2019

ICDAR 2019 Time-Quality Binarization Competition

Rafael Dueire Lins
CIn – UFPE / DC - UFRPE

Ergina Kavallieratou
University of the Eagian

Elisa Barney Smith
Boise State University

Rodrigo Barros Bernardino
UFPE

Darlisson Marinho de Jesus
UFPE

Publication Information

Dueire Lins, Rafael; Kavallieratou, Ergina; Barney Smith, Elisa; Barros Bernardino, Rodrigo; and Marinho de Jesus, Darlisson. (2019). "ICDAR 2019 Time-Quality Binarization Competition". In *2019 International Conference on Document Analysis and Recognition (ICDAR)* (pp. 1539-1546). IEEE. https://doi.org/10.1109/ICDAR.2019.00248

This is an author-produced, peer-reviewed version of this conference proceeding. The final, definitive version of this document can be found online at *2019 International Conference on Document Analysis and Recognition (ICDAR)*, published by IEEE. Copyright restrictions may apply. https://doi.org/10.1109/ICDAR.2019.00248

ICDAR 2019 Time-Quality Binarization Competition

Rafael Dueire Lins
CIn – UFPE / DC - UFRPE
Recife, Brazil
rdl.ufpe@gmail.com

Ergina Kavallieratou
University of the Eagian
Mytilene, Greece
kavallieratou@aegean.gr

Elisa Barney-Smith
Boise State University
Boise, U.S.A.
ebarneysmith@boisestate.edu

Rodrigo Barros Bernardino
UFPE
Recife, Brazil
rodrigobernardino@gmail.com

Darlisson Marinho de Jesus
UFPE
Recife, Brazil
dmj.ufpe@gmail.com

***Abstract* — The ICDAR 2019 Time-Quality Binarization Competition assessed the performance of seventeen new together with thirty previously published binarization algorithms. The quality of the resulting monochromatic image and the execution time were assessed. Comparisons were on both in "real-world" and synthetic scanned images, and in documents photographed with four models of widely used portable phones. Most of the submitted algorithms employed machine learning techniques and performed best on the most complex images. Traditional algorithms provided very good results at a fraction of the time.**

Keywords - Binarization; documents; algorithms; quality evaluation, performance evaluation, historical documents.

I. Introduction

The process by which a color image is converted into its monochromatic version is called binarization. Black and white images are much easier for computers to process, require less storage space and bandwidth when transmitting through computer networks. There is an ever growing variety of binarization methods, which produce images with good quality not only for visual inspection, but also for many applications within the context of document analysis. The huge number of legacy paper documents that are being digitized and processed for information extraction and classification claim that binarization algorithms should not only provide good quality monochromatic images, but that they also must be fast.

The recent article by Tim Roughgarden [1], Beyond Worst-Case Analysis, claims that "*the need for deeply understanding when algorithms work (or not) has never been greater*". In the specific case of document binarization algorithms, the first author of this contest report has long claimed that "no binarization algorithm is good for all kinds of text document images" [2][3]. Thus, in order to make fair comparisons between the time-quality performance of binarization algorithms, it is important to assess the algorithms on different clusters of documents. The end-user should better match the document (or batch of documents) one wants to binarize with such a cluster, to find which algorithms perform best and if they meet the time performance adequate for the proposed document-processing pipeline.

Choosing the images for the different test-sets is the starting point of such a complex problem. Issues to consider include: How to get good-quality ground-truth images to compare with? What kinds of noises [4] affect the original document image? Which was the digitalization hardware or process used? Which measure provides the "best-quality" assessment? How to compare the time-performance of algorithms that were implemented in different programming languages and execute either on different operating systems and hardware platforms? How the time-quality performance of the newly proposed algorithms compare with the more "classical" ones? Assessing the time-quality performance of binarization algorithms in the fairest possible way, addressing all these points raised, is the concern of the ICDAR 2019 Time-Quality Binarization Competition. The competing algorithms will be ranked in each of the test sets according to the quality of the produced images, first. The best quality performing algorithms will have their times compared, in a second step.

II. Participants

Seventeen research groups from over twenty different countries spread in the five continents enrolled in this competition. During the evaluation process, three groups had to withdraw their participation due to implementation problems found. Two of the groups presented three different binarization algorithms. A brief description of the remaining groups and their algorithms follows, in the order of their enrollment in this competition. The indicated affiliation is of the first member of the team.

A. USP - University of São Paulo, São Paulo, Brazil (Nury Yuleny Arosquipa Yanque, Gustavo Enrique Salazar Torres and Roberto Hirata Junior)

This solution uses supervised machine learning techniques. The features vectors are composed of a combination of: binary output values from state-of-art methods like Otsu, Niblack, Sauvola, Su and Howe; binary image output from GridLSTM, proposed by Wesphal; family of texture features called 'Relative Darkness Index', proposed by Wu; and the grayscale intensity value of the original image. These vectors are extracted for every pixel of the gray-scale original image. The dataset composed by these vectors and the foreground/background labels is used to train an

XGBoost classifier that predicts if the pixel belongs to a text or background region. The output image is post-processed by a morphological operation in order to improve the quality of the image.

B. Qatar University, Qatar

(Younes Akbari, Alceu S. Britto Jr., Somaya Al-maadeed and Luiz S. Oliveira)

This binarization methodology relies upon a Segnet network architecture which is fed by multichannel images that correspond to the original image and the image approximations based on the coefficients of three sub bands [5] and the image binarized by the structural symmetric pixels (SSPs) method [6]. Multichannel images were implemented and used as network inputs based on two approaches: single and multiple networks.
Method (1): The original image is decomposed into wavelet sub bands, the original image binarized by the structural symmetric pixels (SSPs) method (single network).
Method (2): Variation of (a) with multiple networks.
Method (3): Variation of (a) where fewer channels are used to reduce computational cost.

