

Machine Learning and Digital Heritage: The CEPROQHA Project Perspective



Abdelhak Belhi, Housseem Gasmi, Abdelaziz Bouras, Taha Alfaqheri,
Akuha Solomon Aondoakaa, Abdul H. Sadka and Sebti Foufou

Abstract Through this paper, we aim at investigating the impact of artificial intelligence technologies on cultural heritage promotion and long-term preservation in terms of digitization effectiveness, attractiveness of the assets, and value empowering. Digital tools have been validated to yield sustainable and yet effective preservation for multiple types of content. For cultural data, however, there are multiple challenges in order to achieve sustainable preservation using these digital tools due to the specificities and the high-quality requirements imposed by cultural institutions. With the rise of machine learning and data science technologies, many researchers and heritage organizations are nowadays searching for techniques and methods to value and increase the reliability of cultural heritage digitization through machine learning. The present study investigates some of these initiatives highlighting their added value and potential future improvements. We mostly cover the aspects related

A. Belhi (✉) · H. Gasmi · A. Bouras
CSE, Qatar University, Doha, Qatar
e-mail: abdelhak.belhi@qu.edu.qa

H. Gasmi
e-mail: housseem.gasmi@qu.edu.qa

A. Bouras
e-mail: abdelaziz.bouras@qu.edu.qa

A. Belhi · H. Gasmi
DISP Laboratory, University Lumière Lyon 2, Lyon, France

T. Alfaqheri · A. S. Aondoakaa · A. H. Sadka
Brunel University, London, UK
e-mail: taha.alfaqheri@brunel.ac.uk

A. S. Aondoakaa
e-mail: akuha.aondoakaa@brunel.ac.uk

A. H. Sadka
e-mail: abdulsadka@brunel.ac.uk

S. Foufou
Le2i Lab, University of Burgundy, Dijon, France
e-mail: sfoufou@u-bourgogne.fr

to our context which is the long-term cost-effective digital preservation of the Qatari cultural heritage through the CEPROQHA project.

Keywords Cultural heritage · Machine learning · Artificial intelligence · CEPROQHA project · 3D-holosopic imaging

1 Introduction

From generation to generation, cultural heritage tends to represent the best way to express and transfer historical information. Cultural heritage items or artifacts have an important moral value and are mostly considered priceless. However, the risks that these artifacts face nowadays are more serious than ever before due to natural disasters, wars, degradation, etc. Many efforts are spent either by governments or by NGOs in order to prevent such risks. But through these efforts, it was rather clear that the physical preservation of cultural artifacts is a costly and labor-intensive process that not only is ineffective for the long term but may also lead to different kinds of issues. Digital preservation, which was first proposed by product lifecycle management (PLM) tools, was suitable in order to keep track and preserve cultural assets using digital tools, especially with the increasingly approved reliability of IT systems and the extreme high-quality digitization tools (photography, 3D, etc.) [1]. Applying these technologies in the cultural context is not straight forward, unfortunately. Several challenges related to the specificities of cultural content that are directly related to the effectiveness of these digitization approaches have arisen. For example, in the cultural context, the metadata is as important as the asset itself. If an asset history is lost, its value is heavily degraded. Much research was thus undertaken in order to find methods and techniques that can effectively label and annotate heritage assets based on their partially available information. This information can be their visual capture (visual features), partially annotated metadata or a combination of both. Through time, the accuracy of such systems kept improving until reaching very high-precision levels. Thanks to the advancements in computer vision and machine learning, and with the maturity of visual acquisition technologies, multiple heritage institutions began using their 2D and 3D collections in information retrieval and data mining tasks [2].

Overall, the main focus either for researchers or heritage specialists remains to how effectively take advantage of the recent innovations in the field of data science and machine learning in order to empower, increase the value, preserve, and promote cultural heritage. In our context and through the CEPROQHA project, our team aims at providing a framework intended for cultural data preservation through the use of advanced and effective digitization techniques such as the 3D-holosopic imaging framework developed by our Brunel university team. We also aim at integrating cutting edge machine learning technologies in our digitization process mostly to increase the quality of the digitization and also to enrich the cultural assets to make their preservation more effective and sustainable for the long term.

