Impact of Image Enhancement Methods on Automatic Transcription Trainings with eScriptorium

Pauline Jacsont and Elina Leblanc

University of Geneva, Switzerland

Corresponding author: Pauline Jacsont , p.jacsont@gmail.com

Abstract

This study stems from the *Desenrollando el cordel* (*Untangling the cordel*) project, which focuses on 19th-century Spanish prints editing. It evaluates the impact of image enhancement methods on the automatic transcription of low-quality documents, both in terms of printing and digitisation. We compare different methods (binarisation, deblur) and present the results obtained during the training of models with the *Kraken* tool. We demonstrate that binarisation methods give better results than the other, and that the combination of several techniques did not significantly improve the transcription prediction. This study shows the significance of using image enhancement methods with *Kraken*. It paves the way for further experiments with larger and more varied corpora to help future projects design their automatic transcription workflow.

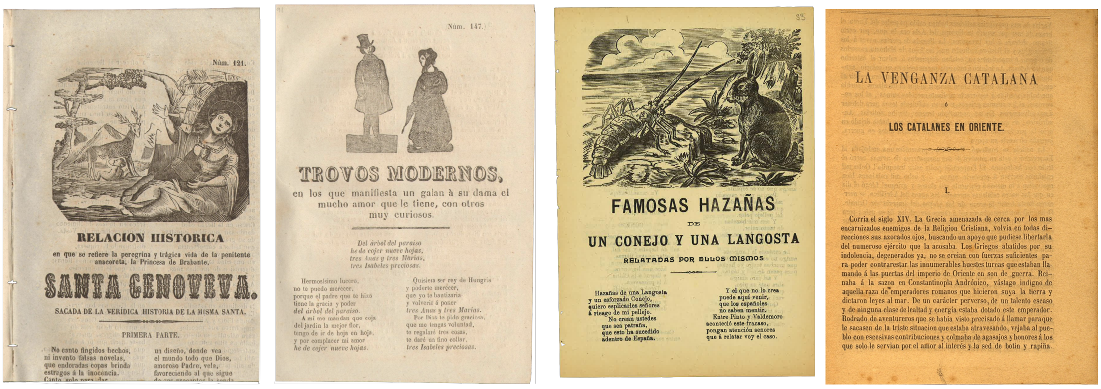
Keywords

image enhancement methods; binarisation; deblur; printed documents; Spanish literature

I INTRODUCTION

The library of the University of Geneva holds a collection of almost 1’000 Spanish chapbooks, printed during the 19th century by several printers across Spain. Chapbooks, also known as *pliegos de cordel*, consist of a few pages (4 to 8) and are in quarto format. They recount real or fictitious events, share songs and poems, or prayers and other religious writings (Figure 1). The Geneva collection is the object of the *Untangling the cordel* project [Leblanc and Carta, 2021], which aims at studying and promoting these documents through a digital library. As one of the main objectives of the project is to analyse the chapbooks’ content, automatic transcription represents a significant step in our editorial workflow to publish diplomatic digital editions using the XML-TEI standard. After testing several tools, including *ABBYY FineReader* and *Transkribus* [Kahle et al., 2017], we chose *Kraken* and its graphic interface *eScriptorium*[1] [Kiessling et al., 2019].

[1] At the beginning of the project, we carried out our automatic transcription experiments with *ABBYY FineReader* and *Transkribus*. These tools allowed us to create our Ground Truth (GT) and transcribe a part of our collection. However, after one year, the University of Geneva developed an automatic transcription platform called *FoNDUE* (cf. https://www.canal-u.tv/chaines/enc/21-fondue-a-lightweight-htr-infrastructure-for-geneva), based on the *Kraken/eScriptorium* tools. We became beta-testers of this new platform, which explains the use of different tools during our project. All the experiments we describe in this paper were performed exclusively with *Kraken*.



RELACION HISTORICA

SANTA GENOVEVA.

TROVOS MODERNOS,

FAMOSAS HAZAÑAS DE UN CONEJO Y UNA LANGOSTA

LA VENGANZA CATALANA

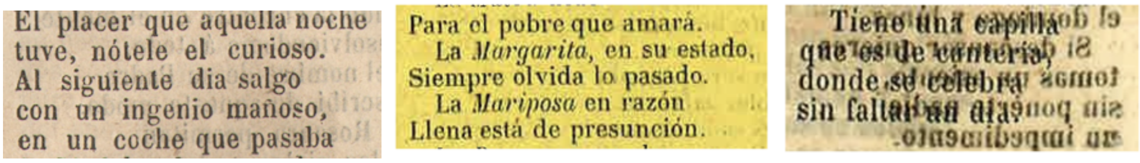
LOS CATALANES EN ORIENTE.