C. DLSI - Universidad de Alicante, Spain

(Jorge Calvo-Zaragoza, Antonio-Javier Gallego)

Image binarization is treated as a two-class classification task at the pixel level. The presented strategy basically consists in learning which label must be given to every single pixel of the image. Foreground and background. Pixels are determined by Convolutional Neural Networks [7]. This image-to-image convolutional architecture is trained to convert an input image into its binarized version. This has a number of advantages such that the classification of each pixel of the image is not produced independently, but also takes into account the label to be assigned to its neighbors. In addition, several pixels can be processed at the same time, thereby leading to higher efficiency than a pixelwise classification approach. An image is passed through it, producing outputs 0 or 1. A thresholding process converts the scores into binary values.

D. Hubei University of Technology, China

(Xiuhong Jia, Wei Xiong, Jingyi Jin, Zijie Xiong, Min Li)

This method, called Doc-DLinkNet, consists of three main steps. First, the original image is cropped into 256×256 patches. Data augmentation strategies such as shape shift and color shift are applied. Second, a D-LinkNet architecture [8] is adopted and trained by using document image patches as input and the corresponding binary maps as ground truths. D-LinkNet is a semantic segmentation neural network, which involves dilated convolution and pre-trained encoder. Finally, the Principal Components Analysis (PCA) method is used to perform image dimensionality reduction and feature extraction, and then generates the final results according to the optimal parameters learned from the training procedure.

E. Universitas Syiah Kuala, Indonesia

(Khairun Saddami)

Method (1) iNICK: An extension of the NICK binarization method [9]. The image standard deviation is used to determine the k value as $k=-\sigma/(255-1.5\sigma)$, where σ is the image standard deviation that represents the image contrast.
Method (2) CNW: Combination of Niblack and Wolf [10]. The threshold $T=(2m+mk((\sigma/m)-(\sigma/R)-1))/2$, where σ is the image standard deviation, m is the mean of local window, R is the maximum standard deviation, k=0.35.
Method (3) CLD: Combined the local adaptive and global thresholding formulas, as described in [11].

F. Larbi Tebessi University, Algeria

(Abdeljalil Gattal)

The proposed method [12] is based on the k-means clustering algorithm, classifying the given data set from image (Img) into three clusters: background, text and noise. City-block distance is used for calculating the distance of pixel value from the particular centroid.

G. Australian National University, Australia

(Hanif Rasyidi)

This model uses a fully convolutional network [13] to analyze the text pattern on the document, and then applies a pixel-based segmentation to produce a binary text image. The model was trained using 115 images from the DIBCO and Nabuco datasets from DIB (https://dib.cin.ufpe.br).
The proposed model contains three parts: the feature extraction backbone, feature merging, and the final output layer for pixel segmentation. This idea is based on the EAST model [14], which uses different backbone and output layers to detect text in the scene images. In the default setup, the ResNet50 model implements a residual connection to prevent the loss of low level information. A variation of HandwitteNet called HandwriteNet-Mobile, that uses a less-costly MobileNetV2 was used as the backbone. The final output layer produces a F-Score with an "image-like" structure to the input image, where each "pixel" contains a value between [0,1]. A threshold value T=0.8 extracts the binary image from the output layer. The binary output, which is smaller than the image is used to match the input size may produce an edge imperfection in the final binary output The final step is to apply a size correction, which may produce an edge imperfection in the output binary image.

H. Hubei University of Technology, China

(Xiong Wei, Wei Xiong, Min LI, Chuansheng Wang, Laifu Guan)

The Doc-UNet method performs three steps: 1. A morphological bottom-hat transform is carried out to enhance the document image contrast, and the size of a disk-shaped structural element is determined by the stroke width transform (SWT) [15]. 2. A hybrid pyramid U-Net convolutional network [16] is performed on the enhanced

This is an author-produced, peer-reviewed version of this conference proceeding. The final, definitive version of this document can be found online at *2019 International Conference on Document Analysis and Recognition (ICDAR)*, published by IEEE. Copyright restrictions may apply. https://doi.org/10.1109/ICDAR.2019.00248

document images for accurate pixel classification. 3. Otsu binarization.

I. Jadavpur University, India

(Showmik Bhowmik, Ram Sarka, and David Doermann)

The GiB method [17], inspired by game theory, performs background separation and binarization. A customized "inpainting" method [18] is used on the grayscale converted input image to remove background information. After that, a two-player game is designed and implemented at the pixel level. An overlapping 3x3-pixel window scans the image. For each window the central pixel is considered the first player and all the other pixels, the second one. The Nash equilibrium state is computed for each game. Two other features are also computed: the central pixel intensity and intensity difference between central pixel and the pixel having maximum intensity among its 8 neighbors. Based on these three features, all the pixels are grouped into three clusters, dynamically using the K-means clustering algorithm. The cluster with the lowest variance is considered the background. If the ratio between the variances of the two remaining unlabeled clusters is less than a threshold, it indicates that they are similar, and are merged. Otherwise, the cluster is foreground.

J. Jadavpur University, India

(Soulib Ghosh, Suman Kumar Bera, Showmik Bhowmik, Ram Sarkar)

This method follows a two-stage approach: background separation and binarization. For background separation, a superset of the foreground is estimated by Niblack's method which acts as a mask. Then, the background surfaced image is obtained which is followed by image normalization. The binarization technique comprises a clustering combination approach. A combination of three popular clustering algorithms is adopted.

