ICFHR 2018 Competition on Handwritten Document Image Binarization (H-DIBCO 2018)

Ioannis Pratikakis[1], Konstantinos Zagoris[1], Panagiotis Kaddas[2] and Basilis Gatos[2]

[1]Visual Computing Group,
Department of Electrical and Computer Engineering,
Democritus University of Thrace,
67100 Xanthi, Greece
e-mail: {ipratika, kzagoris}@ee.duth.gr

[2]Computational Intelligence Laboratory,
Institute of Informatics and Telecommunications,
NCSR "Demokritos",
15310 Athens, Greece,
e-mail: {pkaddas, bgat}@iit.demokritos.gr

Abstract — **H-DIBCO 2018 is the international Handwritten Document Image Binarization Contest organized in the context of ICFHR 2018 conference. The general objective of the contest is to record recent advances in document image binarization using established evaluation performance measures. This paper describes the contest details including the evaluation measures used as well as the performance of the 8 submitted methods along with a brief description of each method.**

Keywords - handwritten document image, binarization, performance evaluation

I. Introduction

A pre-processing stage of the handwritten document image analysis pipeline is image binarization, according to which the pixels are classified into text and background. It is a crucial stage since it affects further stages of character recognition or keyword spotting. The evaluation of a binarization method aids in verifying its effectiveness and studying its algorithmic behaviour. In this respect, it is imperative to create a framework for benchmarking purposes, i.e. a benchmarking dataset along with an objective evaluation methodology in order to capture the efficiency of current image binarization practices for handwritten document images. To this end, following the success of DIBCO series competitions dedicated to handwritten document images, i.e. H-DIBCO 2010 [1], H-DIBCO 2012 [2], H-DIBCO 2014 [3], H-DIBCO 2016 [4] organized in conjunction with ICFHR 2010, 2012, 2014 and 2016, respectively, the follow-up of these contests, namely H-DIBCO 2018 is organized in the framework of ICFHR 2018. As in previous contests, we focused on the evaluation of handwritten document image binarization methods using a variety of scanned handwritten documents that contain representative degradations (e.g. variable background intensity, shadows, smear, smudge, low contrast, bleed-through or show-through) for which we created the binary image ground truth. The authors of submitted methods registered in the competition and downloaded representative samples along with the corresponding ground truth from previous DIBCO contests available in the competition's site (https://vc.ee.duth.gr/h-dibco2018/). In the sequel, all registered participants were required to submit their binarization executable. After the evaluation of all candidate methods, the testing dataset which comprises 10 handwritten images, the associated ground truth as well as the evaluation software are publicly available at the following link: https://vc.ee.duth.gr/h-dibco2018/benchmark.

II. Methods and Participants

Seven (7) research groups have participated in the competition with eight (8) distinct algorithms (Participant 3 submitted two algorithms). Brief descriptions of the methods are given in the following (the order of appearance is the chronological order of the algorithm's submission).

1) Hubei University of Technology, Wuhan, P.R. China (*XIONG Wei, XIONG Zijie, JIA Xiuhong, LI Min*)

First, the morphological bottom-hat transform is performed to compensate the document background with a disk-shaped structuring element, the size of which is determined by the stroke width transform (SWT) [5]. Then, the Howe's binarization method [6] is applied on the compensated document images to further segment the foreground and background pixels. Finally, an image post-processing is carried out to eliminate isolated noise and preserve text stroke connectivity. The proposed technique has been trained to tune the parameters based on the handwritten document images of DIBCO 2017 benchmark dataset.

2) Document Image and Pattern Analysis (DIPA) Center, Islamabad, Pakistan (*Syed Ahsen Raza*)

The proposed method for handwritten documents binarization consists of three main steps.
1. Pre-processing. In this step document image is converted to gray-scale and a bi-cubic interpolation is applied. Then, the image is segmented into small blocks for further processing.
2. Thresholding. In this step adaptively a threshold is calculated for each block using a modified threshold which is calculated using the combination of Sauvola's, Feng's and Otsu's thresholds. This adaptive threshold T is then applied

978-1-5386-5875-8/18/$31.00 ©2018 IEEE
DOI 10.1109/ICFHR-2018.2018.00091

to each block for binarization. Then each block is combined to give the whole binarized image.
3. Post-processing. In this step conditional aspect ratio is used to remove small unwanted artifacts considering them as noise. Two iterations of median filter using linear structuring element are also used to remove very tiny spots from the final image. After these steps, the original image has been completely binarized.

