B.8 Processing History. Potentials of Transformers for 3D Reconstruction of Historical Objects with the Help of Artificial Intelligence

Project

Walpola Layantha Perera [1], Heike Messemer [1], Christiane Clados [2], Matthias Heinz [1]
[1] Technische Universität Dresden, Center for Open Digital Innovation and Participation (CODIP), früher Medienzentrum
[2] Philipps-Universität Marburg, Department of Social and Cultural Anthropology

1 Introduction

The digital preservation of cultural heritage is an important and challenging task for the research community. Reconstructing historical objects, which do not exist anymore, in the form of digital 3D models makes it possible to visualize them and present them to the public. The reconstruction process as well as the visualization lead to a deeper understanding of the lost historical objects. But the process of the digital reconstruction is complex and time consuming as diverse sources have to be consulted and interpreted. Therefore, in this paper the latest technology in the field of artificial intelligence (AI) is used to support researchers in the field of Digital Humanities: A Transformer deep learning model based on questions answering methods is introduced to assist to digitally reconstruct historical objects in 3D. It implies a new dimension of data availability, which supports the knowledge process by making large amounts of data qualitatively accessible.

To demonstrate the potential of Transformers two examples of historic objects were selected: the architecture of Sophienkirche in Dresden, a church stemming from the 13th century, which was destroyed in World War II, and a feline incense burner of the Tiwanaku culture, which flourished between AD 200 and 1100 in the Central Andean Highlands of Peru and Bolivia.

First the two historic objects are shortly introduced followed by an overview of the history and state of the art of Transformers. In the next step the research methodology of the project is presented. The following chapter shows the results and evaluates them. Then the educational value of Transformers is explained. The last chapter summarizes and discusses the results of the project and discusses the value of transformers in context with the reconstruction process and the learning environment.

2 Lost architectural heritage: Sophienkirche, Dresden

The Sophienkirche in Dresden coined the cityscape of Dresden for many centuries before being destroyed in World War II (Schreier & Lauffer, 2014). The manifold architectural history of the church begins in the second half of the 13th century. In this period the building was a single-nave church without towers. According to excavations, the church had a length of 43 meters and a width of about 11 meters. As it became too small, it was extended, becoming a two-bay hall church in the middle of the 14th century. The family of Busmann endowed a chapel as their burial place located at the choir in the south. In 1602 the church (Fig. 1) was consecrated to Sophie, widow of elector Christian I.

Figure 1: Sophienkirche in Dresden, photo Verlag A. und R. Adam, around 1910, SLUB/Deutsche Fotothek (No: df_hauptkatalog_0277757)

The most significant architectural modification meant the conversion to a double tower façade in neo-Gothic style under the architect Christian Friedrich Arnold in 1864. Until its destruction in February 1945 the tower's tops were covered with copper in 1932. Fires in the inner of the church caused by air raids led to its complete destruction and exactly one year later the vault collapsed. Only the surrounding walls, the lower parts of the towers and the spire of the southern tower survived. In the following years the fate of the ruin of Sophienkirche was several times in the focus of discussions in Dresden and finally was demolished in 1962–63 ordered by the SED office. The erection of a memorial of Sophienkirche finally began in 2009.

3 In the following chapter a second historic object is presented.Feline Incense Burner of the Tiwanaku Culture, Tiwanaku IV, Central Andean Highlands

Several scholars (Janusek 2004; Korpisaari 2006; Wallace 1957) have argued that the presence of Tiwanaku incense burners or *incensarios* is one of the most dominant traits indicating the presence of Tiwanaku within the Lake Titicaca Basin. Such vessel types are generally characterized by their hyperboloid shape and painted decoration. They consist of a sculpted animal head (cougar, jaguar, llama) attached to one side of the vessel with scalloped rim, an erect tail on the other that also serves as a loop handle, and a pedestal base. All vessels are made of earthenware. They are often elaborately painted on both sides with identical (winged) profile felines. Excavations have revealed deposits that consisted of feline incense burners and other offerings (juvenile llamas, various ornaments) (Delaere, Capriles & Stanish 2019) indicating that feline incense burners belonged to one of the most favored offerings in Tiwanaku society, and are often associated with mortuary practices.

