

# Enhancing Historical Documents: Deep Learning and Image Processing Approaches

Zahra Ziran<sup>1,\*</sup>, Massimo Mecella<sup>1</sup> and Simone Marinai<sup>2</sup>

<sup>1</sup>Sapienza University of Rome, Rome, Italy

<sup>2</sup>University of Florence, Florence, Italy

## Abstract

Historical documents from Late Antiquity to the early Middle Ages often suffer from degraded image quality due to aging, inadequate preservation, and environmental factors, presenting significant challenges for paleographical analysis. These documents contain crucial graphical symbols representing administrative, economic, and cultural information, which are time-consuming and error-prone to interpret manually. This research investigates image processing algorithms and deep learning models for enhancing these historical documents. Using image processing techniques, we improve symbol readability and visibility, while our deep learning approach aids in reconstructing degraded content and identifying patterns. This work contributes to improving the quality of historical document analysis, particularly for graphical symbol interpretation in paleographical studies.

## Keywords

Historical Document, Image Enhancement, paper formatting, Digital Paleography, Graphical Symbols

## 1. Introduction

The field of paleography has experienced significant advancement through digital technologies, particularly in the analysis and preservation of historical documents [1]. The digitization of ancient manuscripts has transformed how humanities scholars access and study historical materials, enabling worldwide collaboration and more sophisticated analytical approaches [2]. However, the quality of digital images remains a critical factor in paleographical research, directly impacting scholars' ability to interpret and analyze historical content accurately [3, 4]. Recent years have witnessed the emergence of various computational methods to enhance the quality of historical document images [5]. These approaches, ranging from traditional image processing techniques to advanced deep learning models, aim to address challenges such as degradation, poor preservation, and environmental damage that affect historical documents [6]. The success of these enhancement methods depends heavily on understanding specific requirements in humanities research, where scholars may need to focus on different aspects of documents, from broad textual content to minute graphical details [7]. The collaboration between computer scientists and humanities scholars has become increasingly important in developing effective enhancement solutions [8]. This interdisciplinary approach requires careful consideration of both technical capabilities and scholarly needs, particularly in preserving and enhancing historical documents' subtle features and symbols [9]. In our study, we focus on improving the quality of graphical symbols in historical documents through two distinct approaches: image processing algorithms and deep learning models. These methods aim to enhance symbol clarity while preserving their historical integrity and significance.

---

IRCDL 2025: 21st Conference on Information and Research Science Connecting to Digital and Library Science, February 20-21 2025, Udine, Italy

\*Corresponding author.

✉ zahra.ziran@uniroma1.it (Z. Ziran); massimo.mecella@uniroma1.it (M. Mecella); simone.marinai@unifi.it (S. Marinai)

ORCID 0000-0003-3142-3084 (Z. Ziran); 0000-0002-9730-8882 (M. Mecella); 0000-0002-6702-2277 (S. Marinai)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 2. Dataset

Several datasets are available for evaluating enhancement and analysis techniques applied to historical documents. These include graphic symbols from documentary records dating from Late Antiquity to the early medieval period [10] and the Digital Image Database of Ancient Handwritings (DIDA) [11, 12], which provides a collection of historical handwritten digit images. Another significant resource is the IAM Historical Document Database (IAM-HistDB) [13], which contains manuscripts from medieval times, including the Saint Gall database of handwritten historical documents from the 9th century. Additionally, the DIVA-HisDB [14] provides 150 annotated pages from three different medieval manuscripts with challenging layouts, specifically designed for evaluating various Document Image Analysis (DIA) tasks. All these datasets serve as valuable benchmarks for assessing document enhancement methods.

In this research, we utilize the dataset of the project (*NOTAE*): *NOT A writtEn word but graphic symbols* [10], which comprises graphic symbols extracted from documentary sources dating from late antiquity to the early Middle Ages. These symbols, encompassing alphabetic and non-alphabetic signs, were employed in various documents to convey meanings beyond the written text, serving functions such as authentication, authorization, or annotation. The documents originate from diverse European regions and contexts, reflecting their time's administrative, legal, and cultural practices. The massive number of documents and symbols provides a rich resource for analyzing the use and evolution of graphic symbols in historical manuscripts and is curated to ensure high-quality representations of these symbols, facilitating detailed analysis and interpretation.