3) **Larbi Tebessi University, Tebessa, Algeria, Algiers, Algeria** (*Abdeljalil Gattal and Chawki Djeddi*)

Method a: This work is mainly based on the well-known Sauvola's method [7]. It consists of making the method independent of parameters such as k and the size of the sliding window. In step 1, the input grayscale image is binarized with Sauvola's algorithm by setting up k=0.2 (optimal value) and selecting different size of window {7x7,13x13,19x19,..,49x49} pixels. In step 2, each binarized image with different size of window is compared to the binarized image with Basic Image Features (BIFs) [8][9][10] by computing metrics as F-measure and then the best binarized image is selected.

Method b: This work is mainly based on the k-means clustering algorithm which is implemented through city-block distance metric. The data is split into 3 clusters as background, text and noise.

4) **KOBE University, Graduate School of Maritime Sciences, Japan** (*Yuichi Nakata, Naoki Tanaka*)

The proposed method converts a grayscale image to four feature maps which are produced as follows: ridge detected via second order derivatives, absolute values of difference between each pixels and an average of the image, relative pixel values from neighbour local area, and local mean value weighted by gradient strength. These maps are combined using maximum or multiple operator. After that, the unified map is emphasized on its density. Finally the map is normalized and binarized by a fixed threshold.

5) **Ain Shams university, Cairo, Egypt** (*Usama Wadie Aziz Mousa*)

The proposed technique consists of five stages. First, we create two edge detection images after creating the gray image where they are deduced by using a mix of edge detection methods, namely Canny [11], Sobel [12] and Bolan [13]. Second, we create a binarization image by the Niblack method [14]. Third, we select seeds from the Canny edge detection image and create the data part image by using markov random field [15] and the integral image [16]. Fourth, all of these images are combined where any black pixel in the second edge detection image, niblack image or the data part image attachment to a black seed will become a black seed and we repeat until there is no other seed to test. Then the last stage is the post-processing stage where in the combined image, any white pixel attachment to a black pixel with gray value less than the gray value for this black pixel in the gray image will be black to produce the binarized image.

6) **Syiah Kuala University, Indonesia** (*Khairun Saddami, Khairunnisa, Putri Afrah, Merita, M. Irhas and Tuwanku Mohd Iqbal*)

The proposed method is called as improved NICK binarization method (iNICK). In fact, the most Niblack inspiring method has k parameter that is set manually. In our method, we set the k value as an automatic parameter. We use image standard deviation as a guideline for determining the automatic k value. This method has been proposed in ICDAR 2017 [17].

7) **CAI, School of Software, University of Technology Sydney; CVPR Unit, Indian Statistical Institute** (*Chandranath Adak and Michael Blumenstein; Bidyut B. Chaudhuri*)

The proposed method combines local and global adaptive binarization which is motivated by [18] and [19]. We extract the ink pixels from the background using an inpainting-based background estimator and perform an image normalization. This inpainting method is based on a hybrid sparse representation. A dilated version of Niblack [14] output is employed as an inpainting mask. We perform the global binarization on the normalized image similar to [18]. The local binarization is accomplished by an adaptive image contrast processing comprised of ink-edge detection and local threshold estimation as like [19]. Finally, the local and global binarized outputs were merged and very small components were removed.

III. EVALUATION MEASURES

For the evaluation, the measures used comprise an ensemble of measures that are suitable for evaluation purposes in the context of document analysis and recognition. These measures consist of (i) F-Measure (*FM*), (ii) pseudo-FMeasure (F_{ps}), (iii) *PSNR* and (iv) Distance Reciprocal Distortion (*DRD*).

A. F-Measure

$$FM = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{1}$$

where $\text{Recall} = \frac{TP}{TP+FN}$, $\text{Precision} = \frac{TP}{TP+FP}$

TP, FP, FN denote the True Positive, False Positive and False Negative values, respectively.

B. pseudo-FMeasure

Pseudo-FMeasure $\boldsymbol{F_{ps}}$ is introduced in [20] and it uses pseudo-Recall $\boldsymbol{R_{ps}}$ and pseudo-Precision $\boldsymbol{P_{ps}}$ (following the same formula as F-Measure). The pseudo Recall/Precision metrics use distance weights with respect to the contour of the ground-truth (GT) characters. In the case of pseudo-Recall, the weights of the GT foreground are normalized according to the local stroke width. Generally, those weights are delimited between [0,1]. In the case of pseudo-Precision, the weights are constrained within an area that expands to the GT background taking into account the stroke width of the nearest GT component. Inside this area, the weights are greater than one (generally delimited between (1,2]) while outside this area they are equal to one.

