

# An Objective Evaluation Methodology for Document Image Binarization Techniques

*K. Ntirogiannis, B. Gatos and I. Pratikakis*

Computational Intelligence Laboratory,  
Institute of Informatics and Telecommunications,  
National Center for Scientific Research “Demokritos”,  
GR-153 10 Agia Paraskevi, Athens, Greece  
<http://www.iit.demokritos.gr/cil>,  
{[knir](mailto:knir@iit.demokritos.gr),[bgat](mailto:bgat@iit.demokritos.gr),[ipratika](mailto:ipratika@iit.demokritos.gr)}@iit.demokritos.gr

## Abstract

*Evaluation of document image binarization techniques is a tedious task that is mainly performed by a human expert or by involving an OCR engine. This paper presents an objective evaluation methodology for document image binarization techniques that aims to reduce the human involvement in the ground truth construction and consecutive testing. A skeletonized ground truth image is produced by the user following a semi-automatic procedure. The estimated ground truth image can aid in evaluating the binarization result in terms of recall and precision as well as to further analyze the result by calculating broken and missing text, deformations and false alarms. A detailed description of the methodology along with a benchmarking of the six (6) most promising state-of-the-art binarization algorithms based on the proposed methodology is presented.*

## 1. Introduction

Document image binarization is an important step in the document image analysis and recognition pipeline. The performance of a binarization technique directly affects the recognition analysis. Therefore, it is imperative to have an objective evaluation which will account for the performance of the binarization.

Several efforts have been presented that strive towards evaluating the performance of document image binarization techniques. These efforts can be divided in four categories. In the first category, evaluation is performed by one or many human

evaluators [1], [2], [3], [7], [16] while in the second category, the binary result is subjected to OCR and the corresponding result is evaluated with respect to character or word accuracy [5], [6], [17], [18]. The third category uses a binary image as ground truth image and evaluation is performed by comparing the two images (ground truth and binarization result image) at pixel level [4], [19], [20], [21], [22]. In this category, either noise is added to the ground truth image so as to produce the grey scale image (synthetic images), or the original grey scale image is optimally binarized by a human, after many binarization tests and supplementary cleaning, in order to produce a satisfying ground truth image. The last category uses a combination of human-oriented evaluation and OCR results accuracy [7], [23].

Evaluation performed by a human expert is not only subjective but also time consuming. Furthermore, it lacks robustness since it has been observed that in fuzzy situations, the same observer may make different selections for the same dataset in different sessions. In [2], the amount of symbols, text, lines, etc., that are broken or blurred, the loss of complete objects and the noise in both background and print are used as evaluation criteria for a human expert. Moreover, in [3], two widely used evaluation metrics of Precision and Recall, are defined by the manual count of fully restored words. The use of OCR as a means for evaluation is widely used in the evaluation of modern machine-printed documents only, supported by contemporary OCR engines (e.g. ABBYY FineReader) since handwritten OCR yields no satisfactory results [1]. Synthetic images are widely used in binarization evaluation process, however they have to be carefully constructed and they are usually of small amount.

In this paper, an objective evaluation methodology for document image binarization techniques is presented, that avoids the dependence of an OCR engine and reduces the human interference as it is based upon a semi-automated ground truth construction. Not only the evaluation metrics of precision and recall, but also broken and missing text, false alarms and deformations are calculated during the proposed evaluation procedure.

Figure 1 shows all stages of the proposed methodology. Each stage is analyzed in the following sections. Specifically the construction of the ground truth image is described in Section 2. The evaluation metrics of recall and precision are detailed in Sections 3 and 4 respectively. In Section 5 the evaluation results of representative state-of-the-art binarization techniques are presented and in Section 6 the conclusions are discussed.

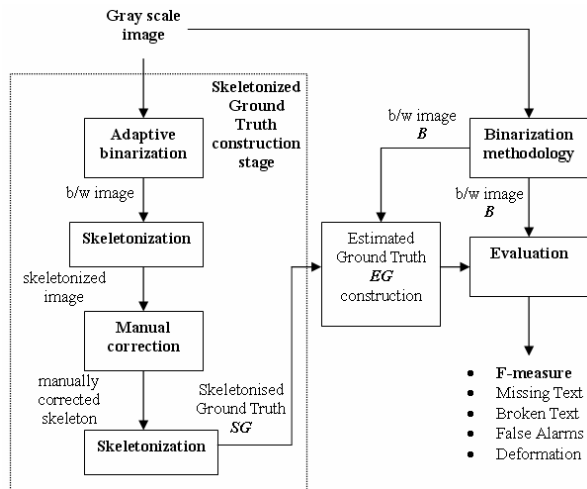


Figure 1. The proposed methodology stages

## 2. Construction of Ground Truth Image Stage

In the proposed methodology, the construction of ground truth plays an important role, since it aids towards the automation of the evaluation procedure. It consists of two distinct stages, namely *Skeletonized Ground Truth* (SG) stage and *Estimated Ground Truth* (EG) stage. These stages will be described in detail, in the following sections.