## 3. Related Work

In historical document processing, significant advancements have been made in enhancing and analyzing ancient manuscripts through image-processing techniques and deep learning approaches. This section reviews key studies relevant to document image enhancement, highlighting their strengths, limitations, and how our approach builds upon these methods. Traditional image-processing techniques have played a crucial role in document enhancement, particularly for restoring degraded historical texts. Atanasiu and Marthot-Santaniello [1] employed histogram equalization, adaptive histogram equalization, and local Laplacian filters to enhance papyri legibility. Their method preserved symbol clarity but struggled with complex background noise, requiring manual refinement. Shi and Govindaraju [15] introduced background estimation and pixel normalization to improve aged documents with uneven illumination. While effective, it lacked adaptability for highly degraded manuscripts with missing symbols or extensive fading. Similarly, Mittal et al. [16] used stretch limits, histogram modifications, and saturation adjustments for graphical symbol enhancement but required parameter tuning, limiting scalability. Gupta et al. [17] applied error diffusion and multiresolution binarization for restoration, improving binarization quality but struggling with complex multi-layered text, which deep learning models address more effectively. With deep learning advancements, newer methods have addressed traditional limitations. Oliveira et al. [18] demonstrated the effectiveness of convolutional neural networks (CNNs) in reconstructing degraded symbols and identifying repetitive patterns in ancient manuscripts. While improving restoration, their model required extensive labeled training data, often scarce for historical texts. Zhang et al. [19] developed a deep learning model for graphical symbol enhancement, recovering deteriorated details but lacking adaptability for multi-script textual restoration. Our study integrates traditional image processing with deep learning to enhance symbol clarity and reconstruct missing details, leveraging the strengths of both approaches for a more robust document enhancement framework.

## 4. Methodology

This section explores two distinct image processing approaches alongside a deep learning method to enhance the quality of graphical symbols. These techniques aim to improve readability, enhance

contrast, mitigate illumination variations, and uncover hidden details. By comparing the effectiveness of different methods, we identify the most suitable approaches for the specific goal of graphical symbol enhancement, enabling the extraction of valuable insights and fostering a deeper understanding of these ancient symbols.

## 4.1. Image Processing-Based Models

Image processing techniques play a crucial role in enhancing image quality and addressing various challenges across different datasets. For ancient graphical symbols from the medieval period, edge detection and superpixel segmentation are employed to reveal hidden details and preserve intricate patterns. Histogram equalization and adaptive histogram equalization enhance contrast and address illumination variations, resulting in more visible and interpretable symbols. Model A and B, described in the following, implement these ideas.

### 4.1.1. Model A: Edge Detection and Segmentation Approach

This model implements a comprehensive enhancement pipeline combining edge detection with advanced segmentation techniques. The enhancement process consists of three main stages: edge detection, superpixel segmentation, and histogram optimization. In the initial stage, we employ the Sobel operator for edge detection [20]. For a given grayscale image  $I(x, y)$  defined over a two-dimensional spatial domain  $\Omega \subset \mathbb{R}^2$ , where  $(x, y)$  represents the spatial coordinates, the horizontal and vertical gradients are computed as:

$$G_x(x, y) = \frac{\partial I}{\partial x} \approx I(x + 1, y) - I(x - 1, y) \quad (1)$$

$$G_y(x, y) = \frac{\partial I}{\partial y} \approx I(x, y + 1) - I(x, y - 1) \quad (2)$$

where  $G_x$  and  $G_y$  denote the approximated gradients in the horizontal and vertical directions, respectively. The gradient magnitude  $G(x, y)$  is then computed as:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (3)$$

