7], the project developed at Dumbarton Oaks—a research institute and library, museum, and historic garden affiliated with Harvard University and located in Washington, DC— focused on a collection of 10,000 images of Syrian monuments in the institution's Image Collections and Fieldwork Archives (ICFA). Drawing on that project, as well as the broader landscape of AI-based categorisation efforts in the fields of art and architecture, authors' insights on the potential of AI to facilitate and enhance archival image access and recording will be shaerd. Many of the Syrian sites in the Dumbarton Oaks collection have been inaccessible

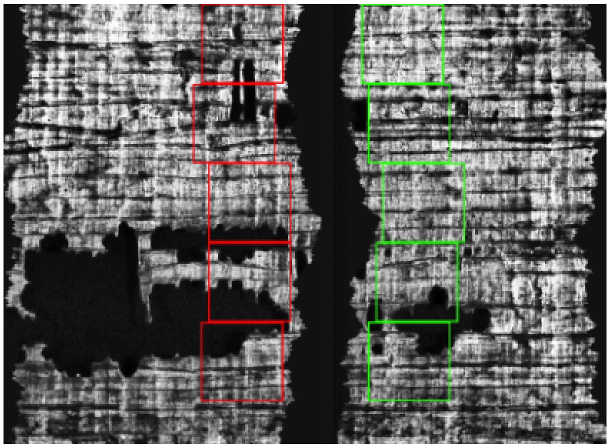

Fig. 1: An example of two adjacent artificially torn fragments with a set of candidate squares to be used in the matching phase.

to researchers and the public for over a decade and/or have been damaged or destroyed. For Dumbarton Oaks, the primary goal was to explore whether AI can improve the speed and efficiency of sharing collections and allow for more sophisticated curation by subject experts who, thanks to automation, would be relieved of the burden of rote processing. The methods and techniques explored included multi-label classification, multi-task classification, unsupervised image clustering, and explainability.

Fig. 2: Explainability heatmaps for predictions of the classes “architecture” (left) and “façades” (right) by the Phase 1 classifier.

- The main contribution in this paper[4] is the completion of the project called ArchAIDE. This project realised an AI-based application to recognise archaeological pottery. Pottery is of paramount importance for understanding archaeological contexts. However, recognition of ceramics is still a manual, time-consuming activity , reliant on analogue catalogues. The project developed two complementary machine-learning tools to propose identifications based on images captured on-site, for optimising and economising this process, while retaining key decision points necessary to create trusted results. One method relies on the shape of a potsherd; the other is based on decorative features. For the shape-based recognition, a novel deep-learning architecture was employed, integrating shape information from points along the inner and outer profile of a sherd. The decoration classifier is based on relatively standard architectures used in image recognition. In both cases, training the algorithms meant facing challenges related to real-world archaeological data: the scarcity of labelled data; extreme imbalance between instances of different categories; and the need to take note of minute differentiating features. Finally , the creation of a desktop and mobile application that integrates the AI classifiers provides an easy-to-use interface for pottery classification and storing pottery data.

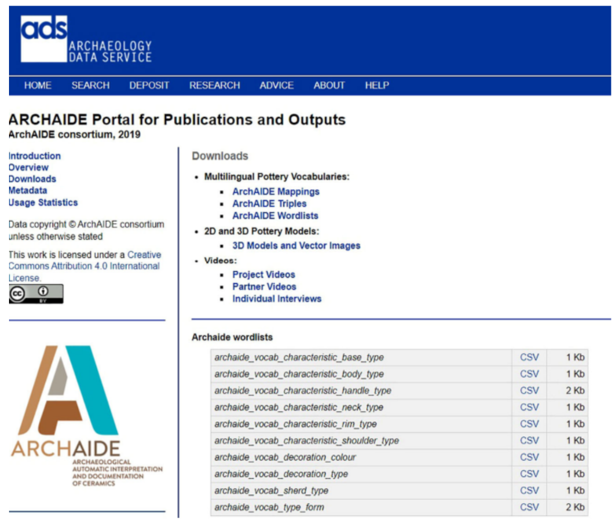

Fig. 3: The ArchAIDE portal is available at the Archaeology Data Service of the University of York.