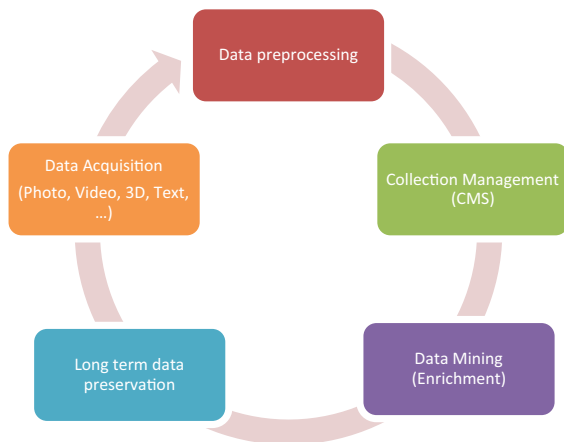
The remainder of this paper is organized as follows. In section two, we present digital heritage where we explain how heritage organizations are leveraging information technology to add value to their collection and the potential applications using the collected data. In section three, we present some of the techniques and tools developed by our team and highlight some on progress and future work related to applications of machine learning in the context of cultural heritage. In section four, we give our conclusion with some perspectives for forthcoming work.

2 Digital Heritage

Digital heritage is the term used for digital information that represents a real physical or moral cultural heritage asset. Nowadays, cultural institutions such as museums rely heavily on digital technologies not only to manage their inventory of assets but also to make their collections more attractive and visible through digital media. Techniques such as 2D and 3D capturing, along with their respective visualization tools, have introduced new methods of content consumption and broadcasting. Digital heritage is widely used not only for entertainment and historical transfer but also for long-term digital preservation and data analytics [3]. Figure 1 outlines the cultural data lifecycle where the input is usually a visual capture of an asset and the output is a digitally preserved copy.

Digital heritage is nowadays widespread across heritage institutions mainly due to the increasing reliability and falling costs of IT systems. As a result, cultural assets are now more accessible for larger audiences than ever before. However, one of the limiting factors is the quality and effectiveness of cultural asset digitization. An asset with unavailable or incomplete metadata is automatically devalued. Consequently, researchers are looking for methods related to recent machine learning techniques to promote and increase the value of underlooked cultural assets.

Fig. 1 Cultural data lifecycle



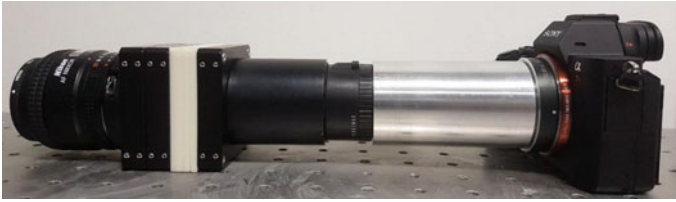


Fig. 2 Sony $\alpha 7$ II camera body fitted with the H3D lens array

3 The CEPROQHA Project Context

In this section, we present the most notable contributions and techniques studied and introduced by the CEPROQHA project for the promotion and the enrichment of digital heritage. These contributions and techniques mostly cover three topics related to cultural data digitization, enrichment, management, and long-term preservation with a focus on machine learning-related techniques. The first topic is related to the improvement at the post-processing stages of the 3D-holoscopic imaging framework. These improvements mainly utilize machine learning techniques such as image super-resolution and video motion interpolation to ensure the cost-effectiveness of the H3D framework while maintaining high-quality standards. In the collection management and enrichment context, our team is working on numerous approaches to complete and curate cultural collections through machine learning-based approaches mainly to save costs and time. Traditionally, heritage institutions have to refer to long-time experts to complete these tasks. Instead, we leverage the power of machine learning in order to annotate, classify, and visually complete missing cultural data [3]. In the following, we present some of the approaches that we designed to tackle the previously mentioned challenges.