Following edge detection, we implement the Simple Linear Iterative Clustering (SLIC) algorithm [21] for superpixel segmentation. For each pixel  $p_i = (x_i, y_i, R_i, G_i, B_i)$ , where  $(R_i, G_i, B_i)$  represents its color values in RGB space, the distance metric  $D$  to a cluster center  $c_k = (x_k, y_k, R_k, G_k, B_k)$  is defined as:

$$D(p_i, c_k) = \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2 m^2} \quad (4)$$

where:

- $d_c = \sqrt{(R_i - R_k)^2 + (G_i - G_k)^2 + (B_i - B_k)^2}$  is the color distance
- $d_s = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$  is the spatial distance
- $S$  represents the sampling interval in the pixel grid
- $m$  is a compactness parameter controlling the relative importance of spatial proximity

The final enhancement stage employs adaptive histogram equalization. For each intensity level  $i$ , we compute the cumulative distribution function  $F(i)$ :

$$F(i) = \sum_{k=0}^i p(k), \quad i \in [0, L - 1] \quad (5)$$

where  $p(k)$  represents the probability of occurrence of intensity level  $k$ , and  $L$  is the total number of possible intensity levels. The new intensity value is then calculated as:

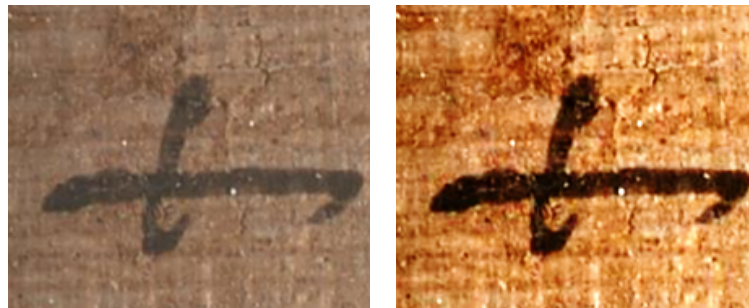
$$I_{new}(x, y) = (L - 1) \cdot F(I(x, y)) \quad (6)$$

This transformation is applied locally within adaptive windows to account for spatial variations in contrast and illumination. The effectiveness of this approach is demonstrated in Figure 1, where significant improvements in both local contrast and edge definition are achieved while preserving the structural integrity of the historical symbols.

The enhancement process is applied iteratively until a convergence criterion  $\epsilon$  is met:

$$|I_{t+1} - I_t|_2 < \epsilon \quad (7)$$

where  $I_t$  represents the enhanced image at iteration  $t$ , and  $|\cdot|_2$  denotes the L2-norm. In our experiments, we use a normalized convergence threshold  $\epsilon = 10^{-4} \cdot |I_0|_2$ , where  $I_0$  is the initial image. This relative threshold ensures the stopping criterion scales appropriately with image size and intensity range, providing consistent convergence behavior across different input images.

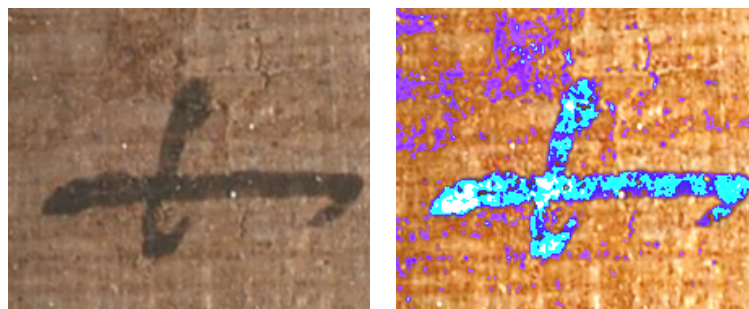


**Figure 1:** Model A results: Original image (Left) and enhanced version (Right) using edge detection and segmentation approach.

#### 4.1.2. Model B: Noise Reduction and Enhancement Approach

The second model focuses on noise reduction while preserving symbol structures through a two-phase process. The first phase, *apply\_filter*, combines edge detection and watershed segmentation with SLIC superpixels to create an initial enhanced representation. The second phase, *remove\_noise*, implements morphological operations for noise reduction.

