Revolution or Evolution? AI-Driven Image Classification of Historical Prints

Vignoli, Michela
michela.vignoli@ait.ac.at
AIT Austrian Institute of Technology, Austria

Gruber, Doris
doris.gruber@oeaw.ac.at
Austrian Academy of Sciences, Austria

Simon, Rainer
rainer.simon@ait.ac.at
AIT Austrian Institute of Technology, Austria

Artificial Intelligence (AI) technologies such as Deep Learning and Transfer Learning are rapidly advancing text and image classification in many disciplines (e.g., Lucas et al. 2022, Huang et al. 2022, Kumar et al. 2020). With the progressing digitisation of cultural heritage, AI increasingly finds applications in arts and humanities disciplines as well (Cetinic et al. 2019, Saleh / Elgammal 2016). It is evident that AI technologies open new possibilities for processing and analysing large, heterogeneous historical data corpora in a semi-automated way (Im et al. 2022). But do they have the potential to “overturn intellectual legacies” and to revolutionise the field?

In the project Ottoman Nature in Travelogues (ONiT 2023) we strive to leverage the application of AI technologies to historical research by developing an interdisciplinary methodological framework for the semi-automated, AI-driven analysis of text–image relations. Our object of study is a large multilingual corpus of travelogues printed 1501–1850 (Rörden et al. 2020) that contains representations of Ottoman “nature” (i.e., flora, fauna, landscapes) both in text and image. This short presentation focuses on our first results achieved for improving image classification of historical prints (i.e., mostly woodcuts, engravings, and etchings). We present our approach to robustly identify and classify key image types in our corpus and discuss challenges and lessons learned during this process.

Despite the availability of numerous well-working pretrained image classification and object detection algorithms (e.g., Wightman 2023, Redmon / Farhadi 2018), there are specific challenges related to historical images that we need to address. Most of today’s image classification algorithms are trained on modern photographic content from large benchmark databases (Im et al. 2022: 5871). As our image data mostly consists of prints, the pretrained algorithms cannot be used out of the box.

To address this challenge, we created an annotated dataset for fine-tuning an existing model (“Transfer Learning”, Gullapalli et al. 2021). We explored two approaches: One using a simpler training dataset annotated with multiple labels (“multi-label image classification”, van Strien 2020); and a second one using a more enriched training dataset annotated additionally with the bounding box locations of the objects to be detected (“object detection”, Smirnov / Eguizabal 2018).

The challenge was to select and annotate enough images with distinct examples per class-label to obtain a viable training data set. Since our data contained many duplicates and images lacking good resolution, we had to deal with bias and class imbalance. A second, perhaps bigger, challenge was of methodological nature. Annotations for historical research are usually very detailed, whereas an AI training dataset should only include annotations of distinct visual features to yield useful results. To create a working AI training dataset, we had to adapt our methodology and develop an understanding for visually distinct representations of “nature” as opposed to ones that might require interpretation. We manually labelled a subset of 409 images from our 16th century data with an ICONCLASS-based classification terminology and prepared the training dataset in an iterative review process (ICONCLASS 2023). In total, we annotated 234 representations of animals (on 57% of the images), 177 plants/vegetation (on 43% of the images), 167 landscapes (on 40% of the images), and 22 maps (on 5% of the images). For the object detection task, we enriched our annotated dataset with bounding boxes for each object of the annotated classes (Fig. 1).

First tests indicate that our approach can provide improved solutions for identifying images containing “nature” representations in our corpus. Our experiments suggest that AI will not overcome the disciplinary canons of DH. Rather, by adopting AI technologies we engage in an interdisciplinary process leading to the development of conventions, practices, vocabularies, and trained models suited to computational humanities (Piotrowski 2020, 10-12). Interdisciplinary collaboration between the humanities and AI developers is indeed the key opportunity for bridging the tension between the numerical modelling that is needed to apply AI, and the interpretive methods for analysing the narratives in historical sources. We should refrain from thinking of AI technologies as mere tools to increase the amount of data and the speed with which it can be processed. We have to critically assess our methods for collecting, annotating, and organising data as they have a decisive impact on both performance and bias of our AI classification tools (Smits / Wevers 2022). Finally, we also need to consider the impact that algorithms, models, and quantifications have on our methods and hermeneutics (van Zundert 2015) and implement ethical AI solutions (Johnson et al. 2022).

Bibliography

Cetinic, Eva / Lipic, Tomislav / Grgic, Sonja (2019): “A Deep Learning Perspective on Beauty, Sentiment, and Remembrance of Art”, in: *IEEE Access* 7: 73694–73710. DOI: 10.1109/ACCESS.2019.2921101.

