CAA2023

50 years of synergy

Amsterdam

3-6 April 2023

Abstract Book

Table of Contents

2018). A low fidelity prototype interface was designed based on interviews with professional archaeological surveyors working in heritage management. The design was also informed by a rapid review of existing ML-powered tools. The prototype was tested in two online workshops with aerial archaeology specialists. The resulting data was analysed using qualitative coding to produce insights grounded in stakeholder feedback.

Results

Key findings of the prototype testing show that trust, accountability, and the provenance of outputs are significant issues for the integration of ML tools into professional practice. For example, an essential requirement for heritage management applications of ML is an audit trail which documents all stages of analysis and interpretation. The prototype testing also presents GIS as a key site of focus for the development of ML-augmented archaeological topographic survey practice. This is in contrast to research which situates automated workflows outside of GIS in notebooks or other code development locations.

Discussion and Conclusions

The implications of these results are that human-centred research is beginning to build a set of features essential to the design of successful ML-integrated topographic survey toolkits and workflows for archaeology.

To develop these resources, insights from the prototype testing will be mapped against an a framework for human-centred ML system design which is based on a synthesis of 30 published guidelines from across communities who research human-centred and social computing (human-computer interaction and computer-supported cooperative work) as well as technological implementation and deployment in real-world scenarios (MLOps). These guidelines are applicable to the archaeological use case because they consider applied ML from a domain-agnostic standpoint.

The ML systems design framework can be used as a basis to produce joined-up understandings of the challenges related to ML tool design, which will help the archaeological ML community make progress by providing practical steps on how to approach solving common problems in human-centred ML. A shared understanding of the challenges of developing ML tools from across disciplines and applications will help to define what problems have common solutions, and what problems are specific to archaeological or wider cultural heritage use cases.

Work-in-progress on future elements of the development of this toolkit will be presented, specifically the development of an open source catalogue of models and datasets as a trust- and accountability-building resource for the archaeological ML community and stakeholders.

References

Benyon, David. 2018. Designing User Experience: A Guide to Hci, Ux and Interaction Design. Harlow, United Kingdom: Pearson Education, Limited. http://ebookcentral.proquest.com/lib/gla/detail.action?docID=5632313.

Fiorucci, Marco, Wouter B. Verschoof-van der Vaart, Paolo Soleni, Bertrand Le Saux, and Arianna Traviglia. 2022. 'Deep Learning for Archaeological Object Detection on LiDAR: New Evaluation Measures and Insights'. Remote Sensing 14 (7): 1694. https://doi.org/10.3390/rs14071694.

17. AI-on-Demand (AIOD) platform and its uses for cultural heritage

Edisa Lozić, ZRC SAZU
Benjamin Štular, ZRC SAZU
Gabriel Gonzalez Castane, Insight SFI Research Centre for Data Analytics

Theme: Education of ML in archaeology

Archaeologists' interest in artificial intelligence (AI) and its subfield of machine learning (ML) is growing exponentially, e.g., (Bickler 2021; Fiorucci et al. 2020; 2022). We agree with the session authors that as archaeologists using ML we need to move beyond isolated case studies and explore how best to combine our efforts. We also share the ultimate goal of building an integrated community of practitioners. This session offers a bottom-up approach to community building, which is almost always the best approach. However, the top-down approach also has its benefits, for example, in providing infrastructure and other resources. We are convinced that the best results are achieved when the bottom-up and top-down approaches can be combined.

One such top-down approach to building an AI community in Europe is the AI-on-Demand (AIOD) platform (https://www.ai4europe.eu). It has been under development since 2019 and has just started a new development cycle under the AI4Europe project (Horizon Europe CORDIS ID 101070000; https://www.ai4europe.eu.), whose mission is to further drive the technical development of the platform and build a supportive community to add value to the AI research community.

The AIOD is intended to serve as a resource to promote European research and innovation in AI. The objective of the platform is to support all solutions and tools that contribute to the ecosystem of excellence and the ecosystem of trust. It will mature to add AI assets and tools that can be used by the wider community to upskill and share knowledge with innovation sectors. It will provide new services and a marketplace for non-experts to experiment with and use AI solutions in their own work.