Figure 2: Artefact V A 16735 of the Ancient Americas Collection of the Ethnologisches Museum Berlin, photograph: Christiane Clados

In order to do a digital 3D reconstruction, artefact V A 16735 (Fig. 2) of the Ancient Americas Collection of the Ethnologisches Museum Berlin was provided for analysis (Clados 2020). V A 16735 is a jaguar head and was part of a jaguar incense burner whose container is not preserved. Different sources were used to determine the incense vessel's original appearance. There are numerous objects of comparison in the Ancient Americas Collection of the Ethnologisches Museum in Berlin and the Metropolitan Museum of Art that allow the accurate reconstruction of the missing parts. Archaeological data provide valuable information on its compositional properties (paste, temper), production, provenience and original function; an iconographic analysis contributed to a deeper understanding of form and surface decoration. Based on absolute chronology, the incense burner in question dates to AD 774 to 991.

The questions answering capabilities of Transformer networks in AI by searching given text documents were tested in order to find out how this technology can assist in building a 3D model of an ancient artifact.

4 Overview and State of the Art of Transformers in Relation to this Paper

Vaswani et al. (2017) proposed a new network architecture called Transformer, which is based solely on attention mechanism, replacing the recurrent layers most commonly used in encowiesder-decoder architecture with multi-headed self-attention.

In 1982 and the following years AI and Neural Networks evolved, Recurrent Neural Networks (RNNs) and later on in the late 1980s Convolutional Neural Networks were developed (Wiest al. 2017).

In recent years, there have been further significant discoveries and innovations. Bowman et al. (2016) introduced a recurrent neural network (RNN) based on a variational auto-encoder generative model that incorporates distributed latent representations of entire sentences. This method is capable of producing diverse and well-formed sentences, which later contributed to automated questions answering.

Kusner, Paige & Hernández-Lobato (2017) found that representing molecules and equations by way of their parse tree outperforms text-based representations. This contributed to representation learning and optimization which can be represented as text. This discovery helped to optimize the automated questions answering in text. Hu et al. (2018) generated text sentences by learning disentangled latent representations with designated semantics. They proposed a new neural generative model which combines variational auto-encoders they were able to obtain meaningful sentences. The importance of this type of research lies in the fact that it was able to enhance the semantic analysis of the texts which leads to improve the tense accuracy of the automated answers for the questions. Liu & Lapata (2019) used Bidirectional Encoder Representations from Transformers (BERT) which is a transformer-based machine learning technique for natural language processing pre-training developed by Google to showcase how it can be usefully applied in text summarization and proposed a general framework for both extractive and abstractive models. They introduced a new document level encoder. The abstractive summarization ability of this method, which was built on top of the encoder by stacking several intersentence Transformer layers helping to summarize large text-based documents. The summarized texts are used for answering questions as well. Beltagy, Lo & Cohan (2019) achieved new state of the art results on using Transformers in the scientific domain, which lead to enhanced automated questions and answering work by using scientific papers. Van Aken et al. (2019) provided clues about which parts of the context the Transformer model considered important for answering a question. They fine-tuned the model for answering questions. They discussed how BERT is answering questions after it is fine-tuned. These findings enhanced multiple natural language processing tasks in the area of question-answering by using Transformers.

The research of this paper will continue from here to find out how the Transformer deep learning model can assist in building 3D objects of ancient artefacts and historical buildings by answering questions instantly after reading research papers and articles.

5 Research Methodology – Digital 3D Reconstruction

The architectonic shape of Sophienkirche between 1864 and its demolition was digitally 3D reconstructed with the CAD program SketchUp and finalized with Cinema4D within the research project UrbanHistory4D (HistStadt4D), funded by BMBF in 2016–2021 (Messemer & Clados 2020).[1] It was integrated into a digital 4D model of Dresden, which can be accessed online via an interactive 4D browser with a timeline. As the main sources for the 3D reconstruction served an online accessible publication (Schreier & Lauffer 2014), a floorplan, dated before 1868 (ibid., fig. 8, p. 11), as well as historical photographs from the 19th and 20th century online accessible via the database of Deutsche Fotothek[2]. Measurements of the church stem from the floor plan and a photograph depicting the front of the building. The walls of the choir were reconstructed according to images of the inner part of the ruin after 1945; the design of the roof is hypothetical (see Fig. 3).

