

PAPER • OPEN ACCESS

## Binarization of Document Images: A Comprehensive Review

To cite this article: Wan Azani Mustafa and Mohamed Mydin M. Abdul Kader 2018 *J. Phys.: Conf. Ser.* **1019** 012023

View the [article online](#) for updates and enhancements.

You may also like

- [A Comprehensive Review on Document Image \(DIBCO\) Database](#)  
W A Mustafa, Wan Khairunizam, I Zunaidi et al.
- [Binarization of Document Image Using Optimum Threshold Modification](#)  
Wan Azani Mustafa and Mohamed Mydin M. Abdul Kader
- [Comparative analysis of off-axis digital hologram binarization by error diffusion](#)  
Pavel A Cheremkhin, Ekaterina A Kurbatova, Nikolay N Evtikhiev et al.



**ECS**  
The  
Electrochemical  
Society  
Advancing solid state &  
electrochemical science & technology

**DISCOVER**  
how sustainability  
intersects with  
electrochemistry & solid  
state science research

# Binarization of Document Images: A Comprehensive Review

**Wan Azani Mustafa<sup>1</sup>, Mohamed Mydin M. Abdul Kader<sup>1</sup>**

<sup>1</sup>Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia

*wanazani@unimap.edu.my*

**Abstract.** Document image binarization is one important pre-processing step, especially for data analysis. Extraction of text from images and its recognition may be challenging due to the presence of noise and degradation in document images. In this paper, seven (7) types of binarization method were discussed and tested on Handwritten Document Image Binarization Contest (H-DIBCO 2012). The aim of this paper is to provide comprehensive review methods in order to binary document images in the damaging background. The results of the numerical simulation indicate that the Gradient Based method most effective and efficient compared to other methods. Hopefully, the implications of this review give future research directions for the researchers.

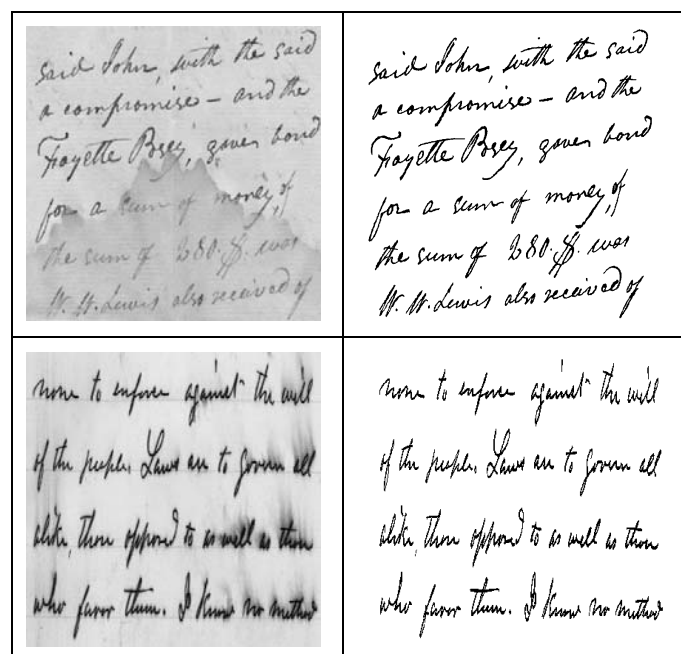
## 1. Introduction

There are many challenges addressed in handwritten document image binarization, such as faint characters, bleed-through and large background ink stains [1]–[4]. Document image binarization is the process that segments the document image into the text and background by removing any existing degradations [1]. Many document image binarization methods have been proposed in the literature. However, selecting the most optimum method for binarization is a difficult task due to the presence of a variety of degradations in document images.