K. Havard University, United States

(Sheng He)

This program is based on Tensorflow and the algorithm DeepOtsu [19]. The neural network is trained to learn the degradations in document images and produce uniform images of the degraded input images, which in its turn allows the network to refine the output iteratively. The stacked refinement (SR), which uses a stack of different neural networks for iterative output refinement, is applied. The binarization map is obtained through use of a global Otsu threshold.

L. Inner Mongolia University, China

(Xu Huali)

This solution used a generative adversarial network. The model consists of a generator and discriminator sub-networks, which are trained in an adversarial way. The generator yields the binarized image, and the discriminator distinguishes the image generated from the real binarized image. The generator adopts a U-Net like structure, in which the encoder uses convolution operation and LeakyReLU as activation function while the decoder uses deconvolution operation and ReLU as the activation function. There are 14 layers in the generator. The discriminator consists of three modules in the form of Convolution-BatchNorm-Relu. Once trained, the generator can be used for image binarization.

The algorithm was trained with images from DIBCO, augmented by flipping, rotating (180°), and changing the values of the RGB channels. The images and their corresponding binarized ones are segmented into 512*512 image blocks. In total, 8,112 pairs of blocks are used for training and 1,000 pairs of blocks for validation. The blocks in each pair are concatenated together to train the model. For testing, each image is segmented into 512*512 blocks and converted into binary individually. The resulting blocks are merged into a complete binary image.

M. Istanbul Technical University, Turkey

(Yasin Yildirim)

This approach consists of three main steps: preprocessing, optimization, and thresholding. Preprocessing: the input image is converted into grayscale and then a 9x9 adaptive Wiener filter is applied to reduce noise. A conjugate gradient descent method is used for optimization. The computation is done on the downscaled pyramidal image version for fast computation. The downscaling ratio is controlled by parameter called 'reduceFactor' (default value is 5). Thresholding: Otsu thresholding is applied as a final step to binarize the document image.

N. Traditional Algorithms

Thirty widely used binarization algorithms, available at the DIB platform (https://dib.cin.ufpe.br), have also been considered in this time-quality analysis. Twenty-three of them are among the top ten in quality for the different test sets: Bernsen [20], Bradley [21], da Silva-Lins-Rocha (dS-L-R) [22], Ergina-G. [25], Ergina-L. [26], Howe [23], Huang [24], Intermodes [27], IsoData [28], Johannsen-Bille (Johann) [29], Kapur-SW [30], Li-Tam [31], Mello-Lins (M-L) [32], Minimum [27], Moments [33], Nick [30], Otsu [34], RenyE. [35], Sauvola [36], Shanbag [37], Triangle [43], Wolf [38], and Wu-Lu [39].

III. TEST SETS

Three large test sets of document images were selected for assessing the binarization algorithms, all publicly available IAPR TC10-TC11 DIB dataset (https://dib.cin.ufpe.br)

Nabuco: 20 historical document images from the late 19^{th} to the early 20^{th} centuries belonging to the bequest of letters from Joaquim Nabuco [40], including handwritten and machine typed documents scanned at 200 dpi resolution. The documents in this set were also clustered according to the tone of the paper: light, medium, and dark. Figure 1 presents a sample of the documents in each cluster. The ground-truth image was generated by applying all the 30 "classical" binarization algorithms also available at the DIB platform. The produced binary images, were visually inspected by the DIB-team, to select the best one, then filtered to remove eventual salt-and-pepper noise, and hand corrected, if needed.

DIB: controlled parameter synthetic images with handwritten and machine typed documents. The influence of strength of the back-to-front interference [4] (weak α=0.6, medium α=0.7, and strong α=0.8) and three paper tones were analyzed. Figure 2 presents a sample of the documents used for the hand-written and machine-typed classes.

Twelve High-quality 300 dpi monochromatic images, which were the seeds for the generation of the synthetic color image, were used as ground-truth.

Camera: The binarization of photographed documents is far more difficult than scanned ones as the resolution of the photo varies from a device to another and is non-uniform due to differences in the distance between the document and the camera, it may suffer the interference from external light sources and even a non-uniform illumination from the in-built strobe flash. This test set encompasses 72 documents, obtained from four different models of portable cell-phones, whose specifications are presented in Table 1. Besides the device model, the documents in this set were clustered according to having the in-built strobe-flash set as "on" or "off". Figure 3 presents samples of the documents used in this test set. The assessment methodology from [41] was adopted for the challenging task of assessing the resulting image quality.

Table 1: Summary of device camera specifications

	Moto Z2	Iphone 6	Iphone SE	Galaxy N4
Megapixels	12	8	12	16
Flash	Dual led	Dual led	Dual led	Dual led
Aperture size	f/1.7	f/2.2	f/2.2	f/2.2
Sensor size	-	1/3 inch	1/3 inch	1/2.6 inch
Pixel size	1.4 µm	1.22 µm	1.22 µm	1.12 µm

IV. Quality Evaluation Methods

To evaluate the binarization algorithms relative to image quality the scanned documents were clustered according to their features (print type, paper texture luminosity, intensity of back-to-front interference, etc.). The quality of the binary images was compared using the PSNR, DRDM, F-Measure (FM) and pseudo-FMeasure (Fps) [44], and Cohen´s Kappa [42] [45]. The final ranking is defined by sorting the ranking summation in ascending order. The consistency of the global ranking with a carefully made visual inspection was also analyzed.

The analysis of the quality of the camera acquired images is still more complex due to the uneven resolution and illumination. Thus, for the Camera dataset, the quality measure proposed in reference [41], which compares the proportion between the black-to-white pixels in the scanned and photographed binary documents was used. Again, visual inspection was applied to check the consistency of the results obtained.