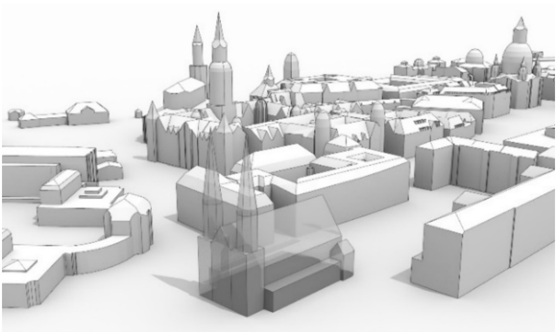

Figure 3: In the foreground 3D model of Sophienkirche, Dresden, depicting the building between 1932 and 1945, 3D model Heike Messemer and Jonas Bruschke, UrbanHistory4D (HistStadt4D), 2019

In the case of the incense burner, only reduced information had been available in order to define its original appearance: a short catalog text (Eisleb 1980) already existed describing the preserved fragment and mentioned it to be part of a Tiwanaku incense burner. This catalog description included cultural affiliation and style, provenience, relative chronology, medium, and dimensions. To restore the original appearance of the container, a comparative analysis was carried out. The main source to reconstruct the original appearance of the incense burner was a nearly identical vessel of the collection of The Metropolitan Museum of Art, Smithsonian Institution.

1 Website of the 4D browser of the junior research group UrbanHistory4D (HistStadt4D): https://4dbrowser.urbanhistory4d.org, accessed on 22. Sept. 2021.

2 Website of Deutsche Fotothek, http://www.deutschefotothek.de/, accessed on 22. Sept. 2021.

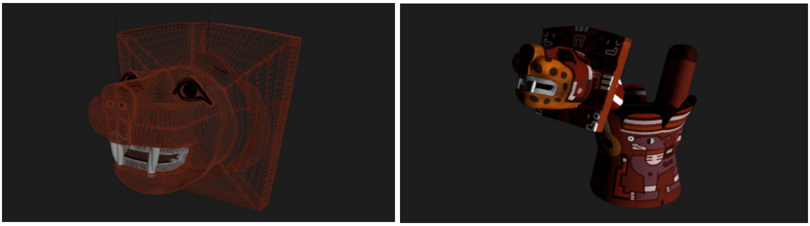

Figure 4: Mesh of 3D reconstructed head of the incense burner (left), and overall appearance without marked hypothetical parts (right)

To 3D reconstruct a true to scale model the 3D graphic software Blender was used (Fig. 4). All photographs were imported and exact measurements were extracted based on MeasureIt. The missing part, the container, was modelled based on photographs of identical incense burners, which have been preserved to this day in different museum collections. The photographs were taken from three heights, then the object was rotated 180 degrees and the entire process was repeated. In the process of texture mapping, the reconstructed parts of the vessel were indicated with less saturation.

6 Research Methodology – Transformer Network

For this research project were used hugging face transformers library and fine-tuned BERT questions answering model, which was built by using Transformers. The next step consisted of fitting the pipeline on corpus using the pre-trained readers. TensorFlow and the following code libraries were used throughout the experiment: Pytorch, Numpy, SKlearn.

Furthermore, BERT for questions concerning the pretrained model was used, which is bert-large-uncased-whole-word-masking-finetuned-squad. Then the BERT Tokenizer was loaded, and then the reference texts were read by the code. In this experiment the reference texts were research papers. At the next step, the code applied the tokenizer to the input text, treating them as a text pair. This was done to tokenize them and to add special tokens in order to create a list of input ids. After that, the code provided segment ids to a model. This was done by noticing the input ids and finding the location of the SEP (separator) token. The next step was to run the questions through the model. The answers were inserted into a textfield.

Figure 5 shows an example of a deep neural network. Such a deep neural network was used in the experiment as well, demanding high computing power. Hence, the experiment was done with a computer which has a Nvidia GTX, Geforce GPU. Moreover, in order to make the process faster CuDNN libraries were installed for the GTX GPU.

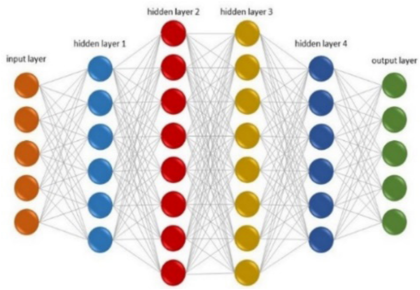

Figure 5: A deep neural network (own source)

7 Results and Evaluation

According to the results the Confusion Matrix calculations have 62.04 % true positives. This means that the system gave true positive answers 62.04 %. The system gave false positive answers 35.46 %. True negative answer rate was 0.96 %. False negative answers rate is 1.54 %. This low rate of false positive answers indicates that the system is considerably good. Figure 6 shows the confusion matrix results, which were generated by the code.