C. PSNR

$$PSNR = 10\log(\frac{C^2}{MSE}) \tag{2}$$

where $MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y)-I'(x,y))^2}{MN}$

PSNR is a measure of how close is an image to another. The higher the value of *PSNR*, the higher the similarity of the two images. Note that the difference between foreground and background equals to *C*.

D. Distance Reciprocal Distortion Metric (DRD)

The Distance Reciprocal Distortion Metric (*DRD*) has been used to measure the visual distortion in binary document images [21]. It properly correlates with the human visual perception and it measures the distortion for all the ***S*** flipped pixels as follows:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN} \tag{3}$$

where NUBN is the number of the non-uniform (not all black or white pixels) 8x8 blocks in the GT image, and DRD_k is the distortion of the *k*-th flipped pixel that is calculated using a 5x5 normalized weight matrix $\boldsymbol{W_{Nm}}$ as defined in [21]. DRD_k equals to the weighted sum of the pixels in the 5x5 block of the ***GT*** that differ from the centered kth flipped pixel at (x,y) in the binarization result image ***B*** (Eq. 4).

$$DRD_k = \sum_{i=-2}^{2}\sum_{j=-2}^{2} |GT_k(i,j) - B_k(x,y)| \times W_{Nm}(i,j) \tag{4}$$

IV. EXPERIMENTAL RESULTS

The H-DIBCO 2018 testing dataset consists of 10 handwritten document images for which the associated ground truth was built manually for the evaluation. The selection of the images in the dataset was made so that representative degradations appear. The document images of this dataset originate from the READ project [22] included in various collections such as:

1. the protocols of the city or municipal council of Bozen, a city in South Tyrol, today Northern Italy from the 15th century to the 19th century [23].
2. Reconstructed Alexander von Humboldt's "Kosmos-Lectures" published in cooperation with the German Text Archive/Deutsches Textarchiv at the Berlin-Brandenburg Academy of Sciences and Humanities [24].
3. Archive Bistum Passau (ABP) collection that contains sacramental register and index pages like baptism, marriage and death entries.
4. Archivio di Stato di Venezia collection that mainly comprises tax and death records from Venice in 1740.

The images that comprise the testing dataset are shown in Figure 1(a).

The evaluation was based upon the four distinct measures presented in Section III. The detailed evaluation results along with the final ranking are shown in Table I. The final ranking was calculated after first, sorting the accumulated ranking value for all measures for each test image. The summation of all accumulated ranking values for all test images denote the final score which is shown in Table I at column "Score". Additionally, the evaluation results for the widely used binarization techniques of Otsu [25] and Sauvola [7] are also presented. Overall, the best performance is achieved by ***Method 1*** which has been submitted by **Xiong Wei, Xiong Zijie, Jia Xiuhong, Li Min** affiliated to ***Hubei University of Technology, Wuhan, P.R. China***. The binarization results of this algorithm for each image of the testing dataset are shown in Fig. 2(a).

For providing a direct link to previous competitions, Table II presents the performance of this year's methods in the DIBCO 2017 dataset.

TABLE I. DETAILED EVALUATION RESULTS FOR ALL METHODS SUBMITTED TO H-DIBCO 2018 AND EVALUATED ON H-DIBCO 2018 DATASET

Rank	Method	Score	FM	F_{ps}	PSNR	DRD
1	1	4	88.34	90.24	19.11	4.92
2	7	11	73.45	75.94	14.62	26.24
3	2	12	70.01	74.68	13.58	17.45
4	3b	14	64.52	68.29	13.57	16.67
5	6	21	46.35	51.39	11.79	24.56
6	5	22	56.08	60.68	11.5	28.99
7	3a	29	43.36	45.92	10.42	40.8
8	4	30	41.87	46.31	10.38	37.36
-	Sauvola	-	67.81	74.08	13.78	17.69
-	Otsu	-	51.45	53.05	9.74	59.07