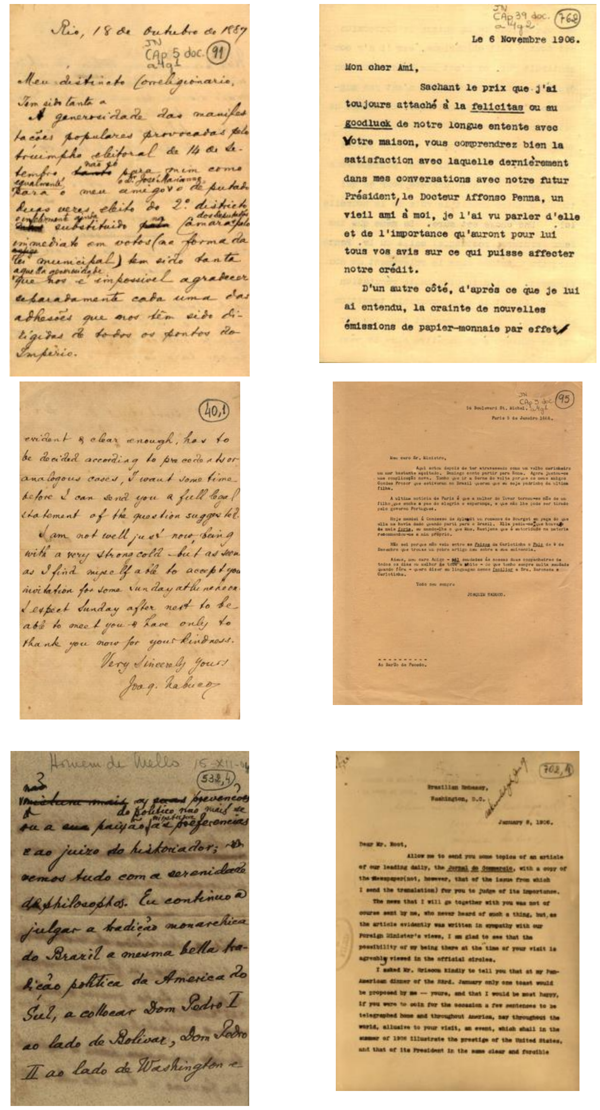

Figure 1: Samples of the images from Nabuco test set, clustered by the printing method (handwritten or machine typed) and color tone of the paper (light, medium, dark).

V. Processing-Time Evaluation

The seventeen new algorithms assessed here were implemented by their authors. The test images were chosen as specified in the Call-for-competitors. Although the test set is a sub-set of the training set, the very large number of documents available most possibly mimic all kinds of "real world" scanned documents, an argument in favor of the expressiveness of the results presented here.

The purpose of the processing time evaluation here is to provide an order of magnitude of time elapsed for binarizing the whole datasets. The training-times for the AI-based algorithms were not computed. The competing

This is an author-produced, peer-reviewed version of this conference proceeding. The final, definitive version of this document can be found online at *2019 International Conference on Document Analysis and Recognition (ICDAR)*, published by IEEE. Copyright restrictions may apply. https://doi.org/10.1109/ICDAR.2019.00248

algorithms were implemented using different programming languages, operating systems, and even for specific hardware platforms such as GPUs. Three different SW/HW platforms were used for the implementation and execution of the competing algorithms:

- **GPU Algorithms - Google Cloud Platform VM**
  OS: Ubuntu 18.04 LTS
  Machine Type: n1-highmem-4
  CPU: Intel Haswell – 4 vCPU   RAM: 26GB
  GPU: nVidia Tesla K80 (Compute capability: 3.7)
  Language: Python 3.6
  Teams: A, C, D, G, H, K, L
- **Windows PC:**
  CPU: Intel(R) Corei7-3610QM 2.30GHz RAM: 8GB
  Language: Matlab
  OS: Windows 7 v. 2018a - Teams: B, I and J
  Windows 7 v. 2017a - Team: E
  Windows 7 v. 2013a - Team: F
  Windows 7 v. 2013b - Team: M
- **Linux PC:**
  OS: Linux Mint 19.1
  CPU: Intel(R) Corei5-4200U 1.60GHz RAM: 12GB
  Language: Java 8: DIB Java Algorithms
  Matlab 2013b: Ergina-Local,
  Ergina-Global, Howe, DIB Matlab Alg.

The time figures that are presented in the results section were “normalized” to allow a fair comparison of the order of magnitude of the processing times. “Normalization” was performed by comparing the execution-time performance of several binarization algorithms in more than one of the three SW/HW platforms above, using the three test data sets and analyzing the elapsed time. The reference SW/HW platform was Intel(R) Corei7-3610QM 2.30GHz RAM: 8GB, running Matlab, on Windows 7 Ultimate.

VI. Results

The results obtained for the quality performance of all 47 binarization algorithms assessed in this competition are presented grouped per test set. The ranking is made in terms of the total score of the algorithms in the six image quality measures listed for the scanned documents and the black-to-white proportion of pixels for the photographed documents. The normalized time that appears next to the names of the algorithms stand for the “normalized” time in seconds for binarizing the batch of the documents in each of the test sets reported, and provides the comparative order of magnitude of the processing times.

A. *Nabuco Dataset*

Table 2 presents the overall performance for the top 10 algorithms on Nabuco Dataset for Handwritten images. The documents were clustered according to the paper texture, as in Figure 1(left). In each column of Table 2, the time figure appears to the right-hand side of the algorithm.

