READ

Recognition and Enrichment of Archival Documents

D6.3 Binarization and Image Enhancement Tools P3

Ioannis Pratikakis, Konstantinos Zagoris DUTH
Basilis Gatos, NCSR

Distribution:

http://read.transkribus.eu/

READ
H2020 Project 674943

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 674943

Project ref no.	H2020 674943
Project acronym	**READ**
Project full title	**Recognition and Enrichment of Archival Documents**
Instrument	H2020-EINFRA-2015-1
Thematic Priority	EINFRA-9-2015 - e-Infrastructures for virtual research environments (VRE)
Start date / duration	01 January 2016 / 42 Months

Distribution	Public
Contractual date of delivery	31.12.2018
Actual date of delivery	04.12.2018
Date of last update	
Deliverable number	D6.3
Deliverable title	Binarization and Image Enhancement Tools P3
Type	Demonstrator
Status & version	Public & Version 1
Contributing WP(s)	WP6
Responsible beneficiary	DUTH
Other contributors	
Internal reviewers	Max Weidemann, Florian Kleber
Author(s)	Ioannis Pratikakis, Konstantinos Zagoris DUTH Basilis Gatos, NCSR
EC project officer	
Keywords	Document Image Binarization, Ground Truth, Document Image Enhancement

Table of Contents

Executive Summary

This deliverable reports on the achievements concerning the tasks of document image binarization at the end of the third year of the READ project. This year DUTH concentrated on increasing the accuracy of previous binarization algorithms (M12 and M24 deliverable).

1. DUTH Binarization

In contrast to the previous year, during the third year, DUTH concentrated on improving M24's Binarization method. Figure 1 shows its architectural structure with red showing the current year changes.

Briefly, these changes are as follows:

- Replacing Otsu algorithm with a better local binarization algorithm (GPP [GAT2006]) for better calculation of the Stroke Width Transform (SWT) [EPS2010]
- Better Support Vector Machines (SVM) [VAP2013] filtering

In our comparative study, we have considered a binarization method which was developed in the frame of the "tranScriptorium" project that is based on the method of Ntirogiannis et al. [NTI2014] adapted to handwritten document images. This method was also used in the evaluation performed during Y1 of the project (see previous year's deliverable D6.1). This algorithm is presented in three variations, denoted as 'NCSR – method (i)-(iii)', respectively.

Table 1 shows the evaluation results using the datasets from the ICFHR2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016) [PRA2016], ICDAR2017 Competition on Document Image Binarization (DIBCO 2017) [PRA2017] and ICFHR2018 Handwritten Document Image Binarization Contest (H-DIBCO 2018) [PRA2018]. Also, the performance of the first ranked method in these competitions is also presented, denoted as '#1 Competition Method'.

The DUTH M12 method although performed very well on H-DIBCO2016 dataset, it did not achieve similar performance on DIBCO2017 dataset and yielding very bad results on H-DIBCO2018. This is because it cannot handle the bleed-through characters and very large borders (H-DIBCO2018) very well. Unfortunately, the distinction between faint characters that should be preserved and bleed-through characters that should be omitted is a very difficult and in some cases near impossible task.

Moreover, because there is not any texture-based filtering on M12, it performs badly on big black borders around the documents.