3.1 The 3D-Holoscopic Imaging Framework Adapted for the Cultural Content

The 3D-holoscopic technology is not recent. Its principle was proposed in 1908 by Lippmann [4]. The technology is often referred to as lightfield imaging. The principle is inspired by Fly's eyes using an evenly spaced macrolens array fitted to a normal camera (either DSLR or mirrorless) [5, 6]. Each of these lenses captures the scene from a slightly shifted angle in comparison with neighboring lenses in the array. The fundamental principle of H3D is described by Fig. 3. The lightfield data is recorded by the CMOS sensor and stored as a 2D capture. At the display stage, the same process used for capture is reversed. A Macrolens array (MLA) is placed in front of the screen and the object can be reconstructed in space [5, 6]. Figure 2 represents the H3D camera prototype developed by the CMCR Laboratory at Brunel University, London.

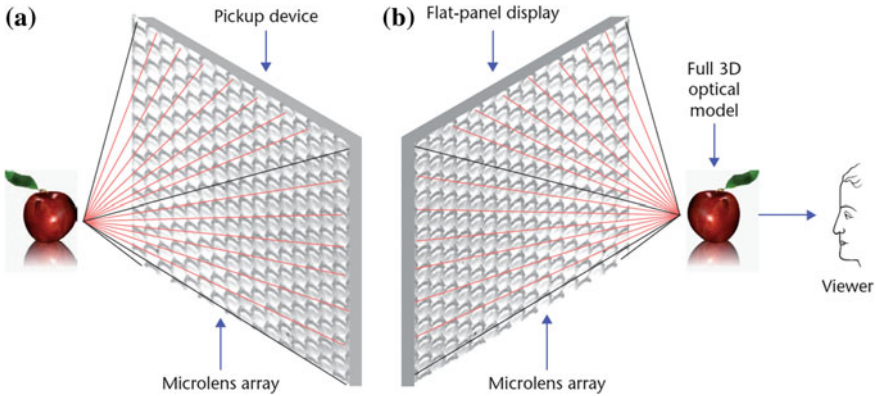


Fig. 3 3D-holoscopic capture and display principle

The principle of the H3D technology and its ability to preserve the depth information with a single capture made it one of the best alternatives to 3D scanning and photogrammetry for the acquisition and display of cultural data due to its cost-effectiveness and ease of use. However, there are still several challenges regarding the requirements of heritage organizations such as the preservation of output quality, etc. The CEPROQHA team focused on adapting the H3D acquisition framework to comply with these requirements, and thus introduced some novelties to the capture; post-processing and display stages of the framework (see Fig. 4) [7].

One of the limitations found in 3D-holoscopic cameras is that the output resolution is smaller in comparison with normal 2D captures. Increasing this resolution through software is thus a must to preserve the main selling point of the H3D technique being its cost-effectiveness. Our team designed and implemented a content adaptation framework to mitigate the low spatial pixel density of H3D captures through using machine learning techniques such as super-resolution to upscale the images.

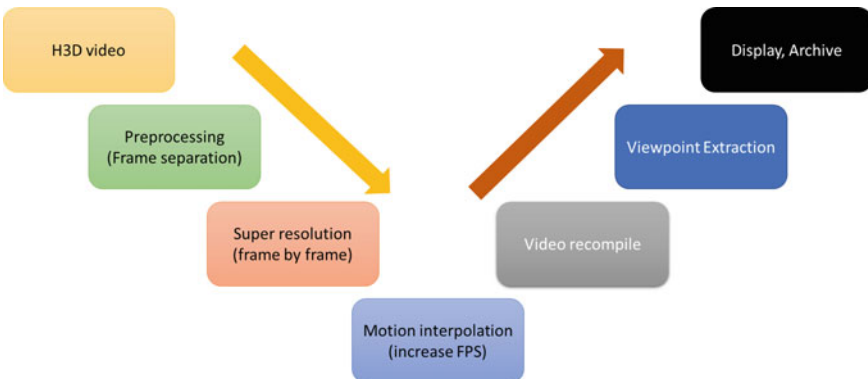


Fig. 4 H3D post-processing framework design

The 360° video scenario was also studied. The video mode of the camera we used has 4 K as the maximum resolution. This is a 1/5 of the 41 megapixels that our CMOS sensor has. We thus developed an adapted 360° video framework that takes 72 still captures shifted by 5° of the assets using the maximum resolution of the camera (41 megapixels). These captures are first upscaled using super-resolution. Then, we apply video motion interpolation to compute the in-between frames. The result is a smooth very high-resolution 360° video that preserves the fine details of the asset.