### 2.1. Skeletonized Ground Truth Image Stage

The idea of building a skeletonized ground truth image originates from a user’s natural interaction, in which a character is presented by its silhouette. For this task, a user would directly draw a one pixel wide

curve approximately in the medial axis of the character. Our proposed skeletonized ground truth image construction stage strives toward automating the aforementioned procedure. To accomplish this, we follow the consecutive stages described in the following.

The grey scale image (Figure 2(a)) is binarized using an adaptive binarization technique [24] (Figure 2(b)). Then, a skeletonization method [8] is used and the resulting skeleton image has one pixel wide text (Figure 3(a)). Due to artifacts in the character, skeletonization does not always represent the complete character. In this case, the user is required to delineate the remaining character or remove spurious parts. To further aid the user during the correction stage, we show both the layers of the skeleton and the grey level image (Figure 3(b)) to guide him/her in the correction process. Finally, a second skeletonization pass, guarantees that ground truth text is one pixel wide (Figure 3(c)).



Figure 2. (a) Original image and (b) the corresponding binary image after adaptive binarization

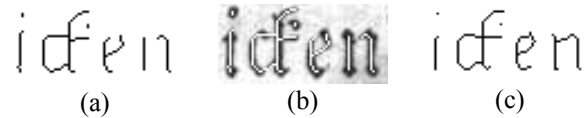


Figure 3. (a) Binary image after skeletonization (skeleton image); (b) simultaneous viewing of skeleton and grey level image layer; (c) skeletonized ground truth image which text is one pixel wide

The skeletonized ground truth image and the corresponding connected component labeled image are defined by the following equations:

$$SG(x, y) = \begin{cases} 0, & \text{background} \\ 1, & \text{text} \end{cases} \quad (1)$$

$$SGC(x, y) = \begin{cases} 0, & \text{if } SG(x, y) = 0 \\ i, & \text{otherwise} \end{cases} \quad (2)$$

where  $i \in K$ ,  $K = \{1, 2, \dots, M\}$  and  $M$  denotes the maximum number of the connected components found in the ground truth image.

After the end of the skeletonized ground truth construction stage, we are able to automatically measure the performance of any binarization algorithm in terms of recall (see Section 3).

## 2.2. Estimated Ground Truth Image Stage

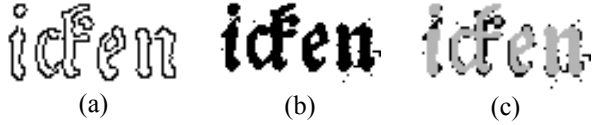
To complete the evaluation process we should calculate the performance of binarization algorithms in terms of precision (see Section 4). Precision requires considering ground truth characters as much close as the original ones. In this paper, we present a methodology to automatically estimate the ground truth for the computation of precision taking into account that a skeletonized ground truth image has been achieved.

Given the skeletonized ground truth image, we apply a dilation constrained by the edge image  $CE$  (Figure 4(a)) and the binary image  $B$  under evaluation (Figure 4(b)) where

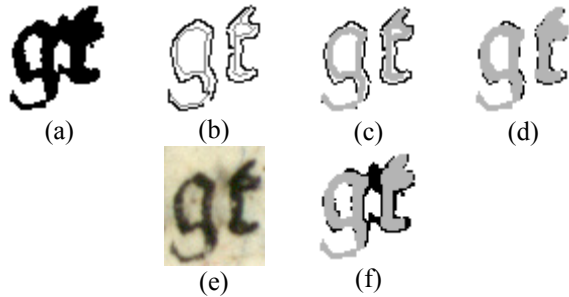
$$CE(x, y) = \begin{cases} 1, & \text{if } (x, y) \in \partial I(x, y) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$B(x, y) = \begin{cases} 0, & \text{background} \\ 1, & \text{text} \end{cases} \quad (4)$$

where  $\partial I(x, y)$  denotes the resulting image after Canny Edge detection [9] on the original grey scale image  $I(x, y)$ .