Figure 1: Examples of chapbooks (From left to right: José María Moreno, Carmona, 1859; José María Moreno, Carmona, [s.d.]; Imp. El Abanico, Barcelone, [s.d.]; J. Jepús, Barcelone, 1884)

Several challenges arose during our initial experiments with this tool. First, the page segmentation phase is faced with the complex layout of the document. Chapbooks' pages can display up to three columns. There can be variation in the layout of the pages even within the same chapbook.

Then, during transcription, the variety of employed fonts used poses a challenge for *Kraken*, especially in the title sections where most of the errors, produced by our model, cluster. Regarding the core text, the quality of the print media (paper and ink) complicates character recognition. Indeed, Spanish chapbooks – also known as cheap prints – were printed in mass on poor-quality paper. This often results in bleed-through, which blends the text on the recto with that of the verso (Figure 2).

In our case, the quality of the digital facsimiles adds another layer of difficulty. Indeed, before the project began, several digitisation campaigns were carried out with various scanners, and by different staff members of the library and the Spanish Unit of the university. Therefore, the resulting digitised corpus is heterogeneous. While some chapbooks (33% of the corpus) are in TIFF with a resolution of 300 dpi, most of the corpus (67%) is in PDF format with a relatively low resolution (72 dpi). For dissemination, these PDF documents were converted into JPG, which resulted in a further deterioration in quality. The resulting images are indeed heavily pixelated, which adds noise to the recognition of the characters by our model (Figure 2).



El placer que aquella noche
tuve, nótele el curioso.
Al siguiente dia salgo
con un ingenio mañoso,
en un coche que pasaba

Para el pobre que amará.
La *Margarita*, en su estado,
Siempre olvida lo pasado.
La *Mariposa* en razón
Llena está de presunción.

Tiene una capilla
que es de cantería
donde se celebra
sin faltar un día

Figure 2: (From left to right) A TIFF image, a pixelated JPG image and a JPG image with bleed-through

In this paper, we focus on one of the problems mentioned above: the low quality of our data, both in terms of printing and digitisation. We propose a comparison of different image enhancement methods to address it, namely binarization and deblurring, and describe their impact on the accuracy of our transcription model.

After presenting related works, the third part explains our workflow and the image enhancement methods we included in our study. The fourth section presents the effectiveness of each method tested to improve our model, and the method we chose for our data. Finally, we discuss the replicability of this experiment with other types of documents. We also address the possibilities offered by image enhancement methods to improve the results of models trained with *Kraken*.

II RELATED WORKS

With the advent of deep learning technologies, image enhancement is widely used to improve the effectiveness of many image processing tasks, such as segmentation, face detection or optical character recognition. In a recent survey about image enhancement methods [Anvari and Athitsos, 2022], the authors distinguish six different tasks to improve the quality of an image, depending on the type of damage: binarization, deblurring, denoise, defading, watermark removal and shadow removal. For instance, binarization and denoising are preferred for document damage (wrinkles, stains, bleed-through); defading and deblurring are mostly used to improve the quality of a digitisation and its exposure.

From our analysis of the scientific literature, binarization appears as the main approach projects adopt to improve their damaged images of printed or handwritten historical texts prior to the automatic transcription. Early binarization approaches can be categorised into global and local thresholding methods. Global methods, amongst which the Otsu method [Otsu, 1979] is the most well-known, apply a single threshold to the entire image. These methods are effective for documents with a high contrast between the foreground and the background. However, they perform poorly with complex and uneven backgrounds [Boiangiu et al., 2011, Gupta et al., 2007]. Local methods, such as Niblack [Niblack, 1986] or Sauvola [Sauvola and Pietikäinen, 2000], determine a threshold at a pixel level, by taking into consideration the color of the surrounding pixels within a specific window, so that local variations in an image can be better taken into account [Boiangiu et al., 2011]. A comparison of these different solutions for the improvement of ATR predictions with *ABBYY FineReader* can be found in Gupta, Jacobson, and Garcia [Gupta et al., 2007]. They found that the Otsu algorithm gave the best results for automatic transcription of English newspapers, digitised at low-resolution and pixelated.