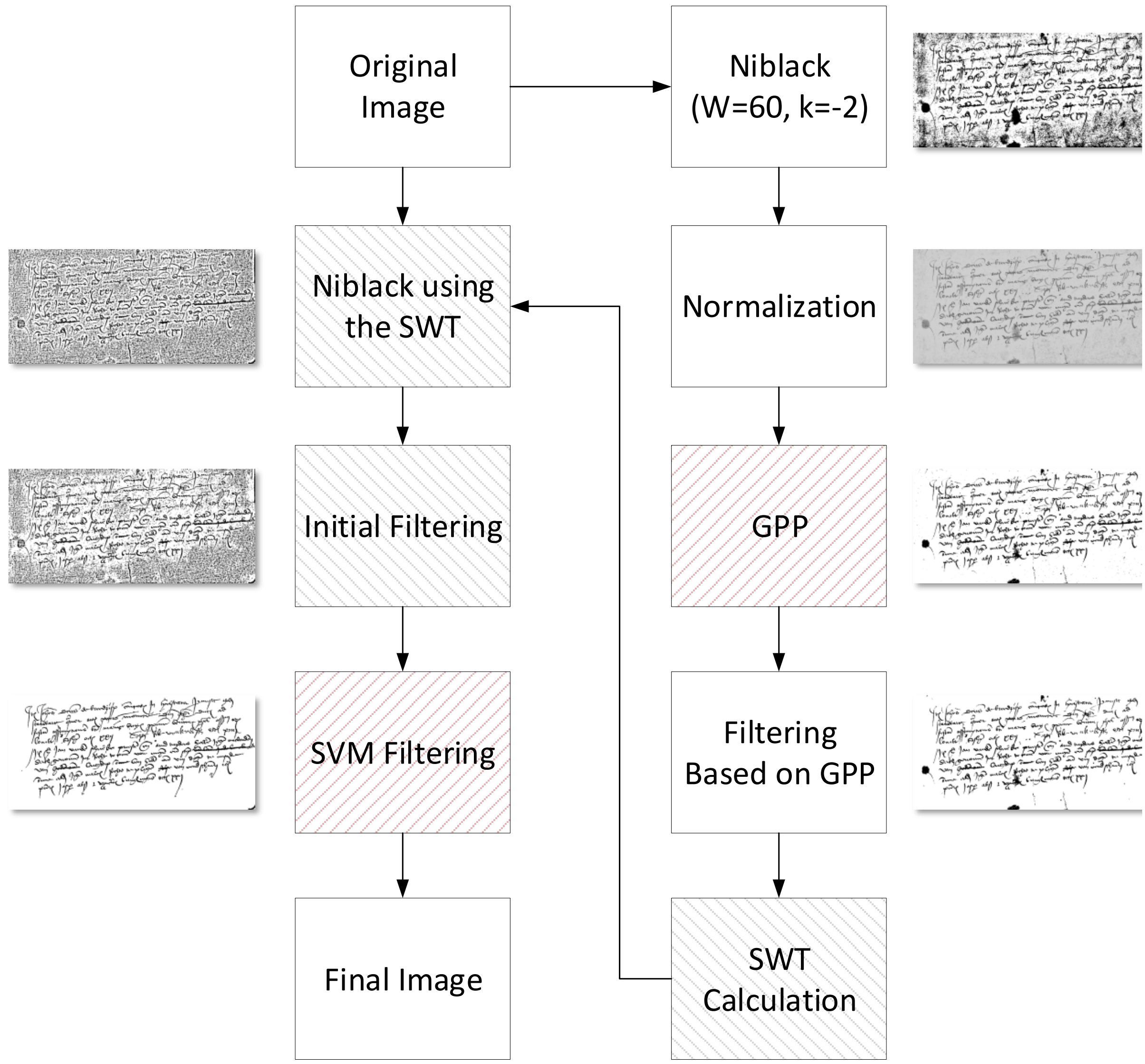

Figure 1: Architecture Diagram of the M36 Binarization method (with red showing the current year changes)

The DUTH M24 and M36 method tries to solve this problem by applying a filtering step in which a HOG-based SVM classifier is used to make a distinction between faint and bleed-through characters. This resulted in a balanced performance for H-DIBCO 2016 dataset between the top achievements of the DUTH and NCSR approaches and resulted in the top performance of DUTH M24 and M36 approach for DIBCO 2017 dataset and H-DIBCO2018 compared to all other approaches produced in READ.

At Table 1, it is also shown the results of the first ranked algorithm for DIBCO 2017 competition which achieves the best performance. It should be mentioned that this top ranked approach relies upon deep learning.

Figure 2 shows the qualitative differences between M24 and M36, namely better SVM filtering.

The DUTH M36 binarization method is developed in C++11 and is available at github under LGPL-3.0:

https://github.com/Transkribus/DUTH

Table 1: Quantitative performance evaluation of M12, M24 and M36 document image binarization methods for H-DIBCO2016, DIBCO2017 and H-DIBCO208 datasets.

	H-DIBCO2016				**DIBCO2017**				**H-DIBCO2018**			
Framework	FM	pseudo-FM	PSNR	DRD	FM	pseudo-FM	PSNR	DRD	FM	pseudo-FM	PSNR	DRD
DUTH M12	90.62	90.95	19.04	**3.86**	83.80	84.04	15.52	8.52	70.47	70.47	70.47	70.47
DUTH M12 (/w enhance-ment)	86.64	91.39	17.67	4.67	82.44	88.30	15.17	712	70.68	76.75	14.02	14.24
DUTH M24	88.81	**91.62**	18.27	4.47	**85.2**	**87.49**	**16.37**	**5.32**	74.6	75.61	14.5	15.5
DUTH M36	89.02	91.61	18.22	4.13	85.31	88.14	16.42	5.16	86.72	89.63	17.53	5.52
NCSR (i)	88.67	87.61	18.42	5.80	79.27	77.64	14.03	13.51				
NCSR (ii)	**90.84**	89.79	**19.31**	4.00	82.94	81.23	15.18	9.81				
NCSR (iii)	84.40	83.49	16.52	9.29	76.7	75.17	13.15	15.73				
#1 Competi-tion Method	*87.61*	*91.28*	*18.11*	*5.21*	***91.04***	***92.86***	***18.28***	***3.40***	***88.34***	***90.24***	***19.11***	***4.92***