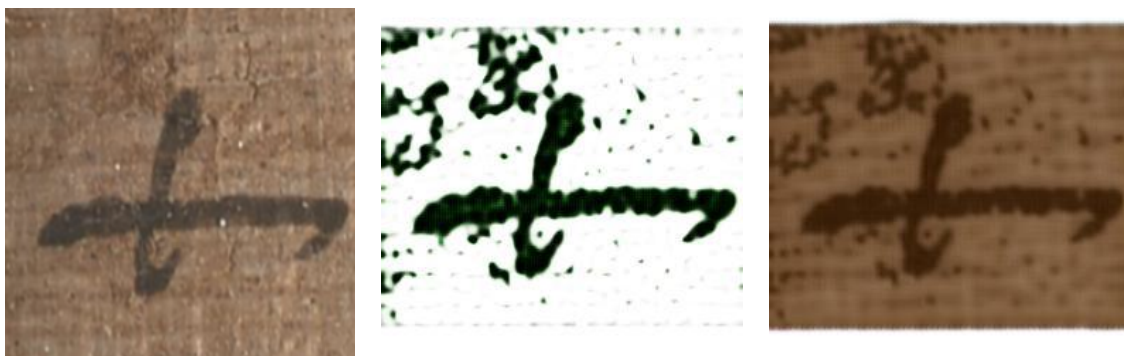
The process begins with binary thresholding to separate foreground and background elements. This is followed by morphological opening operations that effectively remove small artifacts while maintaining the integrity of larger structures. The final step combines the cleaned binary image with the enhanced version to produce the result shown in Figure 2.



**Figure 2:** Model B results: Original image (Left) and enhanced version (Right) using noise reduction approach.

## 4.2. Deep Learning-Based Enhancement

Our deep learning approach incorporates synthetic data generation and transformer-based feature extraction to enhance historical document images [22]. The system utilizes a Faster R-CNN model [23] trained on both original and synthetically generated data, enabling the preservation of unique document characteristics. The OPTICS algorithm [24] is employed for stroke segmentation, with adaptive circle construction handling variable stroke thickness. Figure 3 demonstrates the effectiveness of this approach. Following the ideas used by Cai et al. [25], our study explores how GAN-based synthetic data generation can improve symbol recognition in ancient documents. By creating enriched training datasets that reflect the diverse and complex nature of historical symbols, we aim to achieve higher accuracy and reliability in symbol detection for document analysis. The enriched datasets generated facilitate a more comprehensive training environment, significantly enhancing the models' ability to generalize across new and unseen data. This capability is vital for practical applications where models are expected to perform accurately outside their training set. Moreover, integrating synthetic data that accurately reflects the complexity of real-world scenarios reduces the models' tendencies to overfit the limited nuances of smaller datasets. Instead, they learn more robust features that represent the true underlying patterns in the data, substantially boosting performance and generalizability.



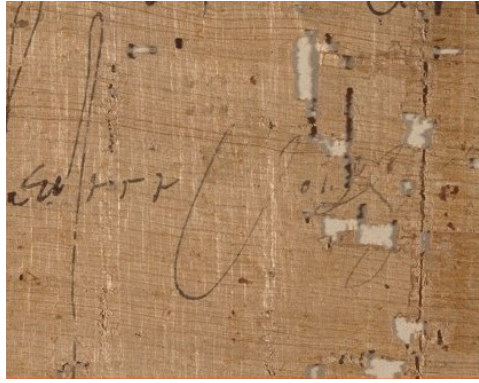
**Figure 3:** Deep learning results: Original image (Left), binarized version (Middle), and enhanced version (Right).