Recently, several studies have proposed to improve these methods, especially for handwritten documents. Boiangiu et al. propose to modify the Niblack method by dynamically defining the window used to determine the threshold, instead of using a predefined value [Boiangiu et al., 2011]. They conclude that this method better harmonises the binarization of images with irregular brightness repartition. Ntirogiannis, Gatos and Pratikakis propose a combination of global and local binarization. This method proves to be efficient in detecting faint characters and removing bleed-through from highly damaged handwritten documents [Ntirogiannis et al., 2014]. Adaptive binarization was also chosen by the developers of *Kraken*. Indeed, binarization is one of the available pre-processing options when a training is launched. This method dynamically calculates the difference between the highest and lowest thresholds for different regions of an image[2]. As for deblurring – which in our case seems to be a relevant approach to improve our pixelated digitisations –, most studies are related to object or face detection. Usually, for this type of task, methods rely on blind deconvolution methods. However, several works showed that convolutional neural networks (CNN) are better at deblurring text images [Hradiš et al., 2015, Gangeh et al., 2019]. Other approaches suggest Generative Antagonistic Network (GAN) methods to deal with a heterogeneous corpus composed of faces pictures and text images to reconstruct high-resolution images from low-resolution ones [Xu et al., 2017].

[2] To our knowledge, no documentation or evaluation has been published about the method chosen by the developers of *Kraken*. These assumptions are based on the code available on the project's GitHub: https://github.com/mittagessen/kraken/blob/master/kraken/binarization.py

However, as pointed out by Anvari and Athitsos [Anvari and Athitsos, 2022], while image enhancement methods prove their effectiveness in improving text images, few papers give details about their impact on the accuracy of automatic transcription models. We can mention some experiments with *ABBYY FineReader* and *Tesseract* in [Gupta et al., 2007, Hradiš et al., 2015, Gangeh et al., 2019, Souibgui et al., 2021]. In [Boiangiu et al., 2011] and [Ntirogiannis et al., 2014], the authors propose an evaluation with their own ATR systems. Furthermore, these works focus their attention on only one method to improve the images of their historical documents.

Therefore, in this paper, we study the benefit of different image enhancement techniques on the predictions of our model for Spanish printed documents. To achieve this, we use *Kraken*, an experiment that, to our knowledge, has never been carried out with this specific tool. It leads us to compare the native binarization method of *Kraken*, primarily thought for handwritten documents, with other solutions.

III METHOD

To evaluate the impact of image enhancement on the effectiveness of our automatic transcription models, we conducted a series of tests comparing different binarization approaches, listed below. The Ground Truth (GT) consists of 198 pages transcribed with *ABBYY FineReader*, and then manually corrected. This corpus was divided into three subsets[3]. Each set is used in a different phase:

- The first set, 80% of GT, is used by *Kraken* to train the model;
- The second set, 10% of the corpus, is used by the tool to evaluate each iteration during training (validation set);
- The last set, 10% of the corpus, is used to evaluate the results of the model on documents it has never encountered (test set).

Each time we pre-processed the data using a notebook[4] for the entire corpus: we chose this method because it was easy to implement and allowed the use of open-source Python libraries – Open CV and Scikit-image – where a pre-processing collection is available. Our experiment is based on the work of Gupta, Jacobson and Garcia on automatic transcription of English newspapers, which resemble our corpus in terms of layout and image quality [Gupta et al., 2007].

To choose the best binarization method for our Spanish chapbooks, we reproduce a part of their work by comparing different solutions. We did not use exactly the same methods but adapted their experiment to our skills and time constraints. Thus, we chose methods that were easy and fast to implement, namely the thresholding binarization of OpenCV, Otsu, Niblack, Sauvola and the native binarization of *Kraken*. We also tested two types of Gamma correction: one to lighten the image (Gamma 1) and the other to darken it (Gamma 2) to experiment with contrasts.

We also tried a deblur technique, which follows the state of the art of Zahra Anvari and Vassilis Athitsos, that focuses, in part, on historical documents that, like those in our corpus, are degraded and damaged [Anvari and Athitsos, 2022].

[3] The details of the three sets are available on our GitHub repository : https://github.com/DesenrollandoElCordel/FoNDUE-Spanish-chapbooks-Dataset/tree/main/Grountruth/Split.