**Figure 4.** (a) Edges of original image; (b) binary image under evaluation and (c) estimated ground truth image (in grey)



**Figure 5.** (a) A connected component of the binary image  $B$ ; (b) the corresponding edges (in black) along with the skeletonized ground truth  $SG$  components (in grey); (c) the dilated  $SG$  components after one dilation (in grey), 15.29% of the edges covered; (d) the dilated  $SG$  components after two dilations (in grey), 54.9% of the edges covered, dilation stops; (e) the original image and (f) the estimated ground truth image (in grey) placed over the binary image (in black)

Specifically, as far as a connected component ( $cc_i$ ) of the binary image  $B$  is concerned (Figure 5(a)), the

corresponding skeletonized ground truth component or components (Figure 5(b)) start to dilate. As soon as half of the edges that belong to  $cc_i$  are covered by the dilated skeletonized ground truth components (Figures 5(c) and 5(d)), dilation stops. The dilated components represent the estimated ground truth components (Figure 5(f)). The construction of the estimated ground truth image requires the aforementioned procedure to be applied in every connected component of the binary image  $B$ . It is worth mentioning that dilated components cannot exceed  $cc_i$ , denoting that a dilated component cannot be larger than the original one.

For the sake of clarity, we provide the following definitions, along with the algorithm for the estimated ground truth image stage in pseudo code.

Let  $A$  be a binary image, and  $A^r$  denotes the corresponding dilated image after  $r$  iterations. Let  $BC$  denotes the connected component labeled image of  $B$ , defined as:

$$BC(x, y) = \begin{cases} 0, & \text{if } B(x, y) = 0 \\ j, & \text{otherwise} \end{cases} \quad (5)$$

where  $j \in L$ ,  $L = \{1, 2, \dots, N\}$  and  $N$  denotes the maximum number of the connected components found in the binarized image.

### Algorithm description (pseudo code)

1.  $A^0(x, y) = SG(x, y)$ ,
2.  $\forall (x, y) \in B(x, y): BC(x, y) = i$   
 $AND \sum_{BC(x, y) = i} B(x, y) \cdot SG(x, y) > 0$
3. Stop = false
4. while (NOT(Stop))
5.  $A^{r+1}(x, y) = (A^r(x, y) \oplus B(x, y)) \cap B(x, y)$   
 $\sum_{BC(x, y) = i} A^{r+1}(x, y) \cdot CE(x, y)$
6. if  $\frac{\sum_{BC(x, y) = i} A^{r+1}(x, y) \cdot CE(x, y)}{\sum_{BC(x, y) = i} CE(x, y)} > \frac{1}{2}$   
 $OR A^{r+1}(x, y) = A^r(x, y)$
7. Stop = true
8. End if
9. End while

where  $i \in L$  and  $\oplus$  denotes dilation.

The aforementioned procedure described by the pseudo code is applied for each connected component of the binary image  $B$  and the estimated ground truth image  $EG$  is fully constructed, as defined in Eq. 6.

Example estimated ground truth images are shown in Figures 4(c) and 5(f).

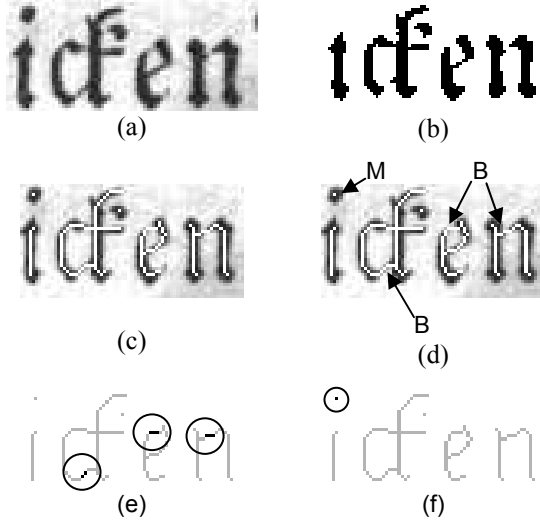
$$EG(x, y) = \begin{cases} 0, & \text{background} \\ 1, & \text{text} \end{cases} \quad (6)$$

### 3. Recall

Recall is defined as the percentage of the skeletonized ground truth image  $SG$  that is detected in the resulting binary image  $B$ . Recall is given by the following equation:

$$\text{Recall} = \frac{\sum_{x=1, y=1}^{x=lx, y=ly} SG(x, y) \cdot B(x, y)}{\sum_{x=1, y=1}^{x=lx, y=ly} SG(x, y)} 100\% \quad (7)$$

These parts that belong to  $SG$  and are not detected can be classified as broken or missing text. Broken text is related with  $SG$  components that are partially detected while missing text denotes all  $SG$  components that have not been detected at all (Figures 6(d), 6(e), and 6(f)). In the following sections, we further analyze the recall result by calculating broken and missing texts information.