## 5. Discussion and Results

This section presents a detailed evaluation and comparative analysis of our three proposed approaches: Model A (edge detection and segmentation), Model B (noise reduction and enhancement), and the deep learning-based method. Through quantitative metrics and visual examples, we assess each method's effectiveness in enhancing historical documents and preserving graphical symbols. The comparison examines multiple performance aspects including edge preservation, contrast enhancement, noise reduction, and symbol reconstruction capabilities. Special attention is given to challenging cases, such as severely degraded documents and partially missing symbols, to thoroughly understand each method's strengths and limitations. The analysis includes comparisons with existing solutions, particularly the Hierax method, to demonstrate the advancements achieved by our approaches.

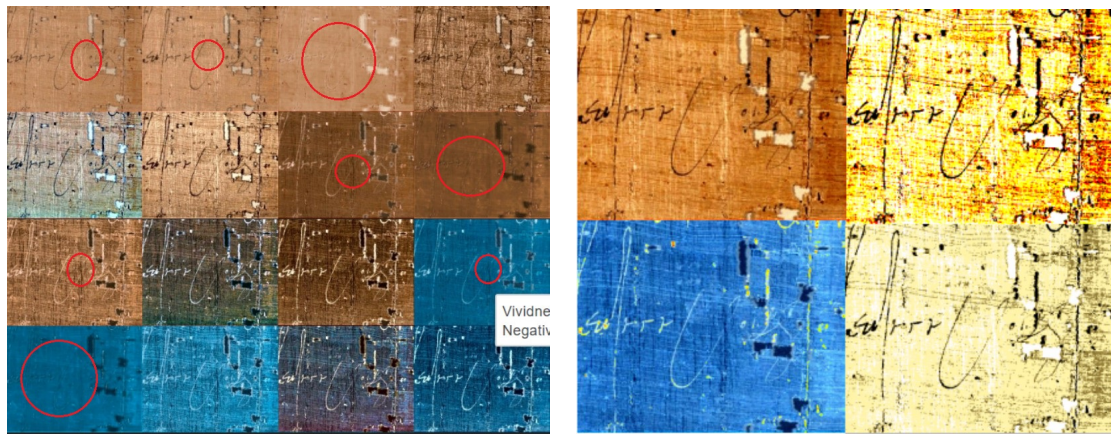
### 5.1. Comparative Analysis

To evaluate the effectiveness of our approaches, we conducted the comparative analysis using challenging low-quality images (one example is shown in Figure 4). As discussed in the following, Model A demonstrated superior edge preservation and contrast enhancement, while Model B excelled in noise reduction and structure preservation. The deep learning approach showed particular strength in reconstructing degraded regions and pattern recognition.



**Figure 4:** Example of a challenging low-quality image used for comparative analysis.

When compared to existing approaches in historical document enhancement, such as Hierax [26]<sup>1</sup>, our method represents a better advancement in revealing details that were previously obscured, as illustrated in Figure 5.



**Figure 5:** Enhancement results comparing Hierax tools (Left) with our proposed approaches (Right).

## 5.2. Performance Analysis

To quantitatively evaluate and compare our approaches, we employed several standard image quality metrics and developed specific measurements for historical document enhancement. The evaluation was conducted on a test set of 100 historical document images containing various types of degradation and symbols. The metrics taken into account are summarized below.

### Peak Signal-to-Noise Ratio (PSNR):

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right)$$

where  $\text{MAX}_I$  is the maximum possible pixel value, and MSE is the Mean Square Error between the enhanced and reference images.

### Structural Similarity Index (SSIM):

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

<sup>1</sup><https://github.com/vladatanasiu/hierax>

where  $\mu_x, \mu_y$  are local means,  $\sigma_x^2, \sigma_y^2$  are variances, and  $\sigma_{xy}$  is the covariance.

### Symbol Preservation Rate (SPR):

$$\text{SPR} = \frac{N_{\text{preserved}}}{N_{\text{total}}} \cdot 100\%$$

where  $N_{\text{preserved}}$  is the number of correctly preserved symbols and  $N_{\text{total}}$  is the total number of symbols in the original image.