3.2 Data Analytics for Cultural Data Enrichment and Curation

Cultural data annotation

The CEPROQHA team designed several annotation and classification approaches for multiple scenarios. These approaches mainly focus on either the full annotation of the metadata or its partial annotation. These approaches rely mostly on the visual features and characteristics of cultural assets to predict the desired labels using a combination of deep learning-based approaches such as convolutional neural networks (CNN) and transfer learning. The frameworks we designed achieved excellent performance and were validated on several datasets of paintings collected from multiple museums and institutions such as the Museum of Islamic Art in Doha, Qatar, Wikiart, the Rijksmuseum and the Metropolitan Museum of New York [8].

- **Multi-task Hierarchical classification**

The main intuition behind this approach came up after analyzing the collected datasets and reviewing the related works on cultural heritage classification. We mainly found out that it is inefficient to deal with different types of assets with the same classification model and tools. Indeed, each type of objects has some metadata properties that can easily be predicted when a specific classifier is concisely implemented, but this classifier cannot be generalized to other types of assets using the same approach. For example, the genre and style fields can be found for a painting, but can never be found for a sword even if both are hosted in the same museum or collection. We thus designed a multi-task hierarchical classification that has mainly two stages. In the first stage, a general type classification CNN takes as input the asset image and predicts its type. In the second stage, the asset image is forwarded to the assigned multi-task classifier for that specific type in order to recover the missing metadata (see Fig. 5) [8, 9].

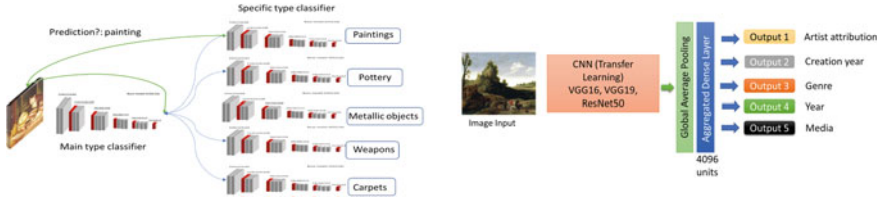


Fig. 5 Hierarchical multi-task classification

- Multimodal classification

Traditional visual classification approaches rely on a single visual capture of a cultural asset to perform prediction on the desired labels. However, in reality, the visual capture of the misannotated asset is generally not the unique information that we have. Often, some labels can be found along this capture and can be used to enrich the input data. This, in fact, is a more realistic scenario which was validated with a handful of heritage institutions. Our team designed and implemented a multimodal classifier for paintings based on convolutional neural networks and transfer learning that takes as input the visual capture of a painting along with the available information at hand. The task of the designed model is to use both visual and textual features to predict the missing labels. We compared this approach with a two-task multi-task classifier with the same output labels and the result was that the multimodal classifier was more effective and yielded a higher validation accuracy across the dataset we used for validation [3] (see Fig. 6).