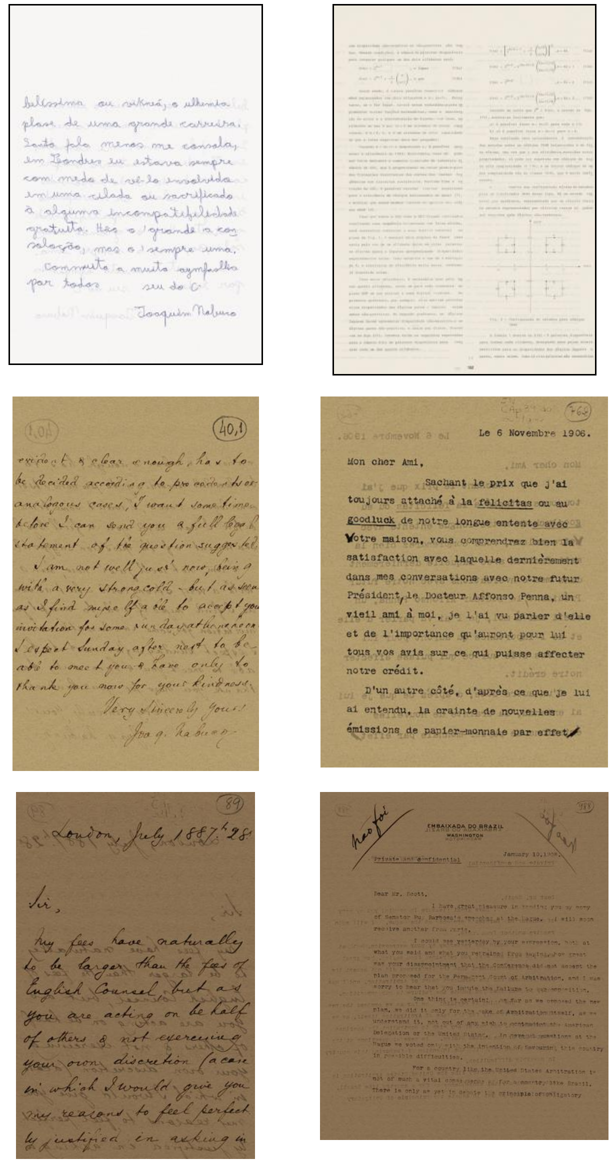

Figure 2: Samples of the synthetic images from the DIB test set, clustered by the printing method (handwritten or machine typed), color tone of the paper (light, medium, dark), and strength of the back-to-front interference (α=0.6, 0.7, 0.8).

Table 2: Nabuco Handwritten Top 10 Quality Algorithms

#	Dark		Medium		Light	
1	Nick	0.36	F	76.65	K	276.52
2	I	87.68	Otsu	0.01	I	96.48
3	J	22.96	IsoData	0.01	D	6.04
4	E(3)	3.36	I	80.75	Sauvola	0.36
5	F	46.69	E(3)	4.47	J	26.70
6	K	234.03	Nick	0.34	Nick	0.38
7	Sauvola	0.34	J	743.37	dS-L-R	0.01
8	D	6.04	dS-L-R	0.01	M-L	0.01
9	Li-Tam	0.01	E(1)	13.90	E(3)	3.13
0	Intermodes	0.01	Moment	0.01	E(1)	3.85

The performance of the top 10 algorithms in binarization quality for the machine typed images in Nabuco dataset are presented in Table 3, together with the times.

Table 3: Nabuco Typed Top 10 Quality Algorithms

#	Dark		Medium		Light	
1	Wolf	0.59	Nick	0.93	M	1.88
2	Wu-Lu	0.02	J	58.95	J	56.13
3	Shanbag	0.02	I	231.09	I	207.10
4	E(1)	4.02	Li-Tam	0.04	Li-Tan	0.03
5	dS-L-R	0.02	M	2.21	E(1)	3.13
6	Sauvola	0.51	Intermodes	0.04	Nick	0.77
7	Minimum	0.02	E(3)	3.24	D	11.79
8	C	8.06	E(1)	3.23	F	48.72
9	Nick	0.57	F	42.95	Howe	46.41
0	J	34.82	IsoData-O	0.05	C	11.55

B. *Synthetic Documents Dataset*

Table 4: Synthetic Top 10 Quality Algorithms with light paper tone and back-to-front interference (α) variation.

#	Hand W (α=0.8)		Machine T (α=0.8)		Printed (α=0.8)	
1	Minimum	0.23	Minimum	0.23	Minimum	0.13
2	Johann	0.24	Johann	0.23	E3	3.75
3	E3	3.99	IsoData	0.24	Intermodes	0.13
4	Otsu	0.23	Otsu	0.23	Johann	0.13
5	Huang	0.24	E2	4.02	Bradley	0.71
6	IsoData	0.23	J	144.29	Otsu	0.13
7	Bradley	1.02	E3	4.07	F	47.98
8	F	50.91	Bradley	1.04	IsoData	0.13
9	Moments	0.23	Huang	0.23	J	110.17
0	Intermodes	0.23	F	51.59	Li-Tam	0.13
#	Hand W (α=0.7)		Machine T (α=0.7)		Printed (α=0.7)	
1	E3	4.08	IsoData	0.23	Otsu	0.13
2	IsoData	0.23	Otsu	0.23	E3	3.80
3	Otsu	0.23	J	146.06	Intermodes	0.13
4	F	50.61	E3	3.90	IsoData	0.13
5	J	142.20	F	52.70	F	48.26
6	Minimum	0.23	Minimum	0.23	Minimum	0.13
7	I	609.90	I	578.56	J	103.36
8	Intermodes	0.23	M	3.91	Li-Tam	0.13
9	M	3.53	Bradley	1.02	I	423.00
0	Li-Tam	0.23	Li-Tam	0.23	Nick	1.51
#	Hand W (α=0.6)		Machine T (α=0.6)		Printed (α=0.6)	
1	F	50.88	E3	4.25	Intermodes	0.13
2	Intermodes	0.23	M	4.15	Li-Tam	0.13
3	Li-Tam	0.23	Intermodes	0.23	I	433.39
4	M	3.54	J	175.41	F	48.57
5	Sauvola	2.46	Li-Tam	0.24	J	121.31
6	J	160.29	Nick	2.19	Otsu	0.13
7	IsoData	0.24	I	602.57	IsoData	0.13
8	Otsu	0.23	IsoData	0.23	E3	3.77
9	I	616.48	Moments	0.23	M	2.76
0	Minimum	0.23	Bradley	1.05	Minimum	0.13

C. *Mobile Dataset*

The results obtained for the binarization of the images for the photographed documents images are presented grouped according to the mobile cell-phone model and the embedded strobe-flash usages are presented in Table 5.

