B.4 Detecting Treasures in Museums with Artificial Intelligence

Project

Walpola Layantha Perera[1], Heike Messemer[2], Matthias Heinz[1], Michael Kretzschmar[3]
[1] Technische Universität Dresden, Media Centre
[2] Julius-Maximilians-Universität Würzburg, Institute of Art History
[3] Technische Universität Dresden, Unit 7.5 Web and Video

1 Introduction

Museums around the world possess hundreds of thousands of priceless objects, which have stories to tell about human history. While students and scholars study them, even the general public is interested in these stories. If there is a way to automate the information delivery system about these objects it will be of immense value, e.g. it will support students to study these objects and speed up research. Adaptive blended learning options are conceivable, which can perfectly merge digital analysis and on-site viewing. Thus, the preparation and post-processing of studied objects is just as conceivable as the adequate acquisition of information for on-site studies. Examples of such solutions would be mobile apps and computer software that can be used for history and archaeology education as well. However, it is important to identify these objects correctly in order to build such solutions. Computer vision technologies in artificial intelligence (AI) can be used for this. Therefore, this paper will show how AI-algorithms can be used for digital humanities in novel ways, such as for detecting museum treasures.

The objective is to identify objects in museums by using computer vision, and building a dataset of high-resolution images of important artworks which are displayed at the New Green Vault, a part of Dresden Castle. The artworks stem from the courtly collection of Dresden electors, which has its origin around 1560. In the beginning, the collection consisted of mostly scientific and technical instruments and was largely extended over the centuries with objects made of gold, gemstones, coral, ivory and many more exotic and rare materials (Spenlé, 2016, pp. 38–39).1 From an art historian point of view, this collection is exceptional as it contains unique artworks created by well-known and specialized goldsmiths, cabinetmakers and many professions more. The cultural and historical value of the objects is priceless – they are indeed treasures.

1 For more information about the collection of the New Green Vault in Dresden, visit the website of the museum: https://gruenes-gewoelbe.skd.museum/en/ausstellungen/neuesgruenes-gewoelbe/ (17/06/2020).

The dataset created for the project contains 105 objects and 70 images of those objects. These outstanding artworks were selected and photographed on site in their glass cases. It was necessary to capture them to form a dataset for the object recognition process, which will be described in the following. Before showing the results, the computer vision process, including the accuracy of the tests and key performance indicators (KPIs), will be explained. In a second step, benefits of recognizing art works with computer vision in the field of art history and in higher education will be described. At the end, implications and future work will be discussed.

2 Recognizing Treasures in Museums

Dresden is a city in Germany, which has lots of historical buildings and monuments, including great museums with amazing treasures. This wealth of information should be able to be easily accessed and the great stories behind those treasures should be told. Therefore, the cluster Artificial Intelligence & Digital Humanities of the Media Centre at the Technische Universität Dresden built the Dresden Treasures Dataset. Together with the photographer Michael Kretzschmar, the cluster team went to Staatliche Kunstsammlungen Dresden and photographed 105 treasures there. 70 images of that photo session were used to build Dresden Treasures Dataset (fig. 1). Each object in the dataset was labelled by using a CSV file.

Figure 1: Sample of the Dresden Treasures Dataset

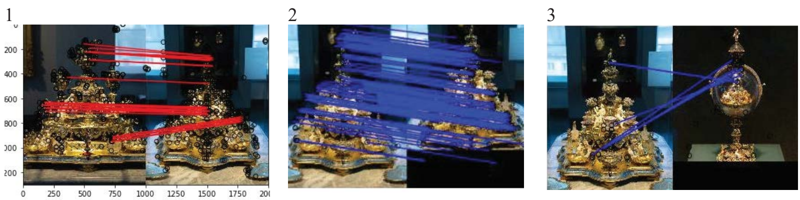
Local feature extraction and description algorithms can be used for computer vision solutions including object recognition and image retrieval (Halavataya, 2020). In 2017 large-scale image retrieval with attentive Deep Local Features (DELF) algorithms was introduced by POSTECH, Korea and Google Inc. (Noh et al, 2017). This algorithm can be used for image retrieval and for automatically annotating images with their visual content by feature extraction processes (Noh et al, 2017). The DELF feature is based on convolutional neural networks (CNNs) (Noh et al, 2017). RANSAC is an iterative approach to estimate parameters of a model from an observed data set that contains outliers (Li & Gans, 2017). This algorithm is important in order to retrieve images that show treasures from a set of image files, which contains images with outliers. DELF and RANSAC are being used in this paper to identify images of a particular treasure and retrieving images from the dataset, which show that treasure.