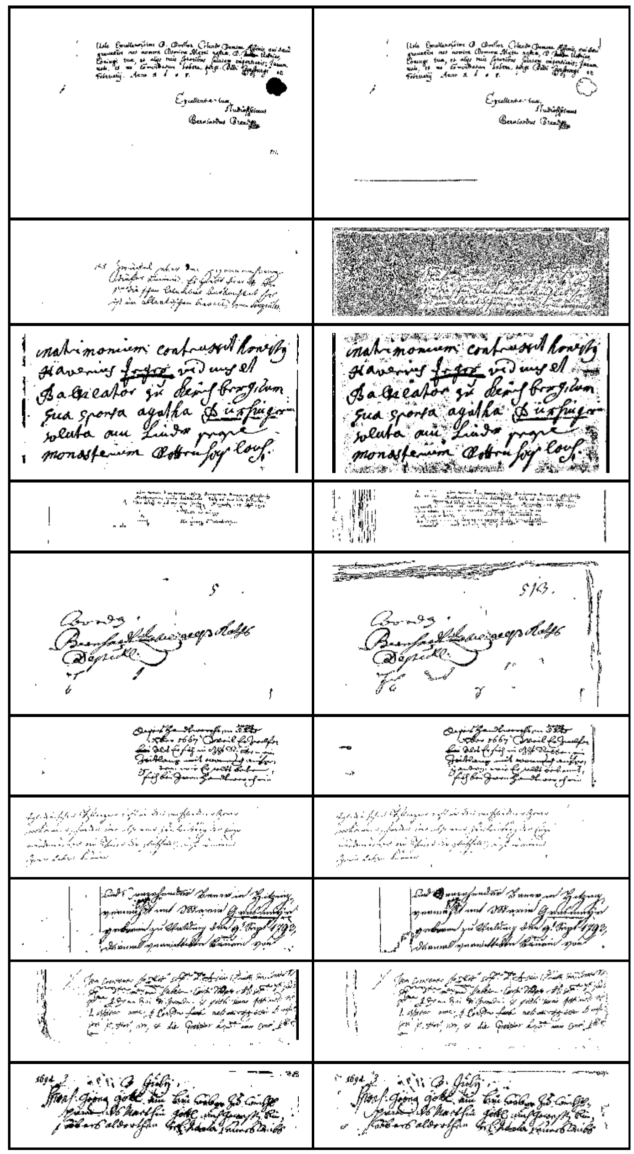
TABLE II. DETAILED EVALUATION RESULTS FOR ALL METHODS SUBMITTED TO H-DIBCO 2018 AND EVALUATED ON DIBCO 2017 DATASET

Rank	Method	Score	FM	F_{ps}	PSNR	DRD
1	1	6	89.37	90.17	17.99	5.51
-	5	6	86.73	92.94	16.58	4.47
2	7	15	84.36	87.34	15.72	7.56
3	6	16	80.75	87.24	15.38	6.22
4	3b	18	82.43	86.74	15.28	6.97
5	4	25	77.48	84.14	15.11	7.02
6	2	29	79.41	82.62	14.04	10.7
7	3a	29	80.26	82.47	14.68	11.16

-	Sauvola	-	86.61	88.67	17.80	5.56
-	Otsu	-	82.52	86.85	16.42	7.49
DIBCO 2017 Winners						
1	10	-	91.04	92.86	18.28	3.40
2	17a	-	89.67	91.03	17.58	4.35
3	12	-	89.42	91.52	17.61	3.56

(a) (b)

Figure 1. (a) The H-DIBCO 2018 testing dataset (b) Their corresponding ground-truth.



(a) (b)

Figure 2. Binarization results from (a) the winner algorithm and (b) the algorithm that ranked at the second position.

V. CONCLUSIONS

The final evaluation can guide us to draw several conclusions that could operate as a fruitful feedback for the research community working on improving handwritten document image binarization. This year's participating methods did not follow the general recent trend which was also adopted at DIBCO 2017, namely the use of supervised approaches with deep neural network architectures. Instead, there were only few training options that appeared in the submitted methods which however, boosted performance among the existing participating methods leading them at the top ranked positions. Similar to the case in h-DIBCO 2016 [4], the winner method relies upon the same already published Howe's method [6] coupled by an effective pre-processing as well as post-processing stage. It has been a gold standard during the DIBCO series evolution that whatever the core approach is, the use of pre-processing and post-processing stages has a major impact on the success of the binarization process. Still, standard approaches like the global Otsu algorithm [25] and the locally adaptive Sauvola algorithm [7] are fully involved in the proposed approaches which in most of the cases they are taken into account in a combined fashion.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union's H2020 Programme READ under grant agreement n° 674943.

REFERENCES

[1] I. Pratikakis, B. Gatos and K. Ntirogiannis, "H-DIBCO 2010 – Handwritten Document Image Binarization Competition", 12th International Conference on Frontiers in Handwriting Recognition (ICFHR'10), Kolkata, India, pp. 727-732, 2010.

[2] I. Pratikakis, B. Gatos and K. Ntirogiannis, "ICFHR 2012 Competition on Handwritten Document Image Binarization (HDIBCO 2012)", 13th International Conference on Frontiers in Handwriting Recognition (ICFHR'12), pp. 813-818, Bari, Italy, 2012.