[4] The OpenCV, NumPy, SciPy and Scikit-image Python libraries are used in the notebook available here: https://github.com/DesenrollandoElCordel/FoNDUE-Spanish-chapbooks-Dataset/blob/main/Varios-GroundTruth-TEST/ImagesTreatments.ipynb.

Each automatic transcription model[5] trained in this test phase was run according to basic training[6]; the command used is : `Ketos train -f alto -t train.txt -e val.txt -d cuda data/*.xml`. For each performed pre-processing, a training was followed by an evaluation test with the set prepared for this purpose. The same XML-ALTO files were used for each test; only the images changed depending on the type of pre-processing.

In a second phase, we conducted a series of tests to evaluate the effectiveness of different image enhancement methods when combined, using the pre-processes that had produced the best results in the first phase.

IV RESULTS

The results of the models trained in the first test phase (i.e. the models created with the images having undergone only one pre-processing) are shown in figure 3. The results are given for the accuracy per character on the training tests (validation test) and the evaluation tests.

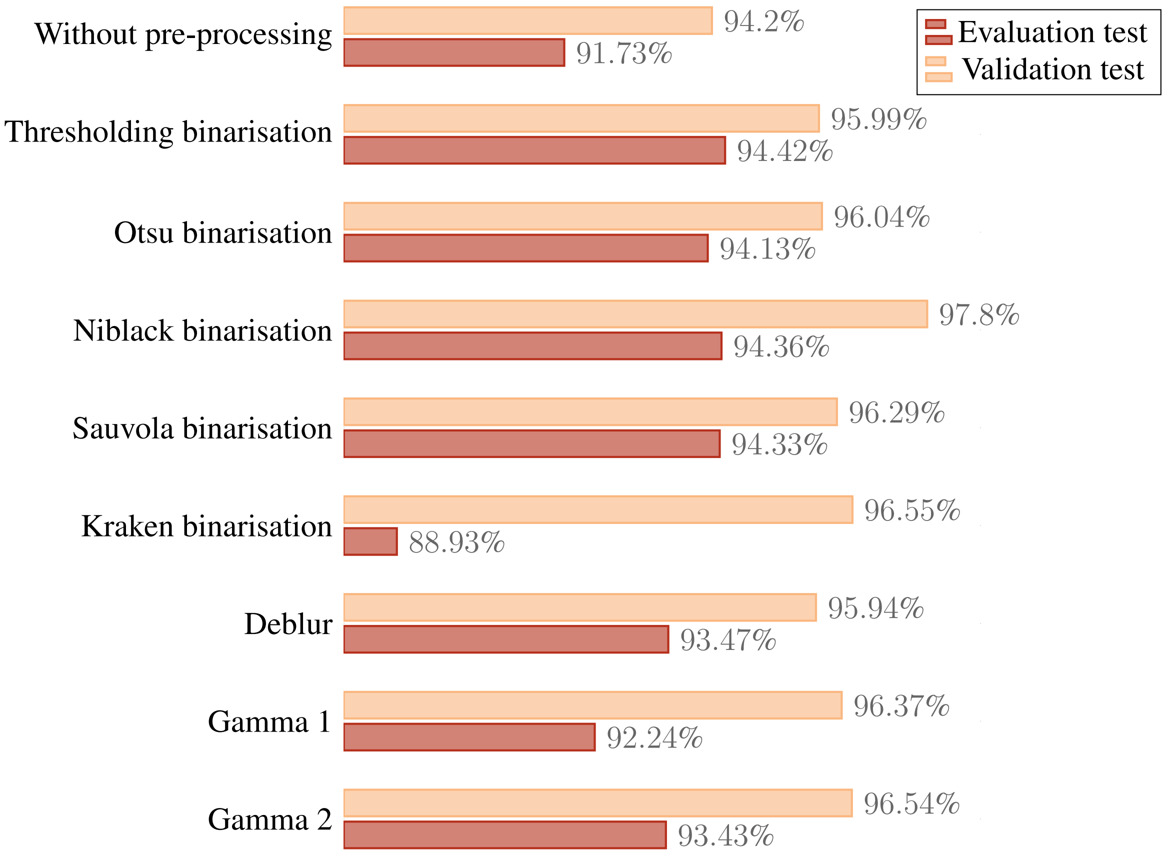
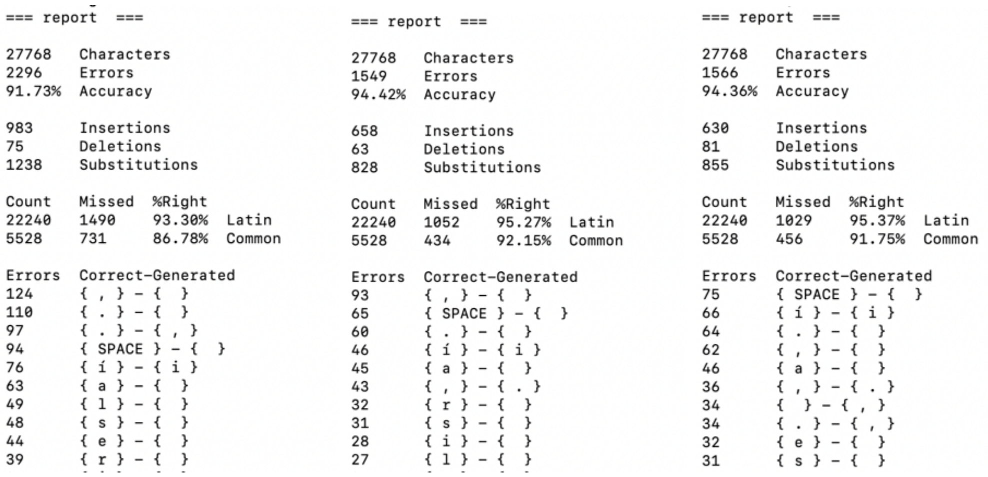