3 Recognizing Treasures in New Green Vault of Dresden Castle

Computer vision in AI can help to present and interpret big data of images of museum objects. Therefore, an AI based method for detecting and retrieving images of museum objects was tested. To retrieve images of a certain object out of the image data set the following methodology of object recognition in computer vision was used. For five treasures this process is explained in detail.

Detecting The Golden Coffee Set, 1697–1701

Figure 2: The Golden Coffee Set, Dresden, 1697–1701, Johann Melchior Dinglinger, Georg Friedrich Dinglinger, Paul Heermann, Inv.No. VIII 203, Staatliche Kunstsammlungen Dresden, Germany (photo: Michael Kretzschmar)



Comparison 1–2 = 1
Comparison 2–3 = 1
Comparison 1–3 = 1

Figure 3: Retrieving images of the Golden Coffee Set

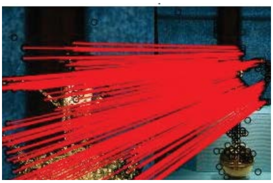
Figure 2 shows an image of The Golden Coffee Set at Staatliche Kunstsammlungen Dresden, Germany, an outstanding artwork of exquisitely worked coffee cups and pots. This highly complex artwork was created between 1697 and 1701 by three different craftsmen: the jeweller Johann Melchior Dinglinger, the enameller Georg Friedrich Dinglinger and the sculptor Paul Heermann. Each of them contributed his specific expertise to create this delicate arrangement of 96 cm height, 76 cm length and 50 cm width with the most precious materials like gold, silver, gemstones, enamel, glass and ivory.

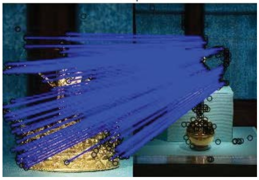
Figure 3 shows correctly retrieving images of The Golden Coffee Set while skipping other images in the images file. Local features of the images were detected and attention based keypoints were selected. Outliers and inliers were realized. In this convolutional neural network-based model one forward pass over the network is sufficient to obtain both key points and descriptors (Noh et al., 2017). After detecting the key points and the descriptors they will be matched with other images in the images folder where the images are retrieved from. The images that have higher inlier counts will be matched with the reference image and retrieved as shown in figure 3. The desired result for test above is 1,0,0 which means that image 1 should be matched with image 2 because they show The Golden Coffee Set. Image 2 should not be matched with image 3, because image 3 does not represent The Golden Coffee Set. Image 1 should not be matched with image 3 either, because as noted above image 3 is not an image of The Golden Coffee Set. However, the predicted result here is 1,1,1 which means there is one wrong image matching, which is image 3.

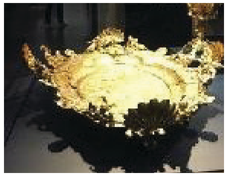
Detecting Crown / Funeral Insignia of Augustus II the Strong, 1733

Figure 4: Crown / Funeral Insignia of Augustus II the Strong (1670–1733), Salomon Wartwald, Warsaw, 1733, Inv.No. P 0340, Staatliche Kunstsammlungen Dresden, Germany (photo: Michael Kretzschmar)

7 8 9







Comparison 7–8 = 1
Comparison 8–9 = 0
Comparison 7–9 = 0

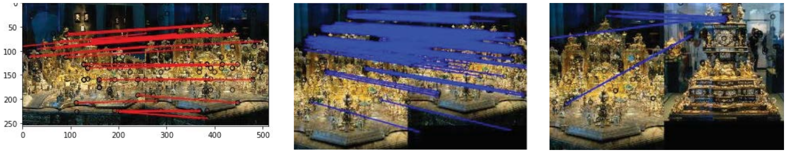
Figure 5: Retrieving images of the Crown / Funeral Insignia of Augustus II the Strong