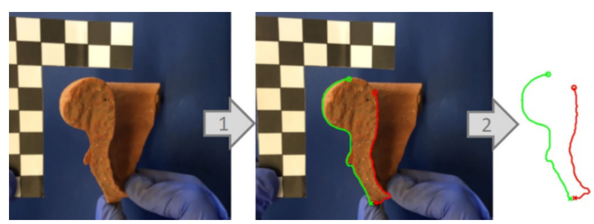

Fig. 4: The automated extraction of the outer (green) and inner (red) profiles from a real-world sherd image.

- The paper[9] indicate that deep learning with CNNs is a highly accessible and effective method for classifying ceramic fabrics based on images of petrographic thin sections and that it can likely be applied on a larger scale. Classification of ceramic fabrics has long held a major role in archaeological pursuits. It helps answer research questions related to ceramic technology, provenance, and exchange and provides an overall deeper understanding of the ceramic material at hand. One of the most effective means of classification is through petrographic thin section analysis. However, ceramic petrography is a difficult and often tedious task that requires direct observation and sorting by domain experts. In this paper, a deep learning model is built to automatically recognize and classify ceramic fabrics, which expedites the process of classification and lessens the requirements on experts. The samples consist of images of petrographic thin sections under cross-polarized light originating from the Cocal-period (AD 1000-1525) archaeological site of Guadalupe on the northeast coast of Honduras. Two CNNs, VGG19 and ResNet50, are compared against each other using two approaches to partitioning training, validation, and testing data. The technique employs a standard transfer learning process whereby the bottom layers of the CNNs are pre-trained on the ImageNet dataset and frozen, while a single pooling layer and three dense layers are added to 'tune' the model to the thin section dataset. After selecting fabric groups with at least three example sherds each, the technique can classify thin section images into one of five fabric groups with over 93% accuracy in each of four tests.

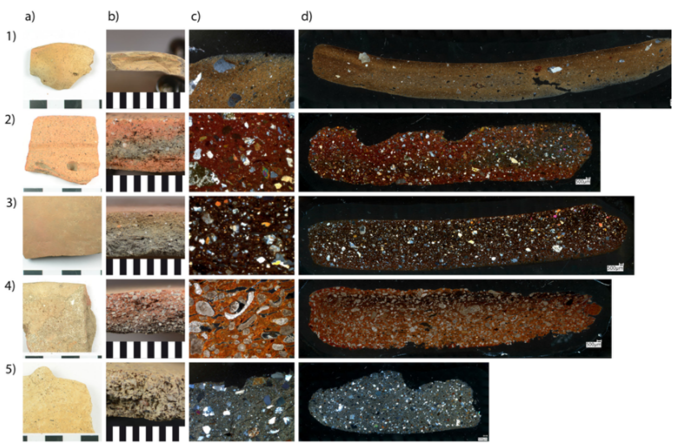

Fig. 5: Examples of the five ceramic fabric types analyzed.

- In the study[10], an alternate approach to archaeological typology which uses DL to classify digital images of decorated pottery sherds into an existing typological framework is presented. This study focuses on a specific kind of ancient painted pottery from the American Southwest, Tusayan White Ware (TWW), but it is believed that it has broader implications for a wide range of geographical settings and artifact types. The results show that when properly trained, a deep learning model can assign types to digital images of decorated sherds with an accuracy comparable to, and sometimes higher than, four expert-level contemporary archaeologists. The technique also offers novel tools for visualizing both the importance of diagnostic design elements and overall design relationships between groups of pottery sherds. This method can objectively match a specific unclassified sherd image to its most similar counterparts through a search of thousands of digital photos. This discovery has important archaeological implications for analyzing time relationships, monitoring stylistic trends, reconstructing fragmentary artifacts, identifying ancient artisans, and studying the evolution and spread of ancient technologies and styles. It also shows how deep learning models can potentially supplement or supplant traditional typologies in favor of more direct groupings and comparisons of artifacts.