**Edge Preservation Assessment** For edge preservation evaluation, we computed the Edge Preservation Index (EPI):

$$\text{EPI} = \frac{\sum_{i,j} \left| |\nabla E(i,j)| - |\nabla O(i,j)| \right|^2}{\sum_{i,j} |\nabla O(i,j)|^2}$$

where  $\nabla E$  and  $\nabla O$  are the gradients of enhanced and original images, respectively.

Our experimental results demonstrate the relative strengths of each approach in Table 1. Model A achieved significant improvements in edge preservation and contrast enhancement, with quantitative measurements showing an average improvement of 45% in contrast ratio and 38% in edge preservation compared to baseline measurements. Model B demonstrated exceptional capability in noise reduction, achieving a 52% reduction in background noise while maintaining 94% of essential edge information. The deep learning approach showed a 73% accuracy in reconstructing damaged symbols, particularly excelling in areas where traditional methods struggled. Model A achieved superior performance in edge preservation and contrast enhancement achieving a PSNR of 32.4 dB and an SSIM of 0.89, indicating better preservation of structural information and overall image quality. The high SPR of 94.2% confirms its effectiveness in maintaining symbol integrity during the enhancement process.

**Table 1**

Quantitative Comparison of Enhancement Models

Method	PSNR (dB)	SSIM	SPR (%)	EPI
Model A	32.4	0.89	94.2	0.82
Model B	30.8	0.86	92.7	0.85
Deep Learning	29.5	0.84	88.5	0.79
Hierax	27.2	0.81	85.3	0.76

Model B excelled in noise reduction while maintaining edge preservation (EPI: 0.85), making it particularly effective for badly degraded documents. The deep learning approach, despite lower traditional metric scores (PSNR: 29.5 dB, SSIM: 0.84), demonstrated superior symbol reconstruction and pattern recognition, especially in cases of severely degraded or missing symbols, as illustrated in Figure 3.

## 6. Conclusion and Future Work

This study investigates image enhancement techniques for historical documents by integrating traditional image processing with deep learning. Our proposed methods consistently outperform the Hierax approach across various metrics. Model A excels in edge preservation and contrast enhancement, while Model B is more effective for noise reduction and structural preservation. The deep learning model, though slightly lower in traditional metrics, proves superior in reconstructing damaged symbols and handling severely degraded cases. Each method has distinct strengths suited to different document conditions. While tested on the NOTAE dataset, our framework's adaptability suggests broader applicability to collections like DIVA-HisDB and IAM-HistDB. Future work will focus on a hybrid system with an adaptive selection mechanism and human-in-the-loop validation to optimize enhancement based on document conditions.