Figure 4 shows an image of the crown, a funeral insignia of Augustus II the Strong (1670–1733), who became Elector of Saxony in 1694 and King of Poland in 1697. The precious crown was crafted by the goldsmith Salomon Wartwald in 1733 in Warsaw. The artwork is completely worked in brass and weighs 1095 gram, having a diameter of 20.2 cm and a height of 25.5 cm. Figure 5 shows correctly retrieving images of the crown while skipping other 69 images in the images folder. As shown in image 7 and 8 of figure 5 the DELF algorithm with the help of convolutional neural networks is detecting and matching the key points and the image descriptors of the crown of Augustus II the Strong. RANSAC algorithm has been used here to identify inliers and outliers separately. The desired image retrieval result for this test was 1,0,0. The predicted output of the test was 1,0,0. Therefore, in this test the desired output was produced.

Detecting the Throne of Grand Mogul Aureng-Zeb, 1701–1708

Figure 6: The throne of Grand Mogul Aureng-Zeb on its original table, Dresden, 1701–1708, 1721–1723 (table), Johann Melchior Dinglinger, Georg Christoph, Georg Friedrich Dinglinger, Johann Benjamin Thomae, Inv.No. VIII 204, Staatliche Kunstsammlungen Dresden, Germany (photo: Michael Kretzschmar)

10 11 12



Comparison 10–11 = 1
Comparison 11–12 = 0
Comparison 10–12 = 1

Figure 7: Retrieving images showing the throne of Grand Mogul Aureng-Zeb

The throne of Grand Mogul Aureng-Zeb is depicted in the photo of Figure 6 and was created between 1701 and 1708 by the jeweller Johann Melchior Dinglinger, goldsmith Georg Christoph, enameller Georg Friedrich Dinglinger and sculptor Johann Benjamin Thomae. The artwork consists of plenty details delicately worked in gold, silver, enamel, gemstones, pearls and lacquer painting, having a size of 58 cm in height, 142 cm in length and 114 cm in width. Figure 7 shows correctly detecting and retrieving images of the throne of Grand Mogul Aureng-Zeb. However, this time the system was capable of avoiding only 68 wrong images. The system is still retrieving one wrong image. This problem can be solved by using more data, in this case images data.

Detecting Equestrian Statuette of Augustus II the Strong, ca 1728–1730

Figure 8: Equestrian statuette of Augustus II the Strong, Dresden, ca 1728–1730, Johann Michael Weinhold, Inv.No. IX 87, Staatliche Kunstsammlungen Dresden, Germany (photo: Michael Kretzschmar)

13

14

15

Comparison 13–14 = 1
Comparison 14–15 = 0
Comparison 13–15 = 0

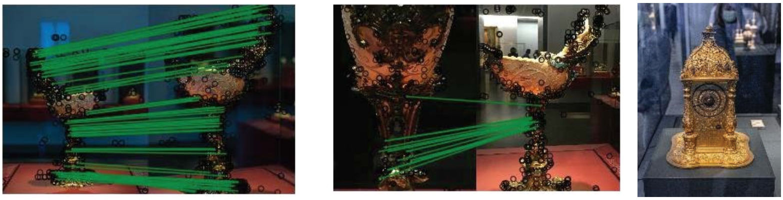
Figure 9: Retrieving images of Equestrian statuette of Augustus II the Strong

In figure 8 the equestrian statuette of Augustus II the Strong dating back to ca 1728–1730 can be seen. Bronze founder Johann Michael Weinhold is the creator of this exquisite work in bronze with a wooden pedestal and a height of 73.3 cm. Figure 9 shows correctly retrieving images of the Equestrian statuette of Augustus II the Strong and skipping other 69 images in the images file.