**Figure 6.** (a) Original grey scale image; (b) resulting binary image  $B$ ; (c) skeletonized ground truth  $SG$  image overlay; (d) broken [B] and missing [M] text parts; (e) Broken parts (in black) indicated in circles and (f) missing part indicated in circle

#### 3.1. Broken Text

Let  $f()$  be the function that defines whether part of a skeletonized ground truth component is partially detected in the binary image  $B$ .  $f()$  is calculated as follows:

$$f(i) = \begin{cases} 1, & \text{if } \sum_{x=1, y=1}^{x=lx, y=ly} SG(x, y) \cdot B(x, y) > 0 \\ SG(x, y) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where  $i \in K$ .

Broken text (BT) can be defined by the percentage of the skeletonized ground truth image  $SG$  parts which are not detected in the resulting binary image  $B$  while belonging to components that are partially detected (Figures 6(d) and 6(e)). Broken text is given by the following equation:

$$BT = \frac{\sum_{x=1, y=1}^{x=lx, y=ly} f(SGC(x, y)) \cdot (1 - B(x, y))}{\sum_{x=1, y=1}^{x=lx, y=ly} SG(x, y)} 100\% \quad (9)$$

#### 3.2. Missing Text

Missing text (MT) is defined by the percentage of the skeletonized ground truth image  $SG$  parts which are not detected in the resulting binary image  $B$  while belonging to components that are not detected at all (Figures 6(d) and 6(f)). Missing text is given by the following equation:

$$MT = \frac{\sum_{x=1, y=1}^{x=lx, y=ly} (1 - f(SGC(x, y))) \cdot SG(x, y)}{\sum_{x=1, y=1}^{x=lx, y=ly} SG(x, y)} 100\% \quad (10)$$

### 4. Precision

Considering the binary image  $B$ , precision estimates the foreground areas that are actually text. In our method, the actual text is considered as the estimated ground truth image  $EG$  as described in Section 2.

Precision is defined as the percentage of the estimated ground truth image that is detected in the binary image (Figure 7(c)).

$$\text{Precision} = \frac{\sum_{x=1, y=1}^{x=lx, y=ly} EG(x, y) \cdot B(x, y)}{\sum_{x=1, y=1}^{x=lx, y=ly} B(x, y)} 100\% \quad (11)$$

The foreground areas of the binary image that are not detected during precision estimation are considered

as either false alarms or deformations which are described in the following section.

#### 4.1. False Alarms

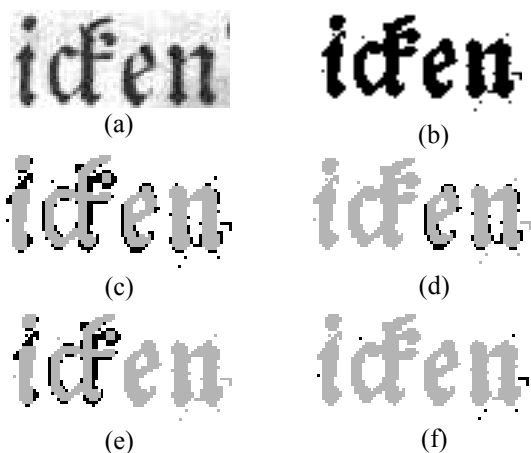
False alarms refer to foreground pixels of the binary image  $B$  that do not belong to estimated ground truth image (Figure 7(f)). They are defined by the percentage of all pixels of the components of the binary image  $B$  that do not have any corresponding pixel with the skeletonized ground truth image  $SG$ .

$$FAlarms = \frac{\sum_{x=1, y=1}^{x=Lx, y=Ly} h(BC(x, y)) \cdot B(x, y)}{\sum_{x=1, y=1}^{x=Lx, y=Ly} B(x, y)} 100\% \quad (12)$$

where  $h(i)$  is a function denoting whether a binary component is not detected in the skeletonized ground truth image.

$$h(i) = \begin{cases} 1, & \text{if } \sum_{\substack{x=1, y=1 \\ BC(x, y)=i}}^{x=Lx, y=Ly} SG(x, y) \cdot B(x, y) = 0 \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where  $i \in L$



**Figure 7.** (a) Original grey scale image; (b) resulting binary image  $B$ ; (c) Estimated ground truth image (in grey) and false alarms along with deformations (in black); (d) Deformations that do not result in component merging (in black); (e) Deformations that do result in component merging (in black) and (f) false alarms (in black).

#### 4.2. Deformations

Components often merge with adjacent background information that was detected as text during binarization. Deformations do not only enlarge

(deform) components but they are also responsible for merging adjacent components (Figures 7(d) and 7(e)).