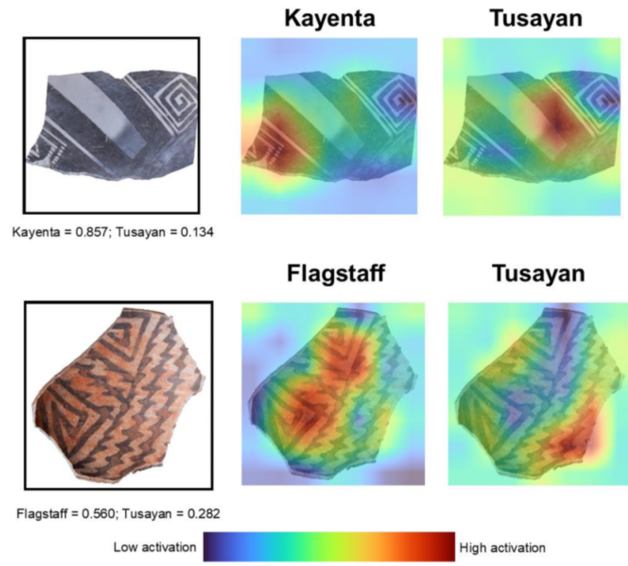

Fig. 6: Grad-CAM heat maps for TWW sherds, showing areas of high (red) and low (blue) model activation for a Kayenta sherd (top) and Flagstaff sherd (bottom). CNN-model-calculated type confidences shown below each sherd.

B. detect objects in archaeological sites

Airborne laser scanning (ALS) is of great use in collecting and documenting detailed measurements from an area of interest. However, it is time consuming for scientists to manually analyze the collected ALS data. One possible way to automate this process is using deep neural networks.

In the paper[8], a hierarchical CNN model is builded to detect objects in archaeological sites using digital terrain models (DTMs) generated from ALS data. The data is acquired from the Harz mining Region in Lower Saxony, where a high density of different archaeological monuments including the UNESCO world heritage site Historic Town of Goslar, Mines of Rammelsberg, and the Upper Harz Water Management System can be found. Objects to be detected are archaeological objects such as hollow ways, streams, pathways, lakes, streets, ditches, heaps, mining shafts, and more, but for this study, the model is fit to detect 4 classes of objects: natural streams, lakes, tracks, and an 'others' class which represents the rest of the objects for which enough labeled data is not available yet. To compare and validate the method in this paper, some experiments on the same data set using two existing deep learning models were conducted. The first model is VGG-16; an image classification network pretrained on ImageNet data. The second model is a stacked autoencoders model. The results of the classification as analyzed in this paper show that our model is suitably tuned for this task as it achieves the best classification accuracy of around 91 percent, compared to 88 percent and 82 percent accuracy by the pretrained and stacked autoencoders models, respectively.

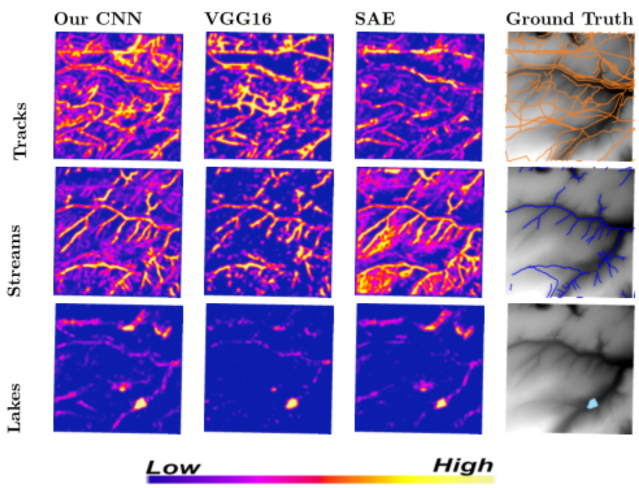

Fig. 7: Heat maps using filter size 48 x 48. Colors show the confidence of the models in detecting objects at that location.

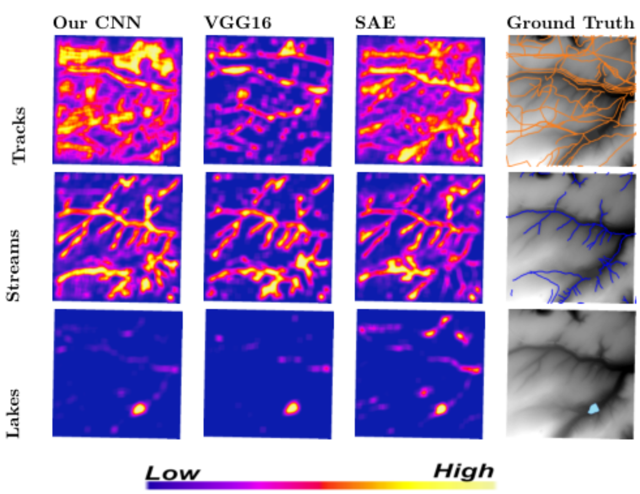

Fig. 8: Heat maps using filter size 98 x 98. Colors show the confidence of the models in detecting objects at that location.