Previous studies have primarily concentrated in order to propose a new method or algorithm due to solve the degradation of document images. In 2008, Nikolaos and Dimitrios review a few enhancement and binarization techniques to find the best approach for the future research [5]. They summarize that combination of pre-processing and binarization algorithm able to improve and finally provide the extensive method. Many researchers agree that it's very difficult to propose a perfect algorithm because the document image in badly condition dealing with many information such as text and structure [5], [6]. Gatos *et al.* discussed the challenges and strategies to improve the document image binarization based on a combination of Multiple Binarization Techniques and Adapted Edge Information [6]. This approach has a number of advantages: firstly, (i) combining the binarization results of several state-of-the-art methodologies; (ii) incorporating the edge map of the grey scale image; and (iii) applying efficient image post-processing based on mathematical morphology for the enhancement of the final result. Research finding by Shijian *et al.* also points towards the edge information to propose a new method [7]. However, they more concentrate on the surface and stroke edge. The result finding more effective compared to the Gatos method [6], Sauvola method [8] and Otsu method [9]. In 2009, Reza and Mohamed published a paper in which they described the new model of a low quality document image using virtual diffusion processes [10]. This technique focuses on the shadow- through and bleed-through problem.



Several studies investigating document binarization based on Otsu modification [11]–[13]. Starting 2010, Nina *et al.* mention the significant combination between the recursive extension of Otsu thresholding and selective bilateral filtering of scanned handwritten images [11]. This approach considers background estimation before applied the post- processing stage [11], [13]. The above findings contradict the study by Zhang and Wu. They examined the modification algorithm based on the Adaptive Otsu method [12]. The proposed based on three main steps; (1) applied the Wilner filter in order to eliminate noise, (2) improved adaptive Otsu's method, and (3) dilation and erosion operators were performed to preserve stroke connectivity and fill possible breaks, gaps, and holes. The advantage of this approach is that faster processing time compared to Recursive Otsu Thresholding Method [11] and AdOtsu method [13]. Similarly, Reza and Mohamed also proposed a new novel method based on the Otsu modification known as AdOtsu method [13]. The main idea of this technique was considered parameterless behavior such as average stroke width and the average line height. The positive result was achieved compared to Sauvola method, Otsu Method, and Lu and Tan method.



**Figure 1.** The document image; (left) the original image and (right) the benchmark image.

Figure 1 (left) has shown the input original image for binarization and figure 1 (right) has shown the binary image for the same. Document image binarization is generally performed in the pre- processing phase of distinctive archive picture handling related requisitions, for example, optical character distinguishment (OCR) and report picture recovery.

In this paper, a comprehensive review of 7 types of binarization methods such as Otsu method, Niblack method, Bernsen method and Gradient Based method was discussed. A few image quality assessment such as Misclassification Error (ME) and Peak Signal Noise Ratio (PSNR) was performed in order to compare the effectiveness of every method. Summary, this paper is organized in the following sections: Section 2 describes the review comparison binarization methods. Section 3 presents the analysis of results and Section 5 gives the conclusion of the results.

## 2. Reviews of Binarization Methods

This section explains a few popular binarization methods that were used in this study. Normally, a poor image quality caused by noise, illumination, artefacts from the camera, and degradation of the image affected the binarization process [14]–[16].

### 2.1 Gradient Based Thresholding

Haniza and Arof presented a new binarization method using gradient based thresholding by constructing a threshold surface [17]. This method contains three important phases: first, construct the inverse image  $T(i, j)$ , second obtaining the  $k$  value between -255 to 255, and finally applying binarization to separate object and background.

$$result(i, j) = \begin{cases} 0, (object) & \text{if } I(i, j) < T(i, j) + k_0 \\ 255, (background) & \text{if } I(i, j) > T(i, j) + k_0 \end{cases} \quad (1)$$

where  $I(i, j)$  is the intensity of the original image and  $k_0$  is the minimum sum of absolute difference intensity. Based on ME, the proposed binarization produced better and effective results compared to other methods.

### 2.2 Niblack Method

The main purpose of Niblack method is to set the threshold value based on local standard deviation and local mean. The threshold for each pixel was determined by [18]:

$$T(x, y) = m(x, y) + k\delta(x, y) \quad (2)$$

where, standard deviation  $\delta(x, y)$  and local mean  $m(x, y)$  were determined by 80 x 80 windowing size [19] and standard  $k$  value is - 0.2. This method does not work correctly if the image suffers from non-uniform illumination.

### 2.3 Otsu Method

The technique obtained the threshold value automatically based on global variance and between-class variance. In the non-uniform image, Otsu assumes the image contains two areas: dark and bright in order to purpose final algorithm [9]. Finally, Otsu thresholding is determined by:

$$k = \frac{\sigma^2 B}{\sigma^2 G} \quad (3)$$

where,  $k$  a threshold value,  $\sigma^2 B$  are a global variance of the entire image, and  $\sigma^2 G$  between-class variance.