In our method, deformations of the binary image are defined by the percentage of all text pixels of the binary image  $B$  that are not detected in the estimated ground truth image  $EG$  and do not belong to false alarms components as described in the previous section.

The deformation of the binary image is defined by Eq. 14 while deformation leading to merging is defined by Eq. 15.

$$Deform = \frac{\sum_{x=1, y=1}^{x=Lx, y=Ly} d(BC(x, y)) \cdot B(x, y) \cdot (1 - EG(x, y))}{\sum_{x=1, y=1}^{x=Lx, y=Ly} B(x, y)} 100\% \quad (14)$$

$$MergeDeform = \frac{\sum_{x=1, y=1}^{x=Lx, y=Ly} m(BC(x, y)) \cdot B(x, y) \cdot (1 - EG(x, y))}{\sum_{x=1, y=1}^{x=Lx, y=Ly} B(x, y)} 100\% \quad (15)$$

where  $d(i)$  and  $m(i)$  are functions denoting whether a binary component corresponds to one or more ground truth components respectively. Functions  $d(i)$  and  $m(i)$  are defined as follows:

$$d(i) = \begin{cases} 1, & \text{if } |SGC(x, y)| = 1 \\ & \forall(x, y): BC(x, y) = i \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

$$m(i) = \begin{cases} 1, & \text{if } |SGC(x, y)| > 1, \\ & \forall(x, y): BC(x, y) = i \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

where  $|SGC(x, y)|$  denotes the Cardinality of  $SGC(x, y)$ .

## 5. Experimental Results

The proposed objective evaluation methodology for document image binarization techniques was applied on a set of grey scale handwritten and machine-printed documents with low quality, shadows, non-uniform illumination, strains, presence of characters from the other side of the page and other significant artifacts. Among all documents, we selected the twenty (20) most representative (ten handwritten and ten machine-printed) and marked the skeletonized ground truth  $SG$  following the procedure described in Section 2. Six (6) of the most promising global and adaptive binarization techniques were chosen for evaluation:

1. Otsu's method (OTS) [10]
2. Bernsen's method (BER) [11]
3. Niblack's method (NIB) [12]
4. Sauvola's method (SAU) [13]

- 5. Adaptive Logical method (AL) [14]
- 6. Adaptive Degraded Document method (GPP)[15]

Our evaluation is based on F-measure which is defined as follows:

$$F = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

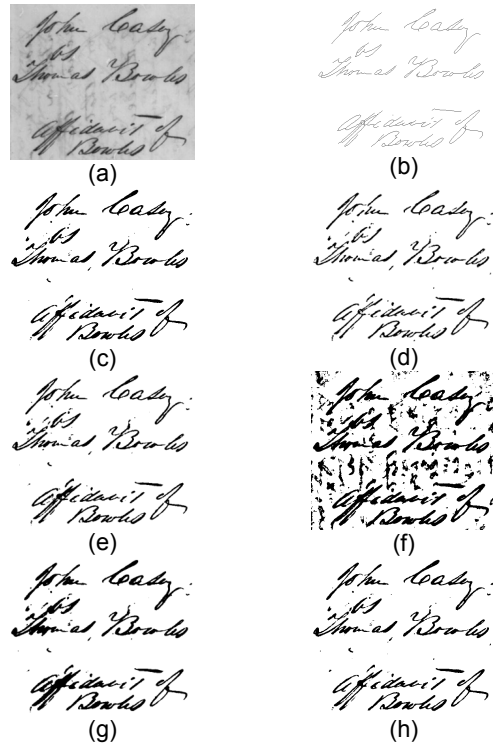
An example of the application of all methodologies to a grey scale handwritten document image is given in Figure 8 where the skeletonized ground truth image **SG** is also demonstrated. The corresponding evaluation results are shown in Table 1. Likewise, in Figure 9, an example of the application of all methodologies to a grey scale machine-printed document image is given and Table 2 demonstrates the corresponding evaluation results.

The evaluation results where the average values for all handwritten and machine-printed test images have been taken into account are presented in Tables 3 and 4 respectively. According to these results, Adaptive Degraded Document method (GPP) [15] had the best overall performance concerning both handwritten and machine-printed document images with F-measures of 85.23% and 93.19% respectively. As far as the handwritten images are concerned, GPP performed slightly better than the Sauvola’s method (SAU) [13] which F-measure is 84.61%, while in the case of machine-printed images, Otsu’s method (OTS) [10] was second with F-measure of 93.10%.

