

# Determining the Influence of Papyrus Characteristics and Data Augmentation on Fragments Retrieval with Deep Metric Learning

In practice, ancient papyri are frequently torn into several fragments, and the task of papyrologists is to assemble and decipher these fragments. Once successfully reconstructed, ancient papyrus offers the opportunity to gather crucial information about past times. However, reassembling by hand is time-consuming because fragments differ in color, structure, and shape. In other words, they do not fit together perfectly - like an artificially designed toy puzzle.

The objective of this thesis is to make the work of papyrologists easier and increase their efficiency by partially automating the reassembling process. To this end, an algorithm is designed to infer a smaller sub-selection of fragments with a high likelihood of being a potential fit. In the following, this algorithm is called puzzle-helper.

In a real-world scenario, a puzzle-helper must fulfill several constraints. First, training such a helper cannot take advantage of a big data set of fragmented and reassembled (labeled) papyri because only a tiny bit of such data exists. Deep Metric Learning (DML) has been used successfully to create puzzle helpers, [1, 2]. Further improvement of DML models requires an immense amount of data. Otherwise, the model will probably not converge. Labeling manually makes the puzzle-helper useless because the total time required by a papyrologist increases. Therefore, the task of creating labeled data has to be automated. In this thesis, two possibilities will be used to alleviate problems that are arising within the low-data regime:

1. Generative-Adversarial-Networks (GANs) are a possibility to generate more artificial data and thereby improve the overall performance in different domains [3].
2. Creating an extensive data set by automatically and artificially tearing papyri into several pieces is plausible as well [4]. More fragments lead to a significant increase in overall model performance [2, 1].

Another constraint is, of course, that the ground truth fragment has to be included in the subset of potential candidates. If the ground truth match is not enclosed, a papyrologist has no chance to identify the correct fragment unless he extends his search again to the entire data set. Another disadvantage is that the approach fails silently. The papyrologist might search long for a match if the expert is convinced that the algorithm works fine. Separating foreground and background is challenging because a fiber often has similar colors as the text. Therefore, several methods will be used and evaluated on the overall accuracy. A separate evaluation is not straightforward and is misleading because we do not have pixel-wise ground truth [5]. Also, the focus of this thesis lies on the influence of specific characteristics and the power of more labeled data available instead of the separating process itself.

Further experiments will examine whether specific characteristics (fibers) can be exploited. Specifically, it will be tested if the fibers can be used to determine the exact position of the fragments to each other. Determining the position is the next logical step towards a fully automated helper. The final goal is to design a puzzle solver function by combining the DML model and a suitable method for position determination.

In summary, the thesis is divided into the following milestones:

- I. Separating text and papyrus fibers by using state-of-the-art binarization [5] and inpainting [6].
- II. Generating larger amounts of data through a GAN-based model and creating a base-line out of artificial torn fragments for evaluation. [7, 3].
- III. Evaluation by means of a DML model using the original data in the original state (text and fibers), in the processed state (text only or fibers only). Also, an evaluation of the artificial data together with the original data in the original state and processed state.
- IV. Review of the state-of-the-art on the possibilities of using certain characteristics for position determination.

The following research questions emerge from the milestones:

- RQI:** Does the accuracy of the puzzle-helper differ significantly based on the binarization and inpainting methods used?
- RQII:** Does the accuracy of the puzzle-helper differ significantly when only the text or only the fibers are used as input as opposed to the unprocessed data?
- RQIII:** Does the accuracy of the puzzle-helper differ significantly when additional artificial data generated by a GAN are used for training?

The implementation will be done in Python.

*Supervisors:* Dr.-Ing. V. Christlein, Prof. Dr.-Ing. habil. A. Maier, Mathias Seuret M. Sc.

*Student:* Timo Bohnstedt

*Start:* November 8th, 2021

*End:* April 8th, 2022

## References

- [1] C. Ostertag and M. Beurton-Aimar, “Using graph neural networks to reconstruct ancient documents,” in *Pattern Recognition. ICPR International Workshops and Challenges*, A. Del Bimbo, R. Cucchiara, S. Sclaroff, G. M. Farinella, T. Mei, M. Bertini, H. J. Escalante, and R. Vezzani, Eds. Springer International Publishing, pp. 39–53.
- [2] A. Pirrone, M. Beurton-Aimar, and N. Journet, “Self-supervised deep metric learning for ancient papyrus fragments retrieval,” vol. 24, no. 3, pp. 219–234. [Online]. Available: <https://doi.org/10.1007/s10032-021-00369-1>
- [3] M. O. Tamrin and M. Cheriet, “Simultaneous detection of regular patterns in ancient manuscripts using GAN-based deep unsupervised segmentation,” in *Pattern Recognition. ICPR International Workshops and Challenges*, A. Del Bimbo, R. Cucchiara, S. Sclaroff, G. M. Farinella, T. Mei, M. Bertini, H. J. Escalante, and R. Vezzani, Eds. Springer International Publishing, pp. 279–291.

- [4] M. Seuret, A. Nicolaou, D. Rodríguez-Salas, N. Weichselbaumer, D. Stutzmann, M. Mayr, A. Maier, and V. Christlein, "ICDAR 2021 competition on historical document classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, S. U. Josep Lladós, Daniel Lopresti, Ed., vol. 12824 LNCS. Springer Science and Business Media Deutschland GmbH, pp. 618–634.
- [5] C. Tensmeyer and T. Martinez, "Historical document image binarization: A review," vol. 1, no. 3, p. 173. [Online]. Available: <https://doi.org/10.1007/s42979-020-00176-1>
- [6] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," vol. abs/1804.07723. [Online]. Available: <http://arxiv.org/abs/1804.07723>
- [7] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," eprint: 1406.2661.