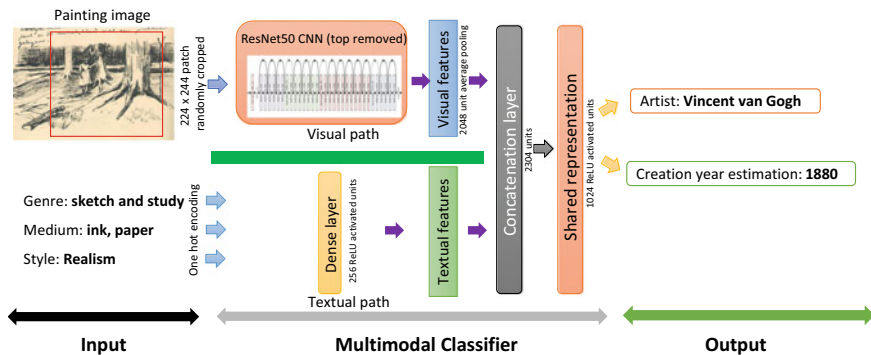


Fig. 6 Our multimodal classifier architecture

Information Extraction for Cultural Ontology Learning

There is a growing need for tools that facilitate the transformation of cultural data to a common form that is shareable and accessible by the domain community as Linked Open Data. Unfortunately, the information manipulated by different organizations is not commonly structured. Each organization has its own data scheme that makes sharing this information with other institutions and organizations difficult. There are several standards for sharing cultural information such as the CIDOC CRM scheme [10]. The manual transformation of this information into a standardized scheme is impractical as it is both time-consuming and labor-intensive. Natural language processing (NLP) techniques are automatic text processing techniques and can help with the metadata scheme transformation. NLP techniques can also help in tackling one of the main limitations that prevent institutions in adopting ontologies as a replacement to standard databases as it is impractical to manually populate ontologies [11]. These activities are usually performed by a domain expert and are labor-intensive and time-consuming. The manual population of ontologies is generally unfeasible except for very small domains. To be practical, the system needs to automate or semi-automate the definition of item metadata from resources such as item descriptions that are available in Web sites, blogs, etc. As these resources are generally available in a free text format, effective methods need to be developed to be able to extract the entities and their relations which are the building block of any ontology. This area of research is called *Ontology Learning* [12].

The two NLP research areas that are today active and mostly relevant to ontology learning are *named-entity recognition* (NER) and *relation extraction* (RE). *Named-entity recognition* (NER) focuses on extracting domain entities from unstructured text such as dates, places, and people names. The definitions of entities are either taken from knowledge sources such as Wikipedia and other sources that are easily and freely accessible or discovered through NLP techniques [13, 14] (see Fig. 7). *Relation Extraction* (RE) on the other hand is concerned with extracting the occurrence of relations in the text which would facilitate the discovery of the relation between the domain entities mentioned in the text [12].

Figure 8 shows an example of the entities extraction and the relation between them from a description of a pot from the Web site of The Metropolitan Museum of Art in New York. Our research concentrates on the transformation of the information available online from museums and cultural heritage institutions. This information is mainly stored as a free text into a more formal structure as an ontology. By representing them as an ontology, this opens the door for many applications such as semantic search, browsing, and visualization, etc.