Figure 3: Models with single pre-processed images

The detailed results (Figure 4) show that the models trained with the binarized images successfully recognise commas and dots. However, some errors remain, such as confusion between *i* and *í* and problems in recognising certain characters such as *a*. To measure the efficiency of

[5] The same tests were also carried out on the segmentation models, but this did not lead to improved results.

[6] The submission script is available at the following address: https://github.com/DesenrollandoElCordel/FoNDUE-Spanish-chapbooks-Dataset/blob/main/Grountruth/submission-script.sh.



```
=== report  ===

27768   Characters
2296    Errors
91.73%  Accuracy

983     Insertions
75      Deletions
1238    Substitutions

Count   Missed  %Right
22240   1490    93.30%  Latin
5528    731     86.78%  Common

Errors  Correct-Generated
124     { , } - {  }
110     { . } - {  }
97      { . } - { , }
94      { SPACE } - {  }
76      { í } - { i }
63      { a } - {  }
49      { l } - {  }
48      { s } - {  }
44      { e } - {  }
39      { r } - {  }
```

```
=== report  ===

27768   Characters
1549    Errors
94.42%  Accuracy

658     Insertions
63      Deletions
828     Substitutions

Count   Missed  %Right
22240   1052    95.27%  Latin
5528    434     92.15%  Common

Errors  Correct-Generated
93      { , } - {  }
65      { SPACE } - {  }
60      { . } - {  }
46      { í } - { i }
45      { a } - {  }
43      { , } - { . }
32      { r } - {  }
31      { s } - {  }
28      { i } - {  }
27      { l } - {  }
```

```
=== report  ===

27768   Characters
1566    Errors
94.36%  Accuracy

630     Insertions
81      Deletions
855     Substitutions

Count   Missed  %Right
22240   1029    95.37%  Latin
5528    456     91.75%  Common

Errors  Correct-Generated
75      { SPACE } - {  }
66      { í } - { i }
64      { . } - {  }
62      { , } - {  }
46      { a } - {  }
36      { , } - { . }
34      {  } - { , }
34      { . } - { , }
32      { e } - {  }
31      { s } - {  }
```

Figure 4: (From left to right) Test reports of the model without pre-processing; with the Thresholding binarization; with the Niblack model

the models, we compared their ability to transcribe a few lines of text on the most complex documents (in this case, a chapbook with poor scan quality, block artefacts specific to the JPG format and smudges during printing).

The results shown in Table 1 confirm those obtained previously. The thresholding and Niblack models perform better but still have difficulty transcribing the letter *a*. Both make the mistake with *una* which is transcribed as *uns*. The model without pre-processing and the one with the Otsu method make more errors, especially on punctuation.