C. More

We believe that there are many other studies using RNNs to study the meaning of certain patterns on ancient artifacts or to try to decipher ancient texts. But this is out of the scope of our survey.

III. A DETAILED CLASSIFICATION

IV. FUTURE

Nowadays, archaeologists utilize AI in many ways, from creating 3D models of historical sights to scanning territories with a laser radar to find ancient graves or from matching and assembling pairs of ancient papyrus fragments containing mostly unknown scripture to detecting objects in archaeological sites using DTMs generated from ALS data. There is no denying that AI becomes more and more popular in the field of archaeology and plays an important role in it.

One direction for the future development of intelligent archaeology is helping archaeologists' work. For example, archaeologists often face such problems as not knowing where exactly to dig. They can define the region but not the exact place where an artifact or a grave lies. It's when the neural network comes in line. Instead of looking through millions of documents by themselves, archaeologists pass this work to neural networks. This technology can sort out information by utilizing a specific algorithm. By analyzing images, this system might not only direct archaeologists in their groundworks but also suggest territories that have similar patterns as potential objects for excavations.

However, knowing that many ML systems — especially deep neural networks — are essentially considered black boxes. This makes it hard to understand and explain the results given by a model. Because of this, it should be noted that AI does not replace the need for experts in archaeology. Instead, AI technology needs the expertise from archaeologists to improve itself and to judge the correctness of the results.

Another direction is to strengthen the identification of the artifacts. In our world, one of the most urgent problems in archaeology is the fact that many artifacts are traded on the dark web. Although many models trained in this direction have achieved quite good results, they are not yet ready to be applied in practice. Currently, the majority of detection operations are performed manually. If an AI can succeed in this direction, it would make an extraordinary contribution to the prevention of art-related illegal activities.

We expect that in the future, artificial intelligence technology will play an increasing, even irreplaceable role in the field of archaeology.

REFERENCES

[1] ABITBOL, R., SHIMSHONI, I., AND BEN-DOV, J. Machine learning based assembly of fragments of ancient papyrus. *J. Comput. Cult. Herit. 14*, 3 (jul 2021).

[2] BICKLER, S. H. Machine learning arrives in archaeology. *Advances in Archaeological Practice 9*, 2 (2021), 186–191.

[3] DRAP, P., SCARADOZZI, D., GAMBOGI, P., AND GAUCH, F. Underwater photogrammetry for archaeology- the venus project framework. In *GRAPP* (2018).

[4] GUALANDI, M. L., GATTIGLIA, G., AND ANICHINI, F. An open system for collection and automatic recognition of pottery through neural network algorithms. *Heritage 4*, 1 (2021), 140–159.

[5] GUYOT, A., HUBERT-MOY, L., AND LORHO, T. Detecting neolithic burial mounds from lidar-derived elevation data using a multi-scale approach and machine learning techniques. *Remote Sensing 10*, 2 (2018).

[6] HU, X., AND YUAN, Y. Deep-learning-based classification for dtm extraction from als point cloud. *Remote sensing (Basel, Switzerland) 8*, 9 (2016), 730–730.

[7] KARTEROULI, K., AND BATSAKI, Y. Ai and cultural heritage image collections: Opportunities and challenges.

[8] KAZIMI, B., THIEMANN, F., MALEK, K., SESTER, M., AND KHOSHELHAM, K. Deep learning for archaeological object detection in airborne laser scanning data.

[9] LYONS, M. Ceramic fabric classification of petrographic thin sections with deep learning. *Journal of Computer Applications in Archaeology 4* (09 2021), 188.

[10] PAWLOWICZ, L. M., AND DOWNUM, C. E. 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* (2021), 105375–.