[3] K. Ntirogiannis, B. Gatos and I. Pratikakis, "ICFHR 2014 Competition on Handwritten Document Image Binarization (H-DIBCO 2014)", 14th International Conference on Frontiers in Handwriting Recognition (ICFHR'14), Crete island, Greece, pp. 809-813, IEEE Computer Society Press, ISBN-978-1-4799-4335-7, 2014.

[4] I. Pratikakis, K. Zagoris, G. Barlas and B. Gatos (2016), ICFHR 2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016), 15th International Conference on Frontiers in Handwriting Recognition (ICFHR'16), October 23-26, 2016, Shenzhen, China, pp. 619-623, IEEE Computer Society Press.

[5] B. Epshtein, E. Ofek, Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, 2010, pp. 2963-2970. doi: 10.1109/cvpr.2010.5540041

[6] N. Howe, "Document Binarization with Automatic Parameter Tuning," International Journal on Document Analysis and Recognition, vol. 16, no. 3, pp. 247-258, 2013.

[7] J. Sauvola, M. Pietikainen, "Adaptive document image binarization", Pattern Recognition, Vol. 33, No. 2, pp. 225-236, 2000.

[8] L. Griffin, M. Lillholm, M. Crosier, and J. van Sande. Basic image features (BIFs) arising from approximate symmetry type. Scale Space and Variational Methods in Computer Vision, pages 343-355, 2009.

[9] A. Gattal, C. Djeddi, Y. Chibani and I. Siddiqi, "Isolated Handwritten Digit Recognition Using oBIFs and Background Features", 12th IAPR International Workshop on Document Analysis Systems (DAS-2016), Santorini, Greece, 2016. pp. 305-310

[10] A. Gattal, C. Djeddi, I. Siddiqi and Y. Chibani," Gender Classification from Offline Multi-Script Handwriting Images using Oriented Basic Image Features (oBIFs)", Expert Systems with Applications, Vol. 99, pp: 155–167, 2018.

[11] Canny, J. F. (1986), A computational approach to edge detection. IEEE Trans Pattern Analysis and Machine Intelligence, 8(6), 679-698.

[12] R. Gonzalez and R. Woods Digital Image Processing, Addison Wesley, 1992, pp 414 - 428.

[13] Su, B., Lu, S., & Tan, C. L., (2010), Binarization of Historical Document Images Using the Local Maximum and Minimum. In Workshop on Document Analysis Systems (DAS) (pp. 159-166). Boston, USA.

[14] W. Niblack, "An Introduction to Digital Image Processing", Prentice Hall, Englewood Cliffs, 1986.

[15] Y. Boykov and V. Kolmogorov, "An experimental comparison of mincut/max-flow algorithms for energy minimization in vision," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, no. 9, pp. 1124–1137, 2004.

[16] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," CVPR, pp. 511– 518, 2001.

[17] K., Saddami, K., Munadi, S. Muchallil, and F. Arnia. Improved Thresholding Method for Enhancing Jawi Binarization Performance. In Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on (Vol. 1, pp. 1108-1113). IEEE.

[18] K. Ntirogiannis, B. Gatos, I. Pratikakis, "A Combined Approach for the Binarization of Handwritten Document Images", Pattern Recognition Letters, vol. 35, pp. 3-15, 2014.

[19] B. Su, S. Lu, C.L. Tan, "Robust Document Image Binarization Technique for Degraded Document Images", IEEE Transactions on Image Processing, vol. 22, no. 4, pp. 1408-1417, 2013.

[20] K. Ntirogiannis, B. Gatos and I. Pratikakis, "Performance Evaluation Methodology for Historical Document Image Binarization", IEEE Transactions on Image Processing, vol. 22, no.2, pp. 595-609, Feb. 2013.

[21] H. Lu, A. C. Kot and Y.Q. Shi, "Distance-Reciprocal Distortion Measure for Binary Document Images", IEEE Signal Processing Letters, vol. 11, No. 2, pp. 228-231, 2004.

[22] READ project (http://read.transkribus.eu/)

[23] Bohisto Bozen Digitized Archives (http://stadtarchiv-archiviostorico.gemeinde.bozen.it/bohisto/Archiv;jsessionid=79E541FF375A0902215E0BC62A3B45A5)

[24] Hidden Kosmos — Reconstructing Alexander von Humboldt's »Kosmos-Lectures« (2014-16) (https://www.culture.hu-berlin.de/de/forschung/projekte/hidden-kosmos/)

[25] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE transactions on Systems, Man and Cybernetics, Vol. 9, No. 1, pp. 62-66, 1979.