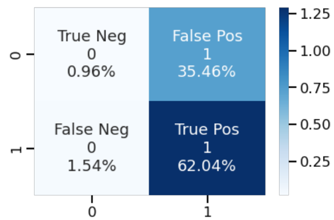

Figure 6: Confusion matrix of the results of Transformers answers to the questions related to the incense burner, and Sophienkirche, Dresden

For the 3D reconstruction of Sophienkirche only one text in German was used as a source of information. To be readable for the Transformer network, it had to be translated into English via DeepL. The questions which were processed included for example: “Where was the Sophienkirche built?”, “When was the double tower façade built?” and “Why was Sophienkirche rebuilt in the middle of the 14th century?”. The first two questions were answered correctly. The third one aimed at a more complex answer (the group of worshippers grew so that the church had to be extended) and was not answered correctly (“danger of decay”, around 1600).

In the case of the Tiwanaku incense burner texts in English were consulted and processed by the Transformer network. The questions were related to form, material, archaeological context, provenience, function and dating of the incense burner. Typical questions such as “What are associated artefacts of Tiwanaku incense burners?”, “Where were Tiwanaku incense burners found/excavated?” or “What is the archaeological context?” were very satisfying answered. It is particularly interesting that questions were even answered correctly when the common Spanish term “incensario” replaced the English term. This indicates that the Transformer network learned to recognize the Spanish term as a surrogate. Questions about the dating of the incense burner were answered less clearly. However, it is very likely that the accuracy of the answers increases with the number of texts analyzed. The following chapter gives a short insight how Transformers and AI in general can enrich educational processes.

8 Educational value of transformers

The described approach about transformer support can be classified in the direction of Explainable AI (XAI), which among other things defends AI-based results that target non-technical users (Díaz-Rodríguez & Pisoni, 2020). In the area of cultural heritage, it can be noted that accessibility has not yet been greatly impacted by the latest AI technology (Díaz-Rodríguez & Pisoni, 2020). Thus, there is still enormous potential here.

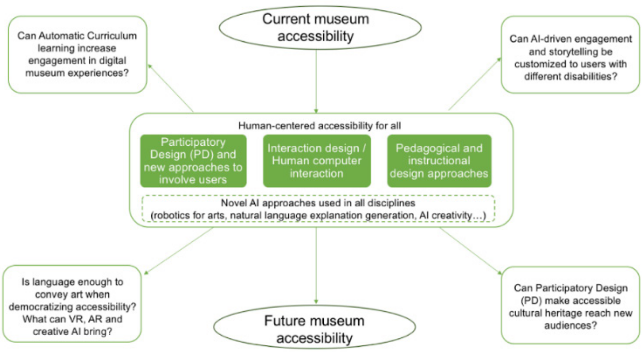

Figure 7: Conceptual framework for AI-enabled accessibility of museums and cultural heritage sites (Pisoni, Díaz-Rodríguez, Gijlers & Tonolli 2021)

When students want to study buildings that no longer exist or do not exist completely, the way to obtain the necessary information is very laborious and time-consuming. Information from various sources such as descriptions of the architecture of the building, plans, sketches, photos, etc. are necessary. These must be obtained, consumed and filtered.

In addition, they must be reassembled at the end in such a way that the building becomes tangible. By integrating Transformers, these processes can be supported and time can be saved. Thus, there is more time for the actual study of the building. Figure 7 shows how AI, and thus Transformers, can enrich educational processes in the cultural domain.

9 Discussion

As was shown a Transformer network was used to support research in the digital humanities in order to create 3D reconstructions of lost historical objects. Two different artefacts were selected: Sophienkirche in Dresden (destroyed 1945 and demolished 1962/63) and a feline incense burner from Tiwanaku dating from AD 774 to 991. For the architecture only one German text, translated into English was used, for the Tiwanaku incense burner several English texts were consulted, some of them also using Spanish terms. The answers showed that they are more likely to be correct and complete the more text was provided.

As the Transformer was not able to process German texts, a future task will be to train it to read other languages than English to open the technology to even more research areas. In the case of the text referring to the incense burner the Transformer already identified a Spanish term replacing the English one. The project proved that the use of Transformer can support the process of 3D reconstruction providing answers of core questions.

The future plan of this research is to automate this process up to ninety percent. This can be achieved by producing software in which 3D model developers can feed text and images data and produce a 3D model. Such software can be built by using Transformers and computer vision in AI. Image retrieval techniques in historical objects detection also can be used in this automation process (Perera et al. 2020).

It was also possible to show the contribution that Transformer and AI can have in the context of educational processes of cultural studies. This technological advance makes it possible to intensify study and research by freeing up more time for new discoveries.