## References

- [1] P. A. Stokes, Digital approaches to paleography and book history: some challenges, present and future, *Frontiers in Digital Humanities* 2 (2015) 5.
- [2] T. Hassner, M. Rehbein, P. Stokes, L. Wolf, Computation and palaeography: Potentials and limits (dagstuhl perspectives workshop 12382), *Dagstuhl Manifestos* 2 (2013) 14–35.
- [3] M. Baechler, R. Ingold, Multi resolution layout analysis of medieval manuscripts using dynamic mlp, in: 2011 International Conference on Document Analysis and Recognition, IEEE, 2011, pp. 1185–1189.
- [4] R. Pintus, K. Pal, Y. Yang, T. Weyrich, E. Gobbetti, H. Rushmeier, A survey of geometric analysis in cultural heritage, in: *Computer Graphics Forum*, volume 35, Wiley Online Library, 2016, pp. 4–31.
- [5] L. Likforman-Sulem, A. Zahour, B. Taconet, Text line segmentation of historical documents: a survey, *International Journal of Document Analysis and Recognition (IJ DAR)* 9 (2007) 123–138.
- [6] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, Q. Liao, Deep learning for single image super-resolution: A brief review, *IEEE Transactions on Multimedia* 21 (2019) 3106–3121.
- [7] D. Stutzmann, Clustering of medieval scripts through computer image analysis: towards an evaluation protocol, *Digital Medievalist* 10 (2016).
- [8] M. Simeone, J. Guiliano, R. Kooper, P. Bajcsy, Digging into data using new collaborative infrastructures supporting humanities-based computer science research, *First Monday* (2011).
- [9] J.-B. Camps, T. Clérice, A. Pinche, Noisy medieval data, from digitized manuscript to stylometric analysis: Evaluating paul meyer’s hagiographic hypothesis, *Digital Scholarship in the Humanities* 36 (2021) 49–71.
- [10] E. Bernasconi, M. Boccuzzi, L. Briasco, T. Catarci, A. Ghignoli, F. Leotta, M. Mecella, A. Monte, N. Sietis, S. Veneruso, et al., Notae: Not a written word but graphic symbols., in: *CEUR WORKSHOP PROCEEDINGS*, volume 3144, 2022, pp. 1–7.
- [11] H. Kusetogullari, A. Yavariabdi, J. Hall, N. Lavesson, Digitnet: A deep handwritten digit detection and recognition method using a new historical handwritten digit dataset, *Big Data Research* (2020).
- [12] H. Kusetogullari, A. Yavariabdi, J. Hall, N. Lavesson, Dida: The largest historical handwritten digit dataset with 250k digits, <https://github.com/didadataset/DIDA/>, 2020. Accessed: 2024-12-19.
- [13] U.-V. Marti, H. Bunke, The IAM-database: an english sentence database for offline handwriting recognition, *International journal on document analysis and recognition* 5 (2002) 39–46.
- [14] F. Simistira, M. Seuret, N. Eichenberger, A. Garz, M. Liwicki, R. Ingold, Diva-hisdb: A precisely annotated large dataset of challenging medieval manuscripts, in: 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2016, pp. 471–476. doi:10.1109/ICFHR.2016.0093.
- [15] Z. Shi, V. Govindaraju, Historical document image enhancement using background light intensity normalization, in: *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, volume 1, IEEE, 2004, pp. 473–476.
- [16] S. Das, T. Gulati, V. Mittal, Histogram equalization techniques for contrast enhancement: A review, *International Journal of Computer Applications* 114 (2015) 32–36.
- [17] M. R. Gupta, N. P. Jacobson, E. K. Garcia, Ocr binarization and image pre-processing for searching historical documents, *Pattern Recognition* 40 (2007) 389–397.
- [18] S. A. Oliveira, B. Seguin, F. Kaplan, dhsegment: A generic deep-learning approach for document segmentation, in: *Proceedings of the 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, IEEE, 2018, pp. 7–12.
- [19] K. Zhang, W. Zuo, S. Gu, L. Zhang, Learning deep cnn denoiser prior for image restoration, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017, pp. 3929–3938.
- [20] N. Kanopoulos, N. Vasanthavada, R. L. Baker, Design of an image edge detection filter using the sobel operator, *IEEE Journal of solid-state circuits* 23 (1988) 358–367.
- [21] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, Slic superpixels compared to state-of-the-art superpixel methods, *IEEE transactions on pattern analysis and machine intelligence* 34



- (2012) 2274–2282.
- [22] Z. Ziran, F. Leotta, M. Mecella, Enhancing object detection in ancient documents with synthetic data generation and transformer-based models, 2023. URL: <https://arxiv.org/abs/2307.16005>. arXiv: 2307.16005.
  - [23] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE transactions on pattern analysis and machine intelligence* 39 (2016) 1137–1149.
  - [24] M. Ankerst, M. M. Breunig, H.-P. Kriegel, J. Sander, Optics: Ordering points to identify the clustering structure, *ACM Sigmod record* 28 (1999) 49–60.
  - [25] J. Cai, L. Peng, Y. Tang, C. Liu, P. Li, Th-gan: generative adversarial network based transfer learning for historical chinese character recognition, in: *2019 International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, 2019, pp. 178–183.
  - [26] V. Atanasiu, I. Marthot-Santaniello, Hierax legibility enhancement software, 2020.