It is worth noting that none of the models manages to transcribe the tilde *n* (*muñeco*) correctly. Therefore, despite the similarity of our documents to those of Gupta, Jacobson and Garcia,

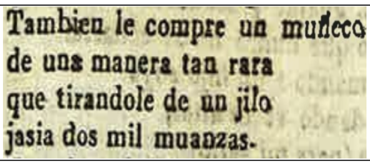
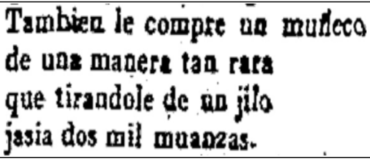
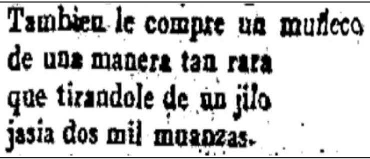
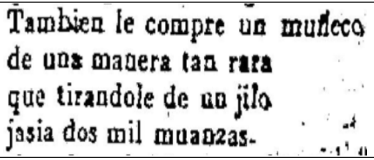
Without enhancement		Tamfién le compre ua mu∅éro, de uns manera tan rará que tirandole de un ∅ilo jasia dos mil muanzas.
Otsu		Tamfiea le compre nu munéso de una manera tan rará que tirandole de un jilo l∅sia dos mil muanzas∅
Thresholding		Tambeien le compre ua mulezo de uns manera tan rara que tirandole de un ∅ilo jasia dos mil muanzas.
Niblack		Tambien le compre un muleco de uns manera tan rara que tirandole de un jilo jasia dos mil muanzas.

Table 1: Transcription of complex lines. Ground truth: Tambien le compre un muñeco / de una manera tan rara / que tirandole de un jilo / jasia dos mil muanzas. Insertions are coloured in blue, substitutions in green and deletions in red.

which suggests that the Otsu method would also be most efficient for us, we do not get the same results. In our case the Niblack method seems to be the most appropriate.

A second series of tests was performed by combining binarization with another image enhancement method. We did not launch any automatic transcription model training if the image obtained was clearly unusable for the tool; this is the case, for example, when combining a deblurring method (sharpening kernel) with the Thresholding binarization (Figure 5).

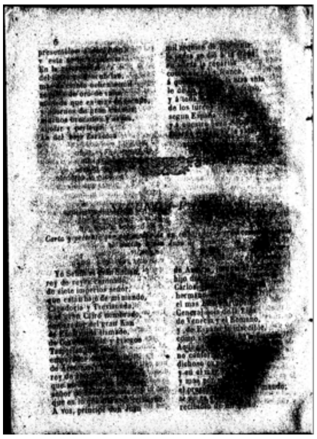


Figure 5: Example of unusable pre-processing (combination of techniques: Sharpening and Thresholding binarization)

Four other models were trained with double pre-processed images. The results obtained are shown in 6, with the accuracy per character in percent.

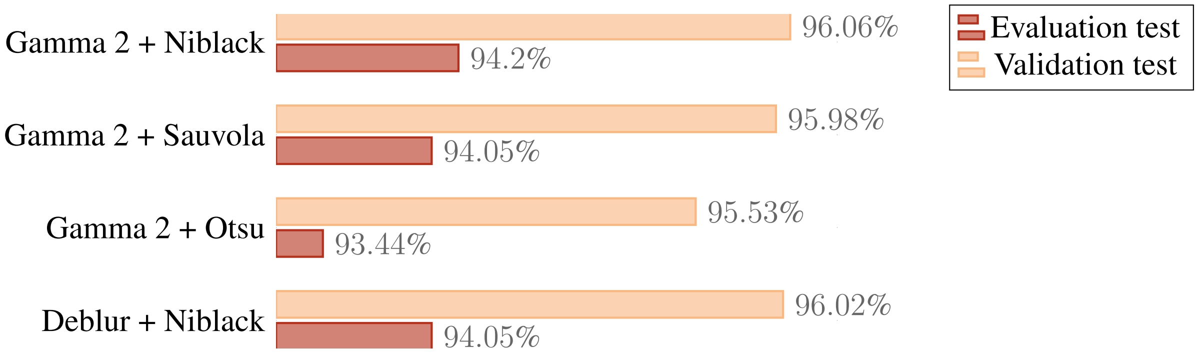


Figure 6: Models with single pre-processed images

Contrary to our initial hypothesis, the multiple-image pre-processing method did not improve the results obtained in the first phase. However, the accuracy of the trained models remains higher than that of the model trained on images without pre-processing.