A more extensive analysis of the evaluation results, concerning all twenty (20) test images (Table 5), shows that Adaptive Level Method (AL) [14] has the best precision (88.44%) (Table 5) and the highest *broken text* value (Tables 3 and 4) that leads to the worst recall (87.55%) (Table 5). This behavior is expected, since AL is based on stroke width estimation. Moreover, in handwritten documents which have characters of nearly the same width, AL has indeed the best precision (88.83%) (Table 3), while in machine-printed images, where characters of different widths may exist, AL has the fourth best precision (88.04%) (Table 4). On the other hand, Niblack’s (NIB) [12] and Otsu’s (OTS) [10] methods, have the best and second best recall rates of 99.15% and 93.90% respectively. However, they have the lowest and second lowest precision rates of 47.13% and 81.01% respectively. In other words, these two methods preserve the characters information, but they cannot effectively remove the background noise, especially in cases where background and characters do not significantly differ. Bernstein’s method (BER) [11] is slightly better than OTS (Table 5), even if it has more *broken text* and *false alarms* (Tables 3 and 4).

The overall rank that is presented in Table 5 is verified by other works, [4] and [15]. In the first case

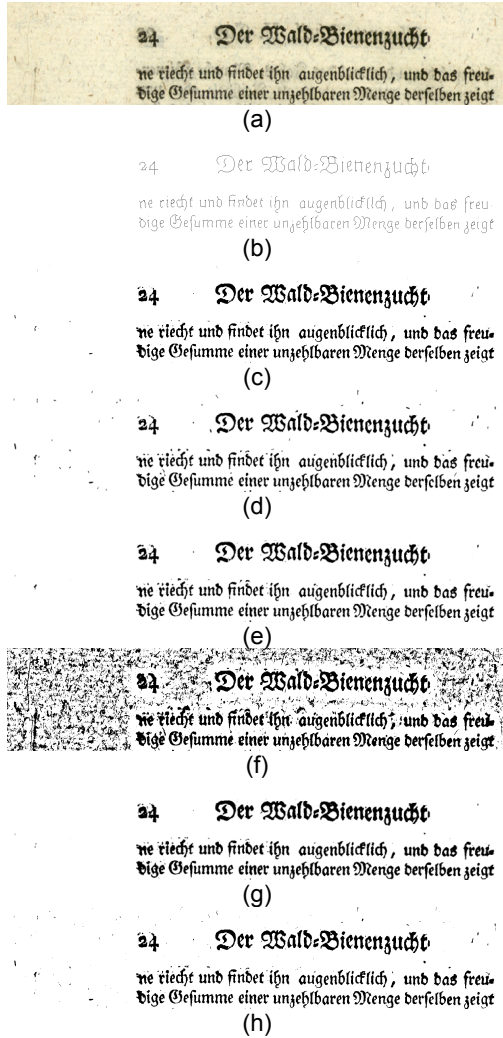
of [4], where a binary ground truth image and pixel accuracy were used as well, SAU, BER, OTS and NIB were evaluated with the same rank, concerning the document test images. In the second case of [15], where OCR word accuracy was used, GPP, SAU, OTS and NIB were also evaluated with the same rank according to the proposed methodology. AL method did not participate in any of the other works mentioned.



**Figure 8.** (a) Original handwritten image; (b) skeletonized ground truth image; (c) GPP binarization; (d) AL binarization; (e) BER binarization; (f) NIB binarization; (g) OTS binarization and (h) SAU binarization.

**Table 1.** All evaluation metrics for every binarization technique concerning the handwritten image of Figure 8.

Figure 8	GPP	AL	BER	NIB	OTS	SAU
<b>F-measure</b>	<b>93.25</b>	87.92	90.39	67.95	91.23	92.84
<b>Recall</b>	96.27	88.74	96.90	<b>99.56</b>	98.58	96.36
<b>Precision</b>	<b>90.42</b>	87.11	84.69	51.58	84.91	89.56
Broken Text	03.73	11.09	02.93	<b>00.44</b>	01.25	03.64
Missing Text	<b>00.00</b>	00.17	00.17	<b>00.00</b>	00.17	<b>00.00</b>
Merge-Deform	00.71	<b>00.14</b>	00.19	07.84	04.81	00.74
Deform	<b>06.81</b>	09.30	11.54	08.39	07.24	<b>06.81</b>
False Alarms	<b>02.06</b>	03.45	03.57	32.19	02.84	02.89



**Figure 9.** (a) Original machine-printed image; (b) skeletonized ground truth image; (c) GPP binarization; (d) AL binarization; (e) BER binarization; (f) NIB binarization; (g) OTS binarization and (h) SAU binarization.