Detecting Nautilus Cup with Venus, so-called Venus Bowl, between 1704 and 1718

Figure 10: Nautilus cup with Venus, so-called Venus Bowl, Dresden, between 1704 and 1718, Gottfried Döring, Cornelis van Bellekin, Paul Heermann, Inv. No. VI 124, Staatliche Kunstsammlungen Dresden, Germany (photo: Michael Kretzschmar)

16 17 18



Comparison 16–17 = 1
Comparison 17–18 = 0
Comparison 16...70 = 0

Figure 11: Retrieving images of the so-called Venus Bowl

Figure 10 shows the so-called Venus Bowl at Staatliche Kunstsammlungen Dresden. This precious artwork was created between 1704 and 1718 by goldsmith Gottfried Döring, mother of pearl carver Cornelis van Bellekin and ivory carver Paul Heermann. They used rare and exclusive materials like nautilus shell, gold, silver, ivory, enamel and diamonds to work this unique bowl of 40.8 cm of height, 23.5 cm of length and 9.4 cm of depth. As shown in figure 11, 16 to 70 images compared with the images of the Venus Bowl, and the algorithms were able to retrieve the images that depict the Venus Bowl out of those 57 images in higher accuracy. The confusion matrix of the result is shown below.

3.1 Technologies being used for treasure detection and image retrieval

After building the dataset the images were processed through TensorFlow and Keras. By using a reference image other images which match the objects were successfully retrieved, and the images which were not matching were skipped. Therefore, the DELF neural network model published by Google was downloaded from the TensorFlow website[2]. The DELF model takes an image as input and describes noteworthy key points and image descriptors, which can be used for image matching, and image retrieval. The following libraries to process the TensorFlow model with a Python code were used: Matplotlib, a Python plotting library which produces publication quality figures;[3] NumPy, the core library for scientific computing with Python;[4] Scipy, a Python-based system of software for science, mathematics, and engineering;[5] Keras, a neural networks API[6].

2 Retrieved from TensorFlow Hub: https://tfhub.dev/google/delf/1 (01/03/2019)

3 Retrieved from Matplotlib: https://matplotlib.org (09/02/2020)

4 Retrieved from NumPy: https://numpy.org (01/03/2020)

5 Retrieved from SciPy: https://www.scipy.org/ (01/03/2020).

6 Retrieved from Keras: https://keras.io (01/03/2020).

As hardware and GPU-accelerated library to operate the following hardware were used: NVIDIA CUDA Deep Neural Network library (cuDNN) and a GeForce GTX 1050 graphic card which is mounted on an ASUS Republic of Gamers laptop motherboard. cuDNN provides highly tuned implementations for standard routines such as forward and backward convolution, pooling, normalization and activation layers[7].

3.2 Confusion matrix results of image retrieval of the treasures

The results of the confusion matrix are as follows: y_actual = 1,0,0, 1,0,0, 1,0,0, 1,0,0,1,0,0, and y_predicted = 1,1,1, 1,0,0, 1,1,1, 1,0,0,1,0,0. y_actual is the desired result, and y_predicted is the predicted result of all five tests shown above from figure 2 to figure 11.

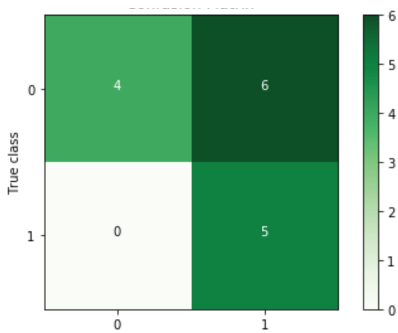


Figure 12: Confusion matrix results of image retrieval of the treasures

The confusion matrix results show that there are: four true negatives which means that the model correctly predicted four negative classes; zero false negatives which means that the model is not predicting any negative classes incorrectly; five true positives which means that the model predicted five positive classes correctly during the above tests; six false positives signifying the indicating that the model six times incorrectly predicted positive classes.

3.3 Key Performance Indicators (KPIs)

```
Precision Score: 0.5555555555555556
Recall Score: 1.0
Accuracy Score: 0.7333333333333333
```

Figure 13: Computer code generated KPIs of the Confusion Matrix

[7] Retrieved from NVIDIA: https://developer.nvidia.com/cudnn (19/03/2020).