These different tests show the importance and influence of image pre-processing methods on automatic transcription predictions. Ultimately, the best method for improving our results, and the method applied to the rest of the images in our corpus, is the Niblack binarization. This model is interesting in that it performs better in the recognition of Latin characters than the thresholding binarization.

V CONCLUSIONS AND FUTURE WORKS

In this paper, we have focused on the impact of image enhancement to improve the performance of ATR models. However, other methods can also improve the models, such as increasing the training data or changing the ATR architecture [Chagué, 2021]. Thanks to the various experiments carried out, we have shown that the use of image enhancement methods can significantly improve the predictions of a model trained with *Kraken*. This empirical approach was carried out on a corpus of printed documents, with low resolution, bleed-through and JPG block artefacts.

By reproducing some of the experiments of Gupta, Jacobson, and Garcia, we found that the best pre-processing for our corpus was not Otsu binarization, but Niblack binarization. The choice and efficiency of image pre-processing in transcription depend largely on the specifics of the corpus and the quality of its digitisation. A similar experiment was carried out in another project, and the results obtained were very different: this experience made with a manuscript[7], that was digitised at high resolution. The results showed that image enhancement methods did not improve the prediction of the models[8].

These results pave the way for further experiments and research. Indeed, it would be useful to reproduce these experiments on a large scale to define recommendations on the best image enhancement methods to use with *Kraken* and other user-friendly automatic transcription tools, depending on the nature of texts and the quality of the images.

Finally, we used traditional and well-known methods for our binarization experiments. However, new approaches using Deep Learning architectures (CNN [Akbari et al., 2020] or GAN [Khamekhem Jemni et al., 2022]) are currently being tested. They are still in their infancy, but it would be interesting to include them in future work to analyse their impact on the accuracy of a model trained with Kraken compared to threshold and adaptive methods.

VI AUTHORS' CONTRIBUTIONS

Pauline Jacsont has conceived and performed all the experiments with the different image enhancement methods. Élina Leblanc has elaborated the general workflow of the project and helped design this study.

VII ACKNOWLEDGEMENTS

We sincerely thank the Fondue project team, Simon Gabay, Pierre Künzli, Christophe Charpilloz and Jean-Luc Falcone, for their constant support and invaluable responsiveness, which made all these experiments possible.

VIII DATASETS AND MODELS

Constance Carta, Pauline Jacsont, Élina Leblanc, Belinda Palacios, and Luana Bermúdez. FoNDUE Spanish chapbooks 19th c. Dataset. In Alix Chagué and Thibault Clérice, editors, *HTR-United*. URL https://github.com/DesenrollandoElCordel/FoNDUE-Spanish-chapbooks-Dataset

[7] This experiment has been done in the context of the digital edition project of a Latin manuscript of the Bern Burgerbibliothek [Sardi, 1561]. This document from the 16th century was digitised in TIFF format (400 dpi).

[8] The results of these tests are available in detail in [Jacsont, 2022]

Pauline Jacsont, and Florian Mittenhuber. FoNDUE-GasparoSardiToponomasia-Dataset. In Alix Chagué and Thibault Clérice, editors, *HTR-United*. URL https://github.com/PaulineJac/GasparoSardiToponomasia/tree/main/HTR

References

Younes Akbari, Somaya Al-Maadeed, and Kalthoum Adam. Binarization of degraded document images using convolutional neural networks and wavelet-based multichannel images. *IEEE Access*, 8:153517–153534, 2020. ISSN 2169-3536. doi: 10.1109/ACCESS.2020.3017783. URL https://ieeexplore.ieee.org/document/9171243/.

Zahra Anvari and Vassilis Athitsos. A survey on deep learning based document image enhancement, 2022. URL http://arxiv.org/abs/2112.02719. arXiv:2112.02719 [cs].

Costin-Anton Boiangiu, Alexandra Olteanu, Alexandru Stefanescu, Daniel Rosner, Nicolae Tapus, and Mugurel Andreica. Local thresholding algorithm based on variable window size statistics. In *Proceedings of the 18th International Conference on Control Systems and Computer Science (CSCS)*, volume 2, pages 647–652, Bucharest, Romania, 2011.

Alix Chagué. Création de modèles de transcription pour le projet lectaurep #2. In *LECTAUREP : L'intelligence artificielle appliquée aux archives notariales (blog)*, Paris, France, 2021. URL https://lectaurep.hypotheses.org/category/experimentations.