It is clear, both observing the images in Figure 3 and the performance figures in Table 5, that the device feature and the strobe-flash "on" or "off" does affect the quality of the document images and their binarization results. The current device embedded software in the market, such as CamScan and EverNote, only acquire and crop the document images, but do not perform binarization.

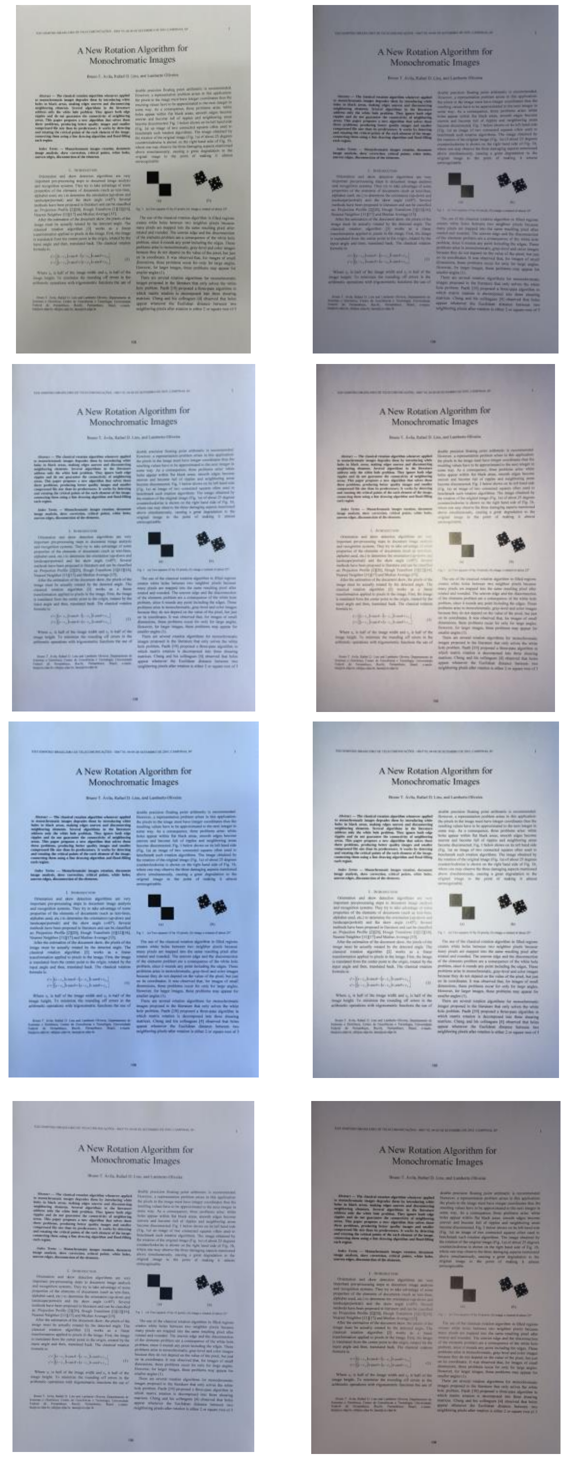

Figure 3: Samples of the images from Camera test set, clustered by the device (Moto Z2, Iphone 6, Iphone SE, Galaxy N4) and set-up of the strobe flash (left-column "off", right-column "on")

Table 5: Top 10 quality algorithms for photographed document images acquired with Iphone6, Iphone SE, Motorola Z1 and Galaxy Note 4 portable cell phones with strobe flash on and off.

#	Iphone 6				Iphone SE			
	OFF		ON		OFF		ON	
1	Ergina-L	1.25	C	21.48	E(2)	4.76	C	23.41
2	C	21.61	E(2)	4.71	C	23.91	E(2)	4.87
3	Bernsen	0.39	Bernsen	0.39	Ergina-L	1.63	B(1)	109.97
4	E(2)	4.86	H	66.1	Bernsen	0.46	B(3)	69.57
5	B(1)	106.43	B(2)	101.77	E(3)	3.92	B(2)	110.32
6	B(3)	67.25	B(1)	101.12	Bradley	0.76	Bernsen	0.41
7	H	66.15	D	21.59	Howe	100.97	H	69.82
8	Ergina-G	1.14	Ergina-L	1.55	B(1)	112.57	D	22.73
9	Huang	0.15	E3	4.06	D	23.16	Ergina-L	1.57
0	Otsu	0.14	Howe	97.68	E1	4.76	Bradley	0.75
#	**Motorola Z1**				**Galaxy Note 4**			
	OFF		ON		OFF		ON	
1	D	31.54	E(2)	5.14	D	32.61	Howe	155.61
2	Ergina-L	1.99	D	29.61	Ergina-G	1.58	E(2)	5.32
3	E(2)	5.26	C	27.62	E(2)	5.45	D	33.08
4	C	30.74	B(1)	136.77	C	32.60	C	32.21
5	Howe	145.89	B(2)	137.30	Howe	155.14	H	0.01
6	B(1)	145.84	Howe	135.31	B(2)	153.39	B(2)	154.32
7	B(2)	146.10	Ergina-G	1.35	B(1)	154.30	B(1)	153.72
8	Bernsen	0.52	H	88.38	Bernsen	0.55	Kapur	0.26
9	Bradley	1.01	Ergina-L	1.82	**Ergina-L**	2.12	Reny	0.25
0	E(3)	4.21	Bradley	0.92	**Triangle**	0.26	Yean	0.25

VII. Conclusions

Document image binarization is an important step in many document processing, indexing, and information extraction systems. This ICDAR 2019 Time-quality binarization competition assessed the quality of binary document images produced by forty-seven algorithms, seventeen new and thirty "classical" ones. Their performance was tested with three different image sets with varying image content, paper tone, back-to-front interference, and image acquisition configuration.