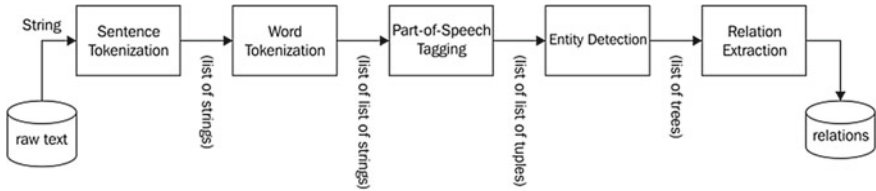


Fig. 7 Typical natural language processing pipeline

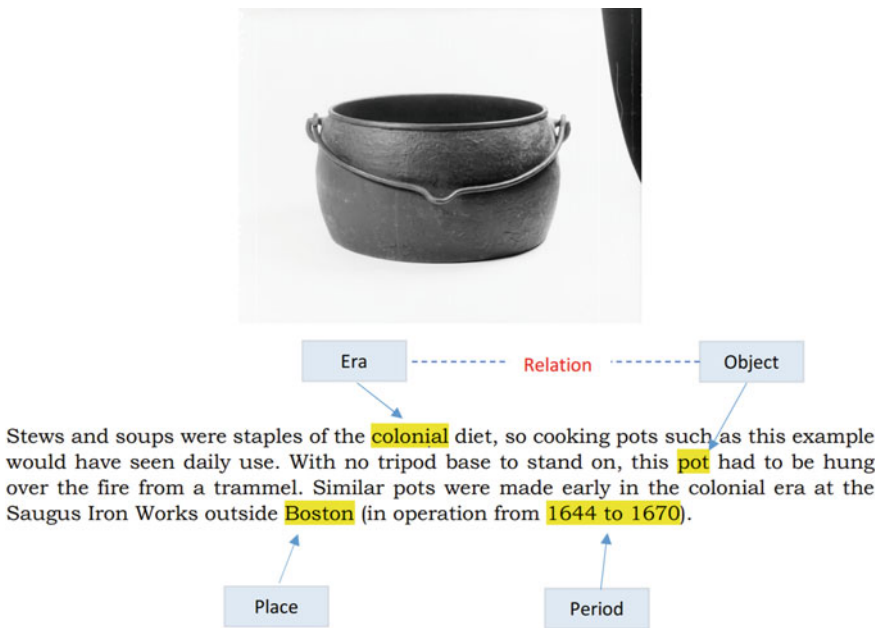


Fig. 8 Named-entity recognition and relation extraction

4 Conclusion

Through this paper, we presented the different challenges and the most important uses of machine learning and data science in cultural heritage context. We presented digital heritage, which represents the set of tools and techniques used to digitize and transfer cultural heritage from the physical to the digital world. This transformation is nowadays necessary as it opens the way to a large spectrum of applications that positively impact cultural heritage. We presented some of the work that was performed by our team in the course of the CEPROQHA project which targets the application of machine learning and data science in the cultural domain. We mostly covered topics such as the automated labeling of misannotated assets, the visual completion of incomplete or damaged assets, and the 3D-holoscopic content adaptation framework, which was improved with super-resolution and motion interpolation. We have also highlighted the importance of using natural language processing to facilitate the adoption of ontologies in cultural heritage context. Some of these techniques still require some work to achieve better maturity and to be ready for real-world use. As future work, we plan to integrate all these tools as well as a tailored digital preservation process in a customized collection management system that links artificial intelligence to cultural heritage.

Acknowledgements This publication was made possible by NPRP grant 9-181-1-036 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors (www.ceproqha.qa).

The authors would also like to thank Mr. Marc Pelletreau, the MIA Multimedia team, the Art Curators and the management staff of the Museum of Islamic art, Doha Qatar for their help and contribution in the data acquisition.

References

1. Belhi, A., Foufou, S., Bouras, A., Sadka, A.H.: Digitization and preservation of cultural heritage products. In: IFIP International Conference on Product Lifecycle Management, pp. 241–253. Springer (2017)
2. Van Noord, N., Hendriks, E., Postma, E.: Towards discovery of the artist's style: learning to recognize artists by their artworks. *IEEE Sig. Process. Mag.* **32**, 46–54 (2015)
3. Belhi, A., Bouras, A., Foufou, S.: Leveraging known data for missing label prediction in cultural heritage context. *Appl. Sci.* **8**, 1768 (2018)
4. Lippmann, G.: Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.* **7**, 821–825 (1908)
5. Swash, M.: *Holoscopic 3D Imaging and Display Technology: Camera/Processing/Display*. Brunel University, London (2013)
6. Aggoun, A., Tseklevs, E., Swash, M.R., Zarpalas, D., Dimou, A., Daras, P., Nunes, P., Soares, L.D.: Immersive 3D holoscopic video system. *IEEE Multimedia* **20**, 28–37 (2013)
7. Alfaqheri, T., Nasri, S.A.E.M., Sadka, A.H.: 3D Holoscopic Imaging for Cultural Heritage Digitalisation. arXiv preprint [arXiv:1810.03916](https://arxiv.org/abs/1810.03916) (2018)

8. Belhi, A., Bouras, A., Foufou, S.: Towards a hierarchical multitask classification framework for cultural heritage. In: 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), pp. 1–7. IEEE 2018
9. Belhi, A., Bouras, A.: Towards a multimodal classification of cultural heritage. In: Qatar Foundation Annual Research Conference Proceedings, pp. ICTPD1010. HBKU Press, Qatar (2018)
10. Doerr, M.: The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. *AI Mag.* **24**, 75 (2003)
11. Gasmi, H., Bouras, A., Laval, J.: LSTM recurrent neural networks for cybersecurity named entity recognition. *ICSEA* **2018**, 11 (2018)
12. Buitelaar, P., Cimiano, P., Magnini, B.: Ontology learning from text: an overview. *Ontology learning from text: methods, evaluation and applications. Front. Artif. Intel. Appl. Ser.* **123** (2005)

13. Belhi, A., Bouras, A., Foufou, S.: Digitization and preservation of cultural heritage: the CEPROQHA approach. In: Software, Knowledge, Information Management and Applications (SKIMA), 2017 11th International Conference on, pp. 1–7. IEEE 2017
14. Nadeau, D., Sekine, S.: A survey of named entity recognition and classification. *Linguisticae Invest.* **30**, 3–26 (2007)