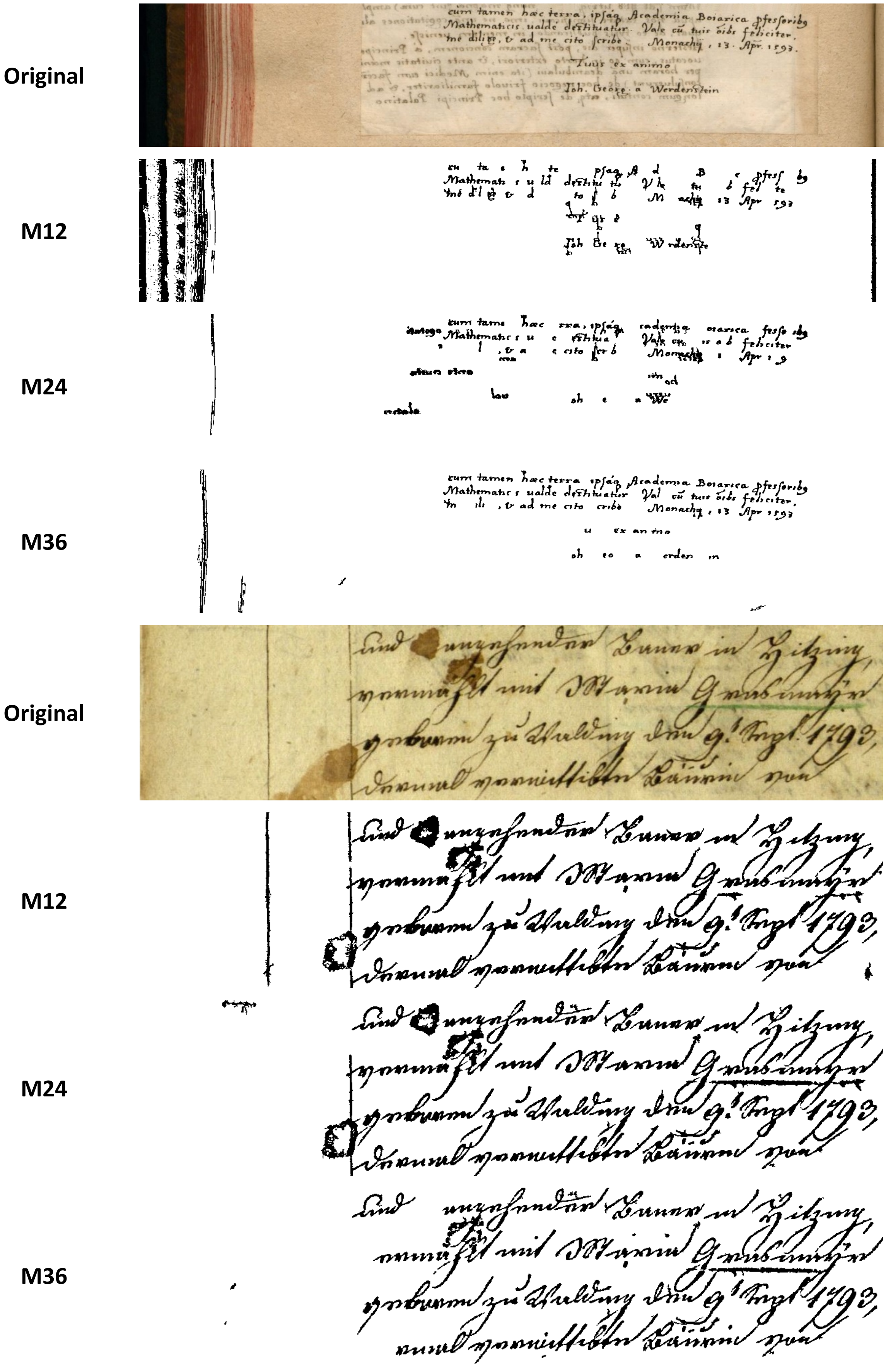

Figure 2: Example results that show qualitative differences between M12, M24, M36

2. References

[BER2001] Bertalmio, M., Bertozzi, A. L., & Sapiro, G. (2001). Navier-stokes, fluid dynamics, and image and video inpainting. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-I). IEEE.

[EPS2010] Epshtein, B., Ofek, E., & Wexler, Y. (2010, June). Detecting text in natural scenes with stroke width transform. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference* on (pp. 2963-2970). IEEE.

[GAT2006] Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. *Pattern recognition*, 39(3), 317-327.

[NTI2014] K. Ntirogiannis, B. Gatos and I. Pratikakis, "A Combined Approach for the Binarization of Handwritten Document Images", *Pattern Recognition Letters - Special Issue on Frontiers in Handwriting Processing*, vol. 35, no.1, pp. 3-15, Jan. 2014.

[PRA2016] Pratikakis, I., Zagoris, K., Barlas, G., & Gatos, B. (2016, October). ICFHR2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016). In *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on* (pp. 619-623). IEEE.

[PRA2017] Pratikakis, I., Zagoris, K., Barlas, G., & Gatos, B. (2017, November). ICDAR 2017 Document Image Binarization Contest (DIBCO 2017). In *Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on* (pp. 1471-1476). IEEE.

[PRA2018] Pratikakis, I., Zagoris, K., Gatos, B., Kaddas P. (2018, August). ICFHR2018 Handwritten Document Image Binarization Contest (H-DIBCO 2018). In *Frontiers in Handwriting Recognition (ICFHR), 2018 16th International Conference on* IEEE.

[VAP2013] Vapnik, V. (2013). *The nature of statistical learning theory.* Springer science & business media.