The quality assessment used several widely accepted image quality measures for the scanned images. The mean processing time for the ten best quality algorithms was taken allowing one to make a, as fair as possible, comparison of their time complexity. For photographed documents, the assessment made measured the proportion of the number of black-to-white pixels in the binary version of the scanned or digitally generated document and the photographed one. Both scanned and photographed binary images for all the algorithms were carefully visually inspected to see the coherence of the results obtained.

The algorithms from competitors H and L generated as output a file of the same format as their input. As the standard output of portable digital cameras is the jpeg file format, such algorithms could not be assessed for camera documents as the output was in grayscale (not monochromatic) as the jpeg noise is automatically added.

Several conclusions may be drawn from the results presented in this binarization competition:

- The claim that no binarization algorithm is good for all document images has been reinforced here.
- Most of the new algorithms presented at this competition are based on some machine-learning or neural-network strategies and yielded good-quality images, at a high cost of processing-time. It is important to remark that the learning/training time of such algorithms was not considered here.
- Surprisingly, several of the "classical" algorithms provide very good quality images, sometimes even the best quality results, and their time performance are at least two orders of magnitude faster than the machine-learning based algorithms.
- The newer algorithms, submitted to this competition, performed very well in the very complex images acquired with portable mobile cell phones, although the chances of embedding such algorithms in mobile application are slim in the short term, due to the architectural limitations of such portable devices.
- It seems that strategies that try to identify the features of the document and chooses the most suitable "classical" and fast algorithm to perform binarization, as presented in some of the new algorithms, will be a trend in this research line for the coming years.

A future assessment for the printed documents would be provided by checking the quality of their OCR transcription.

All the test images and the result of their binarization using the forty seven algorithms assessed here will be made available at the DIB platform (https://dib.cin.ufpe.br) immediately after ICDAR 2019.

The competitors will be invited to make their executable code available at the code repository at the DIB platform.

Acknowledgments

The authors grateful to all the participants of this ICDAR competition, and for all researchers who made the code for their binarization algorithms publically available. They also acknowledge the financial support both of CNPq and FACEPE (Brazil).

References

[1] T. Roughgarden, "*Beyond Worst-Case Analysis*," CACM, pp. 1–14, March 2019.

[2] R. D. Lins, M. M. de Almeida, R. B. Bernardino, D. Jesus, and J. M. Oliveira, "*Assessing Binarization Techniques for Document Images*," ACM Symposium on Document Engineering, 2017, pp. 183–192.

[3] R. D. Lins, R. B. Bernardino, D. M. de Jesus, and J. M. Oliveira, "*Binarizing Document Images Acquired with Portable Cameras*," in CBDAR 2017, 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, pp. 45–50.

[4] R. D. Lins, "*A Taxonomy for Noise in Images of Paper Documents - The Physical Noises*." Image Analysis and Recognition. ICIAR 2009. Lecture Notes in Computer Science, vol 5627. Springer, 2009.

[5] Y. Akbari, S. Al-Maadeed, K. Adam, "*Binarization of degraded document images using convolutional neural networks and wavelet-based multichannel images*", Submitted to Pattern Recognition Journal.

[6] Y. Akbari, A. S. Britto~Jr., S. Al-Maadeed, L. S. Oliveira, "*Binarization of Degraded Document Images using Convolutional Neural Networks based on predicted Two-Channel Images*", Submitted to ICDAR 2019.

[7] J.Calvo-Zaragoza, A.-J. Gallego, *A selectional auto-encoder approach for document image binarization*. Patt. Recog. 86: 37-47, 2019.

[8] L. Zhou, C. Zhang, and M. Wu, *D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction*, in IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018, pp. 192-1924.

[9] K. Saddami, K. Munadi, S.Muchallil, F. Arnia. *Improved Thresholding Method for Enhancing Jawi Binarization Performance*. In ICDAR´17, 2017, Vol. 1, pp. 1108-1113.

[10] K. Saddami, P. Afrah, V. Mutiawani, F. Arnia, F. *A New Adaptive Threshoding Technique for Binarizing Ancient Document*. In International Conference of INAPR, 2018 (Vol. 1, pp. 57-61)

[11] K. Saddami, K. Munadi, Y. Away, F. Arnia. *Combination Local and Global Thresholding Method for Binarizing Ancient Jawi Document*. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK). 2019.

[12] A. Gattal, F. Abbas, M. R. Laouar, *Automatic Parameter Tuning of K-Means Algorithm for Document Binarization*, 7th International Conference on Software Engineering and New Technologies, 2018.

[13] J. Long, E. Shelhamer, and T. Darrell, "*Fully convolutional networks for semantic segmentation IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[14] X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, and J. Liang, "*East: an efficient and accurate scene text detector*," IEEE Computer Vision and Pattern Recognition, 2017, pp. 5551–5560.