**Table 2.** All evaluation metrics for every binarization technique concerning the machine-printed image of Figure 9.

Figure 9	GPP	AL	BER	NIB	OTS	SAU
F-measure	<b>93.42</b>	91.67	93.41	46.00	91.59	92.15
Recall	99.24	98.35	98.88	99.91	<b>99.93</b>	99.29
Precision	88.24	85.85	<b>88.52</b>	29.88	84.53	85.96
Broken Text	00.76	01.64	01.12	00.09	<b>00.07</b>	00.71
Missing Text	<b>00.00</b>	00.01	<b>00.00</b>	<b>00.00</b>	<b>00.00</b>	<b>00.00</b>
Merge-Deform	01.14	<b>00.41</b>	01.03	02.50	02.22	01.50
Deform	09.65	10.03	08.72	<b>04.84</b>	12.37	09.07
False Alarms	00.97	03.71	01.73	62.78	<b>00.88</b>	02.87

**Table 3.** The average value of all evaluation metrics for every binarization technique concerning all handwritten images.

Hand/ten	GPP	AL	BER	NIB	OTS	SAU
F-measure	<b>85.23</b>	81.99	77.70	50.87	77.02	84.61
Recall	85.36	79.82	87.83	<b>98.47</b>	89.40	86.51
Precision	87.98	<b>88.83</b>	75.86	35.42	73.59	85.89
Broken Text	14.04	19.53	11.90	<b>01.42</b>	10.30	12.89
Missing Text	00.60	00.65	00.27	<b>00.01</b>	00.30	00.60
Merge-Deform	01.05	<b>00.29</b>	13.42	12.93	15.36	00.61
Deform	10.02	08.94	<b>06.89</b>	09.14	08.96	10.22
False Alarms	<b>00.95</b>	01.94	03.83	42.51	02.09	03.28

**Table 4.** The average value of all evaluation metrics for every binarization technique concerning all machine-printed images.

Printed	GPP	AL	BER	NIB	OTS	SAU
F-measure	<b>93.19</b>	91.39	92.74	72.41	93.10	92.30
Recall	98.33	95.27	97.16	<b>99.82</b>	98.40	98.64
Precision	88.64	88.04	<b>88.95</b>	58.84	88.42	86.92
Broken Text	01.66	04.72	02.83	<b>00.18</b>	01.60	01.36
Missing Text	00.01	00.01	00.01	<b>00.00</b>	<b>00.00</b>	<b>00.00</b>
Merge-Deform	00.30	<b>00.09</b>	00.14	01.82	00.45	00.46
Deform	09.82	09.04	<b>08.16</b>	09.21	10.03	09.81
False Alarms	01.24	02.83	02.75	30.13	<b>01.10</b>	02.81

**Table 5.** The average value of F-measure, Recall and Precision, for every binarization technique concerning all test images.

All	GPP	AL	BER	NIB	OTS	SAU
F-measure	<b>89.21</b>	86.69	85.22	61.64	85.06	88.46
Recall	91.85	87.55	92.50	<b>99.15</b>	93.90	92.58
Precision	88.31	<b>88.44</b>	82.41	47.13	81.01	86.41

## 6. Conclusion

This work is based on a semi-automatic procedure for the construction of the ground truth as well as a fully automated evaluation scheme. The evaluation results of the proposed methodology, are extensively analyzed and verified by other works, [4] and [15], as presented in the previous section. However, it is likely to evaluate different binary images with different ground truth images, since the ground truth image is constructed in relation with the corresponding binary image  $B$  and the edge image  $CE$ . Even though, the difference of the ground truth images cannot seriously affect the evaluation results.

In future work, we will focus on the development of a more accurate technique for the estimation of



ground truth image  $EG$ , that will not depend on the binary image  $B$  under evaluation.

## Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n° 215064 (project IMPACT).