Mehrdad J. Gangeh, Sunil R. Tiyyagura, Sridhar Dasaratha, Hamid Motahari, and Nigel P. Duffy. Document enhancement system using auto-encoders. In *Proceedings of the 33rd Conference on Neural Information Processing Systems*, Vancouver, Canada, 2019. URL https://www.academia.edu/66039620/Document_Enhancement_System_Using_Auto_encoders.

Maya R. Gupta, Nathaniel P. Jacobson, and Eric K. Garcia. Ocr binarization and image pre-processing for searching historical documents. *Pattern Recognition*, 40(2):389–397, 2007. ISSN 0031-3203. doi: 10.1016/j.patcog.2006.04.043. URL https://www.sciencedirect.com/science/article/pii/S0031320306002202.

Michal Hradiš, Jan Kotera, Pavel Zemcík, and Filip Sroubek. Convolutional neural networks for direct text deblurring. In *Proceedings of BMVC*, volume 10, 2015. doi: 10.5244/C.29.6.

Pauline Jacsont. Mise en valeur du patrimoine textuel grâce aux éditions numériques "le cas du codex 174 de la bibliothèque de la bourgeoisie de berne". Master's thesis, Université de Genève, 2022.

Philip Kahle, Sebastian Colutto, Günter Hackl, and Günter Mühlberger. Transkribus - a service platform for transcription, recognition and retrieval of historical documents. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 04, pages 19–24, 2017. doi: 10.1109/ICDAR.2017.307. ISSN: 2379-2140.

Sana Khamekhem Jemni, Mohamed Ali Souibgui, Yousri Kessentini, and Alicia Fornés. Enhance to read better: A multi-task adversarial network for handwritten document image enhancement. *Pattern Recognition*, 123:108–141, 2022. ISSN 0031-3203. doi: 10.1016/j.patcog.2021.108370. URL https://www.sciencedirect.com/science/article/pii/S0031320321005501.

Benjamin Kiessling, Robin Tissot, Peter Stokes, and Daniel Stökl Ben Ezra. escriptorium: An open source platform for historical document analysis. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, volume 2, pages 19–19, 2019. doi: 10.1109/ICDARW.2019.10032.

Elina Leblanc and Constance Carta. Le projet "démêler le cordel" : une bibliothèque numérique pour l'étude de la littérature éphémère espagnole du xix^{e} siècle. In *Humanistica 2021, Rennes, 10-12 mai 2021*, pages 100–101, 2021. URL https://hal.archives-ouvertes.fr/hal-03526522.

Wayne Niblack. *An Introduction to Digital Image Processing*. Prentice-Hall International, 1986. ISBN 978-0-13-480674-7. Google-Books-ID: XOxRAAAAMAAJ.

K. Ntirogiannis, B. Gatos, and I. Pratikakis. A combined approach for the binarization of handwritten document images. *Pattern Recognition Letters*, 35:3–15, 2014. ISSN 0167-8655. doi: 10.1016/j.patrec.2012.09.026. URL https://www.sciencedirect.com/science/article/pii/S016786551200311X.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979. ISSN 0018-9472, 2168-2909. doi: 10.1109/TSMC.1979.4310076. URL http://ieeexplore.ieee.org/document/4310076/.

Gasparo Sardi. Toponomasia. 1561. URL https://katalog.burgerbib.ch/detail.aspx?ID=340662.

J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern Recognition*, 33(2):225–236, 2000. ISSN 0031-3203. doi: 10.1016/S0031-3203(99)00055-2. URL https://www.sciencedirect.com/science/article/pii/S0031320399000552.

Mohamed Ali Souibgui, Yousri Kessentini, and Alicia Fornés. A conditional gan based approach for distorted camera captured documents recovery. In Chawki Djeddi, Yousri Kessentini, Imran Siddiqi, and Mohamed Jmaiel, editors, *Pattern Recognition and Artificial Intelligence*, Communications in Computer and Information Science, pages 215–228, Cham, 2021. Springer International Publishing. ISBN 978-3-030-71804-6. doi: 10.1007/978-3-030-71804-6_16.

Xiangyu Xu, Deqing Sun, Jinshan Pan, Yujin Zhang, Hanspeter Pfister, and Ming-Hsuan Yang. Learning to super-resolve blurry face and text images. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 251–260, 2017. doi: 10.1109/ICCV.2017.36. ISSN: 2380-7504.