

Image Classification of Paris Bible Data

Abo Bakir Shuan, Yasmin Hawar

y.hawarshuan@gmail.com

University of Southern Denmark, Denmark

Meinecke, Christofer

cmeinecke@informatik.uni-leipzig.de

Leipzig University, Germany

Jänicke, Stefan

stjaenicke@imada.sdu.dk

University of Southern Denmark, Denmark

The Paris Bible holds significant importance as a primary source for examining the religious attitudes and customs of Medieval Europe. These richly illustrated manuscripts were produced in thousands and spread throughout the continent. Eventually, they became the template for the Bible as we know it today. More than ten thousand original copies are currently preserved in libraries, museums, and in private collections around the world, of which a share have been digitized (Boillet et al., 2019). This digitization has created opportunities for interdisciplinary, computationally based research, and large scale, macro level analysis, for instance by use of machine learning techniques.

The effectiveness of advanced supervised machine learning techniques hinges on the availability of sizable amounts of data with accurate labeling. They rely heavily on principles that may seem at odds with humanistic methods. Particularly, they depend on disambiguation on every level, from the encoding to the structuring of information (Burdick et al., 2012). However, this is a challenging prerequisite when dealing with humanities data, which is often characterized by a high level of ambiguity. The analysis of visual art through computational methods is susceptible to several sources of bias due to the unique and heterogeneous nature of the artworks (Zhang et al., 2022). Thus, when digitized collections are accompanied by meta-data prepared by domain experts from different institutions, it inevitably leads to ambiguity and inconsistency across collections, making them difficult to merge for the application of computational methods (Meinecke et al., 2022).

This issue is at the core of ongoing research by Meinecke et al. who aim to bridge the gap between digitized cultural heritage collections from different sources and supervised machine learning practices, suggesting a visualization-based solution to aid domain experts in streamlining conflicting meta-data, using the *Initiale* and the *Mandragore* digitized Paris Bible collections.

Given the evident interest in interdisciplinary collaboration between the study of cultural heritage collections and large-scale computational analysis, our objective was to evaluate the feasibility of utilizing supervised learning methods for image classification of the *Initiale* digital Paris Bible collection. To accomplish this, we collaborated with domain professionals from The Paris Bible Project¹ who were specifically interested in classifying images according to biblical figures appearing in illustrations and their hand gestures. We utilized pre-trained convolutional neural networks and data balancing techniques to undertake three

separate classification tasks for the *Initiale* digitized Paris Bible collection:

- multi-class classification of the six prominent biblical figures Saint Paul, King David, Saint Jean, King Salomon, Saint Jerome, and Moses.
- multi-class classification of saints, prophets and kings.
- multi-label classification of five hand gestures: open hand, index pointed, covered hand, hand on chest, and folded hands.

The images in *Initiale* collection vary in terms of quality and layout, with some being full-page scans or double-page scans and others being close-ups of illustrations within a page. This characteristic poses a challenge when creating random subsets for training, validation, and testing. As for the meta-data, the Bible, just as any other book, follows a character hierarchy, with more prominent biblical figures appearing in illustrations more frequently, which results in skewness in the true distribution data. Furthermore, the issue of varying and infrequent classes arises due to the images being partially annotated.

As a pre-processing step, we found it necessary to remove any classes with a frequency under 100 instances. We applied stratified sampling methods to obtain training, validation, and test splits where the distribution of classes reflected that of the original data frame. Finally, to address the problem of class imbalance during training, we tested the Synthetic Minority Over Sampling Technique (Chawla, 2002) and training with weighted classes. We found that our models were most successful in performing multi-class classification of saints, prophets, and kings with an accuracy score of 66%. For classification of six prominent biblical figures, our best performing model successfully classified 53% test instances. Finally, in the multi-label scenario of classifying hand gestures, our most successful model achieved a hamming score of 60%. While the differences between models trained with or without accounting for class imbalance were mostly marginal, global measures were best when models were trained without balancing measures. However, class-wise performance improved for minority classes when training with weighted classes.

Previously mentioned characteristics and limitations of our data together with our results clearly pointed towards a mismatch between the supervised learning paradigm and the highly complex nature of Paris Bible data. We suggest looking towards active learning instead. While supervised learning models require large amounts of annotated data, active learning models only require a small amount of training data. The key assumption behind active learning is that a machine learning algorithm can perform better with less training data if it is allowed to choose the data from which it learns (Wu et al., 2020). However, with the *Initiale* meta-data containing over 1700 unique annotations with number of annotations per image ranging from 1 to over 60, it will require experts to direct their attention towards a smaller subset of labels.

Active learning gives room for focused, systematic annotation of smaller subsets of data at a time which makes annotation more manageable and less error prone. While human effort still plays a significant role in active learning, the process allows for learning, sampling, and annotation to occur at the same time, which can ultimately result in better performance with the same amount of human labeling efforts (Wang and Hua, 2011), potentially leading to a better return on investment for domain experts.

We believe that our process along with the challenges we encounter, demonstrate the highly non-trivial nature of the intersection between humanities data and computational methods. Additionally, it serves as a practical demonstration of the applied machine learning process and its requirements for domain experts

within humanities, who wish to motivate the application of digital humanities methods within their area of expertise.

Notes

1. The Paris Bible Project was founded during the 2019 pandemic, as a consequence of worldwide lockdowns of public spaces including museums, libraries, and educational institutions. The group, who have academic backgrounds within Digital Humanities, Medieval Studies, and Art History, study the production and diffusion of medieval Latin Bibles in Europe.

Bibliography

Boillet, Mélodie / Bonhomme, Marie-Laurence / Stutzmann, Dominique / Kermorvant, Christopher (2019): "HORAE: an annotated dataset of books of hours", in: *Proceedings of the 5th International Workshop on Historical Document Imaging and Processing*.

Burdick, Anne / Drucker, Johanna / Lunenfeld, Peter / Presner, Todd / Schnapp, Jeffrey (2016): *Digital Humanities*. The MIT Press.

Chawla, Nitesh V. / Bowyer Kevin W. / Hall, Lawrence O. / Kegelmeyer William P. (2002): "Smote: Synthetic minority over-sampling technique" in: *Journal Of Artificial Intelligence Research* 16, 321-357.

Meinecke, Christofer / Guéville, Estelle / Wrisley, David W. / Jänicke, Stefan (2022): "A visual analysis framework for composing a hierarchical classification for medieval illuminations" *Submitted to Digital Scholarship in the Humanities (DSH)*.

Wang, Meng / Hua, Xian-Sheng (2011): "Active learning in multimedia annotation and retrieval: A survey" in: *ACM Transactions on Intelligent Systems and Technology* 2, 2, 10: 1–21.

Wu, Jian / Sheng, Victor S. / Zhang, Jing / Li, Hua / Dadakova, Tetiana / Swisher, Christine L. / Cui, Zhiming / Zhao, Pengpeng (2020): "Multi-label active learning algorithms for image classification: Overview and future promise" in: *ACM Computing Surveys* 53, 2.

Zhang, Zhuomin / Li, Jia / Stork, David G. / Mansfield, Elizabeth C. / Russell, John / Adams, Catherine / Wang, James Z. (2022): "Reducing Bias in AI-based Analysis of Visual Artworks" in: *IEEE BITS the Information Theory Magazine, Special Issue on Information Processing in Arts and Humanities* 2, 1: 36-48.