## References

- [1] E. Kavallieratou, and S. Stathis, "Adaptive Binarization of Historical Document Images", *18<sup>th</sup> International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, China, 2006, vol. 3, pp. 742-745.
- [2] D. Trier, and T. Taxt, "Evaluation of Binarization Methods for Document Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, March 1995, vol. 17, no. 3, pp. 312-315.
- [3] Q. Wang, and C. L. Tan, "Matching of Double Sided Document Images to Remove Interference", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, Kauai island of Hawaii, USA, 2001, vol. 1, p. 1084.
- [4] M. Sezgin, and B. Sankur, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation", *Journal of Electronic Imaging*, January 2004, vol. 13, no. 1, pp. 146-168.
- [5] J. He, Q. D. M. Do, A.C. Downton, and J. H. Kim, "A Comparison of Binarization Methods for Historical Archive Documents", *Proceedings of the 8th International Conference on Document Analysis and Recognition (ICDAR '05)*, Seoul, South Korea, August-September 2005, vol. 1, pp. 538-542.
- [6] Y. Zhu, C. Wang, and R. Dai, "Document Image Binarization Based on Stroke Enhancement", *Proceedings of the 18<sup>th</sup> International Conference on Pattern Recognition (ICPR '06)*, Hong Kong, China, August 2006, vol. 1, pp. 955-958.
- [7] F. Chang, K. H. Liang, T. M. Tan, and W. L. Hwang, "Binarization of Document Images Using Hadamard Multiresolution Analysis", *Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR '99)*, Bangalore, India, September 1999, pp. 157-160.
- [8] H. J. Lee, and B. Chen, "Recognition of Handwritten Chinese Characters via Short Line Segments", *Pattern Recognition*, May 1992, vol. 25, no. 5, pp. 543-552.
- [9] J. Canny, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 1986, vol. 8, no. 6, pp. 679-698.
- [10] N. Otsu, "A Thresholding Selection Method from Gray-level Histogram", *IEEE Transactions on Systems, Man and Cybernetics*, March 1979, vol. 9, pp. 62-66.
- [11] J. Bernsen, "Dynamic Thresholding of Grey-level Images", *Proceedings of the 8th International Conference on Pattern Recognition*, Paris, France, 1986, pp. 1251-1255.
- [12] W. Niblack, "An Introduction to Digital Image Processing", Prentice-Hall, Englewood Cliffs, NJ, pp. 115-116, 1986.
- [13] J. Sauvola, and M. Pietikainen, "Adaptive Document Image Binarization", *Pattern Recognition*, 2000, vol. 33, no. 2, pp. 225-236.
- [14] Y. Yang, and H. Yan, "An Adaptive Logical Method for Binarization of Degraded Document Images", *Pattern Recognition*, 2000, vol. 33, pp. 787-807.
- [15] B. Gatos, I. Pratikakis, and S. J. Perantonis, "Adaptive Degraded Document Image Binarization", *Pattern Recognition*, March 2006, vol. 39, no. 3, pp. 317-327.
- [16] B. H. Kang, G. S. Han, H. G. Kim, J. S. Kim, C. R. Yoon, and M. S. Cho, "Fuzzy Inference and Logical Level Methods for Binary Graphic/Character Image Extraction", *1998 IEEE International Conference on Systems, Man and Cybernetics*, San Diego, California, USA, October 1998, vol. 5, pp. 4626-4629.
- [17] C. Wolf, J. M. Jolion, and F. Chassaing, "Text Localization, Enhancement and Binarization in Multimedia Documents", *Proceedings of the 16<sup>th</sup> International Conference on Pattern Recognition (ICPR '02)*, Quebec City, Canada, August 2002, vol.2, pp. 1037-1040.
- [18] H. Cao, and V. Govindaraju, "Handwritten Carbon Form Preprocessing Based on Markov Random Field", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, USA, June 2007, pp. 1-7.
- [19] C. Bastos, C. Mello, J. Andrade, D. Falcão, M. Lima, W. Santos, and A. Oliveira, "Thresholding Images of Historical Documents with Back-to-Front Interference based on Color Quantization by Genetic Algorithms", *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI '07)*, Patras, Greece, October 2007, vol. 1, pp. 488-491.
- [20] H. Kohmura, and T. Wakahara, "Determining Optimal Filters for Binarization of Degraded Characters in Color Using Genetic Algorithms", *18th International Conference on Pattern Recognition (ICPR '06)*, Hong Kong, August 2006, vol. 3, pp. 661-664.
- [21] T. Obafemi-Ajayi, G. Agam, and O. Frieder, "Ensemble LUT Classification for Degraded Document Enhancement", *Proceedings of SPIE, 15th Document Recognition and Retrieval Conference*, San Jose, California, USA, January 2008, vol. 6815.
- [22] LY. Fan, C. L. Tan, and LX. Fan, "Edge-preserving Prefiltering for Document Image Binarization", *International Conference on Image Processing*, Greece, October 2001, vol. 1, pp. 1070-1073.
- [23] V. Wu, R. Manmatha, and E. M. Riseman, "Textfinder: An Automatic System to Detect and Recognize Text in Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, November 1999, vol. 21, no. 11, pp. 1224-1229.
- [24] M. Kamel, and A. Zhao, "Extraction of Binary Character/Graphics Images from Grayscale Document Images", *CVGIP: Computer Vision Graphics and Image Processing*, May 1993, vol. 55, no. 3, pp. 203-217.