Gullapalli, Ujwal / Chen, Lei / Xiong, Jinbo (2021): “Image Classification with Transfer Learning and FastAI”, in: Xiong, Jinbo / Wu, Shaoen / Peng, Changgen / Tian, Youliang (eds.): *Mobile Multimedia Communications* . Cham: Springer International Publishing 796–806. DOI: 10.1007/978-3-030-89814-4_59.

Huang, Guan-Hua / Fu, Qi-Jia / Gu, Ming-Zhang / Lu, Nan-Han / Liu, Kuo-Ying / Chen, Tai-Been (2022): “Deep Transfer Learning for the Multilabel Classification of Chest X-ray Images”, in: *Diagnostics* 12 (6): 1457. DOI: 10.3390/diagnostics12061457.

ICONCLASS (2023). *Iconclass Illustrated Edition* . https://iconclass.org/ [11.04.2023].

Im, Chanjong / Kim, Yongho / Mandl, Thomas (2022): “Deep learning for historical books: classification of printing technology for digitized images”, in: *Multimedia Tools and Applications* 81 (4): 5867–5888. DOI: 10.1007/s11042-021-11754-7.

Johnson, Marina / Albizri, Abdullah / Harfouche, Antoine / Fosso-Wamba, Samuel (2022): “Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence”, in: *International Journal of Information Management* 64: 102479. DOI: 10.1016/j.ijinfomgt.2022.102479.

Kumar, Pakhee / Ofli, Ferda / Imran, Muhammad / Castillo, Carlos (2020): “Detection of Disaster-Affected Cultural Heritage Sites from Social Media Images Using Deep Learning Techniques”, in: *Journal on Computing and Cultural Heritage* 13 (3): 1–31. DOI: 10.1145/3383314.

Lucas, Luis / Tomás, David / Garcia-Rodriguez, Jose (2022): “Exploiting the Relationship Between Visual and Textual Features in Social Networks for Image Classification with Zero-Shot Deep Learning”, in: Sanjurjo González, Hugo / Pastor López, Iker / García Bringas, Pablo / Quintián, Héctor / Corchado, Emilio (eds.): *16th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2021)* . Cham: Springer International Publishing 369–378. DOI: 10.1007/978-3-030-87869-6_35.

ONiT (2023). *Ottoman Nature in Travelogues* . https://onit.oeaw.ac.at/ [11.04.2023].

Piotrowski, Michael (2020): “Ain’t No Way Around It: Why We Need to Be Clear About What We Mean by ‘Digital Humanities’”, in: SocArXiv. April 14. S. p. https://osf.io/d2kb6 [11.04.2023].

Redmon, Joseph / Farhadi, Ali (2018): “YOLOv3: An Incremental Improvement”, in: ArXiv. April 8. S. p. https://arxiv.org/abs/1804.02767 [11.04.2023].

Rörden, Jan / Gruber, Doris / Krickl, Martin / Haslhofer, Bernhard (2020): “Identifying Historical Travelogues in Large Text Corpora Using Machine Learning” in: Sundqvist, Anneli / Berget, Gerd / Nolin, Jan / Skjerdingstad, Kjell Ivar (eds.): *Sustainable Digital Communities* . Cham: Springer International Publishing 801–815. DOI: 10.1007/978-3-030-43687-2_67.

Saleh, Babak / Elgammal, Ahmed (2016): “Large-scale Classification of Fine-Art Paintings: Learning The Right Metric on The Right Feature”, in: *International Journal for Digital Art History* 2. DOI: 10.11588/DAH.2016.2.23376.

Seydlitz, Melchior von (1580): *Gründtliche Beschreibung Der Wallfart nach dem heiligen Lande [...].* Görlitz: Ambrosius Fritsch. http://data.onb.ac.at/rep/1062FBA6 [11.04.2023].

Smirnov, Stanislav / Eguizabal, Alma (2018): “Deep learning for object detection in fine-art paintings”, in: *2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo)* . Cassino FR, Italy: IEEE. 45–49. DOI: 10.1109/MetroArchaeo43810.2018.9089828.

Smits, Thomas / Wevers, Melvin (2022): “The agency of computer vision models as optical instruments”, in: *Visual Communication* 21 (2): 329–349. DOI: 10.1177/1470357221992097.

van Strien, Daniel (2020): “Image labeling vs classification models”, in: *Daniel van Strien* . October 12. https://danielvanstrien.xyz/models/labels/loss%20functions/2020/10/12/labelling_vs_classification_models.html [11.04.2023].

Van Zundert, Joris J. (2015): “Screwmeneutics and Hermenumericals: The Computationality of Hermeneutics” in: Schreibman, Susan / Siemens, Ray / Unsworth, John (eds.): *A New Companion to Digital Humanities* . Chichester, UK: John Wiley & Sons, Ltd 331–347. DOI: 10.1002/9781118680605.ch23.

Wightman, Ross (2023): “PyTorch Image Models”, in: GitHub. April 6. https://github.com/rwightman/pytorch-image-models [11.04.2023].