[15] B. Epshtein, E. Ofek, Y. Wexler, "*Detecting text in natural scenes with stroke width transform*," IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 2963-2970.…

[16] X. Kong, G. Sun., Q. Wu, *et al*. "*Hybrid pyramid U-Net model for brain tumor segmentation*," 10th IFIP TC-12 International Conference on Intelligent Information Processing, China, 2018, pp. 346-355.

[17] S. Bhowmik, R. Sarkar, B. Das, and D. Doermann, "*GiB: a Game theory Inspired Binarization technique for degraded document images*," IEEE Trans. Image Process., vol. 28(3): 1443–1455, 2019.

[18] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "*A combined approach for the binarization of handwritten document images*," Pattern Recognit. Lett., vol. 35, pp. 3–15, 2014.

[19] S. He and L. Schomaker, "DeepOtsu: Document enhancement and binarization using iterative deep learning," Pattern Recognit., vol. 91, pp. 379–390, 2019.

[20] J. Bernsen, "*Dynamic thresholding of gray-level images*," in Int. Conference on Pattern Recognition, 1986, pp. 1251–1255.

[21] D. Bradley and G. Roth, "*Adaptive Thresholding using the Integral Image*," J. Graph. Tools, vol. 12, no. 2, pp. 13–21, Jan. 2007.

[22] J. M. M. da Silva, R. D. Lins, and V. C. da Rocha, "*Binarizing and filtering historical documents with back-to-front interference*," ACM symposium on Applied computing - SAC '06, 2006, pp. 853–858.

[23] N. R. Howe, "*Document binarization with automatic parameter tuning*," Int. J. Doc. Anal. Recognit., vol. 16, no. 3, pp. 247–258, Sep. 2013.

[24] L. K. Huang and M. J. J. Wang, "*Image thresholding by minimizing the measures of fuzziness*," Pattern Recognit., vol. 28(1):41–51, 1995.

[25] E. Kavallieratou, "*A binarization algorithm specialized on document images and photos*," Proc. Int. Conf. Doc. Anal. Recognition, ICDAR´05, pp. 463–467, 2005.

[26] E. Kavallieratou and S. Stathis, "*Adaptive binarization of historical document images*," Proc. - Int. Conf. Pattern Recognit., vol. 3, pp. 742–745, 2006.

[27] J. M. S. Prewitt and M. L. Mendelsohn, "*The analysis of cell images*," Ann. N. Y. Acad. Sci., vol. 128, no. 3, pp. 1035–1053, Dec. 2006.

[28] F. R. Velasco, "*Thresholding Using the Isodata Clustering Algorithm*," Mar. 1979.

[29] J. Johannsen, G and Bille, "*A threshold selection method using information measures*," *Int. Conf. Pattern Recog.*, 1982, pp. 140–143.

[30] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "*A new method for gray-level picture thresholding using the entropy of the histogram*," Comput. Vision, Graph. Image Process., vol. 29(1), p. 140, Jan. 1985.

[31] C. H. Li and P. K. S. Tam, "*An iterative algorithm for minimum cross entropy thresholding*," Pattern Recognit. Lett., vol. 19, no. 8, pp. 771–776, 1998.

[32] C. A. B. Mello and R. D. Lins, "*Image segmentation of historical documents*," Visual 2000, 2000.

[33] W.-H. Tsai, "*Moment-preserving thresolding: A new approach*," Comput. Vision, Graph. Image Process., vol. 29(3):377–393, 1985.

[34] N. Otsu, "*A threshold selection method from gray-level histograms*," IEEE Trans. Syst. Man. Cybern., vol. 9, no. 1, pp. 62–66, 1979.

[35] P. Sahoo, C. Wilkins, and J. Yeager, "*Threshold selection using Renyi's entropy*," Pattern Recognit., vol. 30, no. 1, pp. 71–84, 1997.

[36] J. Sauvola and M. Pietikäinen, "*Adaptive document image binarization*," Pattern Recognit., vol. 33, no. 2, pp. 225–236, 2000.

[37] A. G. G. Shanbhag, "*Utilization of Information Measure as a Means of Image Thresholding*," *CVGIP Graph. Model. Image Process.*, vol. 56, no. 5, pp. 414–419, 1994.

[38] C. Wolf, J.-M. Jolion, and F. Chassaing, "*Text localization, enhancement and binarization in multimedia documents*," in Object recognition supported by user interaction for service robots, 2003, vol. 2, pp. 1037–1040.

[39] W. Lu, M. Songde, and H. Lu, "*An effective entropic thresholding for ultrasonic images*," Pattern Recognition, 1998. Proceedings. 14th Int. Conf., pp. 1552–1554, vol. 2, 1998.

[40] R. D. Lins, "*Two Decades of Document Processing in Latin America*," Journal of Universal Computer Science, vol. 17, no. 1, pp. 151–161, 2011.

[41] R. D. Lins, R. B. Bernardino, D. M. de Jesus, and J. M. Oliveira, "*Binarizing Document Images Acquired with Portable Cameras*," in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, pp. 45–50.

[42] R. B. Bernardino, R. D. Lins, D. M. de Jesus, "*A Quality and Time Assessement of Binarization Algorithms*", 15th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2019.

[43] G. W . Zack, W . E. Rogers, and S. A. Latt, "*Automatic measurement of sister chromatid exchange frequency*". J. Histochem Cytochem., vol. 25, no. 7, pp. 741–753, 1977.

[44] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "*Performance Evaluation Methodology for Historical Document Image Binarization,*" IEEE Trans. Im. Process., vol. 22(2): 595–609, 2013.

[45] R. G. Congalton, "*A review of assessing the accuracy of classifications of remotely sensed data*", Remote Sens. Environ., vol. 37, no. 1, pp. 35–46, Jul. 1991.