The AIOD platform is a community resource and its success depends on the active engagement of end users. The platform is therefore open to any individual or organisation interested in technically developing aspects of the platform, leading or contributing to various components, sharing research outputs or simply sharing news updates or information about upcoming events. There are many opportunities to contribute and help shape the direction and features of the AIOD platform in the years to come.

Cultural heritage, and by extension archaeology, was envisaged from the outset as one of the eleven 'industrial verticals' (user domains) of AIOD. However, this is the first time that it has been included as one of the "model verticals" (case studies). The goal of the cultural heritage model vertical is to showcase success stories and how content can be presented most effectively. In addition, we want to identify the needs of cultural heritage in general and archaeology in particular, to identify where new research can be applied, and to help AI researchers to understand the specific challenges of archaeological applications. Ultimately, the aim is to build and foster a cultural heritage community within the AIOD platform.

In the first part of our presentation, we will introduce the AIOD platform and the plans for its development over the next four years. In the second part of the presentation we will focus on the services most suited to archaeology and how they can be used by archaeologists. The key goal of our presentation is to support the aim of the authors of this session, which is to continue building an integrated community of practitioners.

Bickler, Simon H. 2021. 'Machine Learning Arrives in Archaeology'. Advances in Archaeological Practice 9 (2): 186–91. https://doi.org/10.1017/aap.2021.6.

Fiorucci, Marco, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. 2020. 'Machine Learning for Cultural Heritage: A Survey'. Pattern Recognition Letters. http://dx.doi.org/10.1016/j.patrec.2020.02.017.

Fiorucci, Marco, Wouter B. Verschoof-van der Vaart, Paolo Soleni, Bertrand Le Saux, and Arianna Traviglia. 2022. 'Deep Learning for Archaeological Object Detection on LiDAR: New Evaluation Measures and Insights'. Remote Sensing 14 (7): 1694. https://doi.org/10.3390/rs14071694.

160. Creating an additional class layer with machine learning to counter overfitting in an unbalanced ancient coin dataset

Gampe, Sebastian
Tolle, Karsten

Machine Learning (ML) applied to ancient coin data brings different challenges. We have discussed some of them already at our talk at CAA 2022 in Oxford. In this paper we want to focus on the problem of dealing with an extremely unbalanced dataset for training a convolutional neural network (CNN) and the progress we have made since then.

One aim of our current project Data quality for Numismatics based on Natural language processing and Neural Networks (D4N4) is to train and improve a supervised CNN in order to recognize types and/or mints for the "Corpus Nummorum" (CN) dataset. This dataset features about 19,000 different coin types and more than 40,000 coin images from four different ancient landscapes ("Thrace", "Moesia Inferior", "Troas" and "Mysia"). One of our main problems is the coin per type ratio which is approximately two coins per type. Our previous tests have shown that a CNN based model cannot learn a type properly without a threshold of about twenty coins. By applying this threshold less than 200 types remain for training (Gampe 2021) (Gampe and Tolle 2019). So, we established another recognition model based on the coins' mints. Every mint class includes all corresponding coin types which increases our training set by a large number of coins (about 5,700 for the coin types and about 37,000 for the mints). This way we cover 94 of our 120 mints in the dataset.

Although we have achieved good results with the mint approach (78% Top-1 Accuracy) new problems have emerged. Some of our mint classes are too dominant due to a high amount of coin images in the dataset. Mints like "Pergamon" and "Perinthos" come with about 3,000 images while other mint classes just passed the 20 picture threshold. Hence, we discovered an overfitting in the model's results. So our idea was to split these "big" classes into an additional class layer without reducing the number of training images in the dataset. The goal is to create a dataset with a better balance between the classes. In addition, these new classes might be potentially interesting for the domain experts as well. Currently, we work on three approaches for generating the additional class layer:

- The first method is the use of the unsupervised "Deep Clustering" (DC) method (Caron et al. 2019) for the distribution of images to smaller classes of the same mint. DC incorporates a k-means approach which creates pseudo-labels for a CNN. It also needs a predetermined number of clusters which has to be found in several experiments. However, a manual filtering for clusters with images that are as homogeneous as possible will still be necessary.