Figure 13 shows the confusion matrix results, which were generated by the code. According to the results the Confusion Matrix calculations have 73.3% accuracy. The accuracy is the ratio of: (true positives + true negatives) / (true positives + true negatives + false positive + false negatives) and indicates the fraction of predictions the used model got right. It also shows that it has 55% precision, which indicates the proportion of positive identifications that was actually correct. The precision is the ratio of: true positives / (true positives + false positives). The recall score is 100%. The recall is the ratio of: true positives / (true positives + false negatives) and indicates the fraction of the total amount of relevant classes that were actually retrieved (Manliguez, 2016).

4 Art History and Higher Education – Benefits for Digital Humanities

The examples show that it is possible to recognize valuable historical objects in museums depicted on digital images by computer vision. This opens up new opportunities to analyse artworks for art historian scholars, students and interested stakeholders.

In the field of art history research projects using artificial intelligence are increasingly developed since about the 2010s (Bell & Ommer, 2018; Manovich, 2015, p 14). There are several areas, which can benefit from automated tools and the use of AI as The Alan Turing Institute indicates (2020): “The cultural heritage sector is experiencing a digital revolution driven by the growing adoption of non-invasive, non-destructive imaging and analytical approaches generating multidimensional data from entire artworks. The ability to interrogate this wealth of data is essential to reveal an artist’s creative process, the works’ restoration history, inform strategies for its conservation and preservation and, importantly, present artwork in new ways to the public.”

Object recognition in computer vision can be an important research area for art history as shown in the preceding chapter. Photographs showing artworks can be matched and compared automatically to find images displaying the same object. This technology can be used to identify artworks, to find visually similar objects and to detect characteristic details of artworks – to name just a few. In the case of large datasets this technology is essential to handle huge amounts of images and their data to analyse them.

This also opens up new perspectives for online image databases – important tools for art historians – as they could display, order and filter their collection with innovative visualizations for the users based on computer vision technologies. Furthermore, AI is a technological trend and an emerging technology for academic teaching and learning which will have a significant impact on higher education (Brown et al., 2020). AI elements can be found, for example, in test generators, plagiarism detection systems, word processing software, presentation software, learning management systems, student information systems, library services and admission services (Brown et al., 2020).

But can AI also enrich learning and research processes during studies and not just the systems behind them? For example, learning processes can be enriched by algorithms that measure students' performance metrics and create adaptive, individual, needs-based learning paths (Brown et al., 2020).

The correct identification of the displayed objects by means of computer vision in AI illustrates the possibility of being able to view museum treasures digitally. These technical procedures can open up possibilities for new teaching-learning arrangements. Conceivable are blended learning arrangements, which offer a flexible way of teaching and learning (Jantos et al., 2016). In this way, museum treasures can be studied virtually prior to an on-site visit in order to make the most intensive use of the time spent in the museum by focusing on questions related to the object. The exchange with fellow students and lecturers can be in focus through the digital (joint) preparation. Here the inverted classroom method opens up various learning possibilities for different learning needs (Engel et al., 2017). The follow-up of studies on objects can also be enriched by the virtual possibilities. Also conceivable are augmented scenarios in which presence research is virtually enriched. Students can also view and study the museum treasures virtually together and be in exchange. This computer-supported collaborative learning can provide students with additional support (Breitenstein et al., 2018).

However, with all this enrichment, the use of AI technologies in higher education is not without controversy, especially in systems that use student data and make intervention decisions based on performance metrics (Brown et al., 2020). Furthermore, AI is controversial in the context of data protection, ethics and access to student data (Brown et al., 2020).

5 Conclusion

In museums worldwide innumerable artworks of historical and cultural importance are being studied by researchers, students and the general public. It is of high importance to provide easy access to them in order to study these treasures remotely, too. Furthermore, the objects need correct metadata. Technical applications can decisively help to accelerate these processes and thus to promote research in the fields of (art) history and archaeology. This paper discussed how to identify museum treasures by computer vision.

A dataset containing digital images of treasures from the Staatliche Kunstsammlungen Dresden was built and labelled using a Comma Separated Values (CSV) file. After that, images of these treasures were processed through two algorithms, whose task was to identify objects. Accuracies of the results were compared using Key Performance Indicators (KPIs). This entire process was demonstrated by using computers.