Literature

Beltagy, I., Lo, K. & Cohan, A. (2019). SCIBERT: A Pretrained Language Model for Scientific. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), November 2019, Hong Kong, China, pp. 3615–3620. DOI: https://doi.org/10.18653/v1/D19-1371

Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A. M., Jozefowicz, R., Bengio, S. (2016). Generating Sentences from a Continuous Space. Retrieved from https://arxiv.org/abs/1511.06349 (22.09.2021).

Clados, C. (2020). Digital 3D Reconstruction and Simulation of Fragmentary Museum Artefacts: A Tiwanaku Incense Burner from the Ethnologisches Museum of Berlin. Baessler-Archiv, Vol. 66, 111–122.

Delaere, C., Capriles, J. M. & Stanish, C. (2019). Underwater ritual offerings in the Island of the Sun and the formation of the Tiwanaku state. In Proceedings of the National Academy of Sciences of the United States of America, April 23 2019, 116 (17), pp. 8233–8238. DOI: https://doi.org/10.1073/pnas.1820749116

Díaz-Rodríguez, N. & Pisoni, G. (2020). Accessible Cultural Heritage through Explainable Artificial Intelligence. In Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP ‘20 Adjunct). Association for Computing Machinery, New York, NY, USA, 317–324. DOI: https://doi.org/10.1145/3386392.3399276

Eisleb, D. (1980). Altperuanische Kulturen III; Tiahuanaco. Berlin: Ethnologisches Museum.

Hu, Z., Yang, Z., Liang, X., Salakhutdinov, R. & Xing, E. P. (2017) Toward Controlled Generation of Text. Retrieved from https://arxiv.org/abs/1703.00955 (22.09.2021).

Janusek, J. W. (2004). Identity and Power in the Ancient Andes: Tiwanaku Cities through Time. New York: Routledge.

Korpisaari, A. (2006). Death in the Bolivian High Plateau: Burials and Tiwanaku Society. Oxford: British Archaeological Reports.

Kusner, M. J., Paige, B., Hernández-Lobato, J. M. (2017). Grammar Variational Autoencoder. Retrieved from https://arxiv.org/abs/1703.01925 (22.09.2021).

Liu, Y. & Lapata, M. (2019). Text Summarization with Pretrained Encoders. Retrieved from https://arxiv.org/abs/1908.08345 (22.09.2021).

Messemer, H. & Clados, C. (2020). 3D-Rekonstruktion als Werkzeug der Quellenreflexion. In C. Schöch (ed.): DHd 2020. Spielräume. Digital Humanities zwischen Modellierung und Interpretation. 7. Jahrestagung des Verbands Digital Humanities im deutschsprachigen Raum, 123–126.

Mikaelian, F. (2019). BERT for QA fine-tuned on SQuAD v1.1. Retrieved from https://github.com/cdqa-suite/cdQA/releases/tag/bert_qa (22.09.2021).

Perera, W. L., Messemer, H., Heinz, M. & Kretzschmar, M. (2020). Detecting Treasures in Museums with Artificial Intelligence. In T. Köhler, E. Schoop & N. Kahnwald (eds.), Gemeinschaften in Neuen Medien. Von hybriden Realitäten zu hybriden Gemeinschaften. 23. Workshop GeNeMe'20. Gemeinschaften in Neuen Medien. Dresden, 07.–09.10.2020. Dresden: TUDpress, S. 36–48.

Pisoni, G.; Díaz-Rodríguez, N.; Gijlers, H. & Tonolli, L. (2021). Human-Centered Artificial Intelligence for Designing Accessible Cultural Heritage. Appl. Sci., 11, 870. https://doi.org/10.3390/app11020870

Schreier, D. & Lauffer, M. (2014). Die Sophienkirche. In Verlorene Kirchen. Dresdens zerstörte Gotteshäuser. Eine Dokumentation seit 1938 (2nd edition), Dresden, 8–13.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention Is All You Need. In Proceedings of NIPS 2017. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. Retrieved from https://arxiv.org/abs/1706.03762 (22.09.2021).

Wallace, D. T. (1957). The Tiahuanaco horizon styles in the Peruvian and Bolivian highlands. PhD dissertation. Berkeley: University of California.

van Aken, B., Winter, A. Löser, A. & Gers, F. A. How Does BERT Answer Questions? A Layer-Wise Analysis of Transformer Representations. In CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, November 2019, pp. 1823–1832, DOI: https://doi.org/10.1145/3357384.3358028

Wiest, L. (2017). Recurrent Neural Networks – Combination of RNN and CNN. Retrieved from https://wiki.tum.de/display/lfdv/Recurrent+Neural+Networks+-+Combination+of+RNN+and+CNN (22.09.2021).