As indicated in this paper, for art history the use of AI and its vast range of technologies can have a huge benefit for research. Especially with the immense increase of data it will become more important to process this amount of information, to analyse, retrieve and present it. Computer vision, object recognition and artificial neural networks have the potential to accelerate art history research and to open up new research questions supported by technology's abilities to handle big data.

In the field of higher education AI starts to play a more and more important role in different kinds of ways. For example, blended learning arrangements have the potential to enrich and foster the students' study of artworks. In virtual environments they can interact with the historical objects and even work collaboratively together in groups and gain a different view on the artworks they visit later on in the museum to study the original. Furthermore, those digital learning applications offer the possibility for students to study the objects on their own pace.

Future work in the domain of treasure recognition and image retrieval can focus on a new algorithm for treasure detection. Moreover, it is possible to use existing algorithms to build museum object exploration apps and software, which are enriched with scholarly information and stories behind those treasures. Those apps and software will be useful tools for art history researchers and students. Furthermore, building a new algorithm for recognizing museum objects will lead to new findings in computer science.

Literature

Bell, P. & Ommer, B. (2018). Computer Vision und Kunstgeschichte – Dialog zweier Bildwissenschaften. In P. Kuroczyński, P. Bell, L. Dieckmann (Eds.), Computing Art Reader: Einführung in die digitale Kunstgeschichte (Vol. 1). Heidelberg, pp. 61–75. arthistoricum.net.

Breitenstein, M., Dyrna, J., Heinz, M., Fischer, H. & Heitz, R. (2018). Identifying success factors for cross-university computer-supported collaborative learning. In Proceedings of IAC in Dresden 2018. Prague, pp. 123–134. Czech Institute of Academic Education z.s.

Brown, M., McCormack, M., Reeves, J., Brook, D. C., Grajek, S., Alexander, B., Bali, M., Bulger, S., Dark, S., Engelbert, N., Gannon, K., Gauthier, A., Gibson, D., Gibson, R., Lundin, B., Veletsianos, G. & Weber, N. (2020).2020 EDUCAUSE Horizon Report. Teaching and Learning Edition. Louisville, CO: EDUCAUSE.

Engel, M., Heinz, M. & Sonntag, R. (2017). Flexibilizing and Customizing Education using Inverted Classroom Model. Information Systems Management, 34(4), 378–389. https://doi.org/10.1080/10580530.2017.1366221

Jantos, A., Heinz, M., Schoop, E. & Sonntag, R. (2016). Creating a scalable Flipped Classroom Arrangement to enable Part Time Studies to Further Human Development and Management Education. In Proceedings of the International Conference “New Challenges of Economic and Business Development – 2016: Society, Innovations and Collaborative Economy”. Riga, Latvia, pp. 289–299. University of Latvia.

Halavataya, K. (2020). Local Feature Descriptor Indexing for Image Matching and Object Detection. Pattern Recognition and Image Analysis, 30(1), 16–21.

Li, Y., & Gans, N. R. (2017). Predictive RANSAC: Effective Model Fitting and Tracking Approach under Heavy Noise and Outliers. Computer Vision and Image Understanding, 161, 99–113.

Manliguez, C. (2016). Generalized Confusion Matrix for Multiple Classes. URL https://www.researchgate.net/publication/310799885_Generalized_Confusion_Matrix_for_Multiple_Classes, DOI: 10.13140/RG.2.2.31150.51523

Manovich, L. (2015): Data Science and Digital Art History. International Journal for Digital Art History, 1, 13–35.

Noh, H., Araujo, A., Sim, J., Weyand, T., & Han, B. (2017). Large-Scale Image Retrieval with Attentive Deep Local Features. In IEEE International Conference on Computer Vision, Venice, pp. 3476–3485. IEEE. doi: 10.1109/ICCV.2017.374.

Spenlé, V. (2016). The Kunst- and Wunderkammer. Origin and development in the Renaissance and Baroque. In G. Laue (Ed.), The Kunstkammer. Wonders are Collectable, Munich, pp. 12–105. Kunstkammer Georg Laue.

The Alan Turing Institute (2020). AI for Arts. How can AI technology shape art investigation, including art history, conservation and preservation, and art presentation? URL https://www.turing.ac.uk/research/interestgroups/ai-arts (26/06/2020).