### 2.4 Nick Method

Khurshid *et al.* proposed new strategies to improve the Niblack method by shifting the thresholding value downward [20]. The threshold is obtained based on the following equation:

$$T(x, y) = m + k \sqrt{\frac{I^2 - m^2}{N}} \quad (4)$$

The  $k$  factor value is similar with Niblack and the windowing size is defined as 15 x 15, while  $I$  and  $m$  represent the intensity pixel and mean of grey scale image.  $N$  represents the image size.

### 2.5 Bradley Method

By using the integral image as the input image, this method is an improvement of Wellner's method [21] and it is robust to illumination changes within the image. The key idea of the algorithm is that every image's pixel is set to black if its brightness is  $T$  percent lower than the average brightness of surrounding pixels in the window of the specified size, otherwise it is set to white. The default windowing size is 15 by 15 and  $T$  is 10 [22].

$$T = m \left( 1 - \frac{k}{100} \right) \quad (5)$$

### 2.6 Bernsen Method

The Bernsen algorithm is based on the estimation of a local threshold value for each pixel. This value is assigned as the local threshold value only if the difference between the lowest and the highest grey level value is bigger than a threshold  $k$ . Otherwise, it is assumed that the window region contains pixels of one class (foreground or background). The default windowing size ( $w$ ) is 3-by-3 and  $k$  is 15 [23]. The final equation as follows;

$$T(x, y) = \frac{Z_{\max} + Z_{\min}}{2} \quad (6)$$

where,  $Z_{\min}$  and  $Z_{\max}$  are the lowest and highest grey level pixel values.

### 2.7 Local Adaptive Thresholding

Local Adaptive Thresholding is a basic and simple algorithm to separate the foreground from the background with non-uniform illumination. For each pixel in the image, a threshold has to be calculated. If the pixel value is below the threshold, it is set to the background value, otherwise, it assumes the foreground value. The default local windowing size ( $w$ ) is 15 by 15 and local threshold ( $T$ ) is a 0.05 [14].

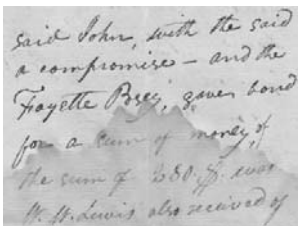
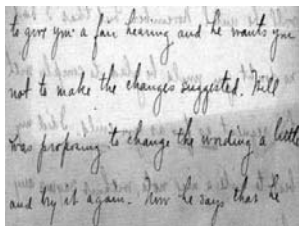
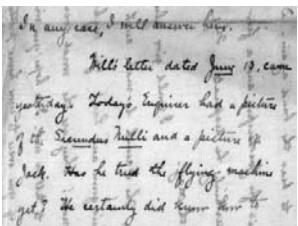
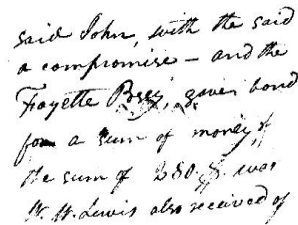
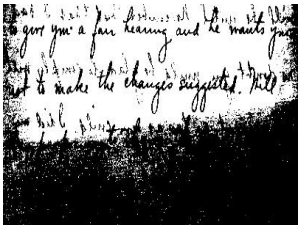
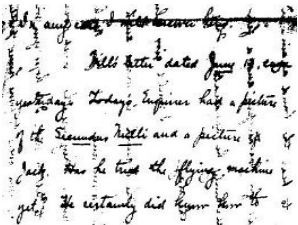
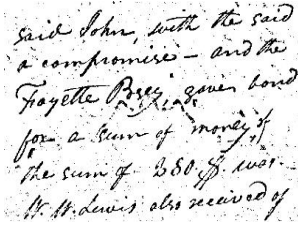
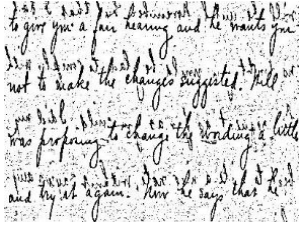
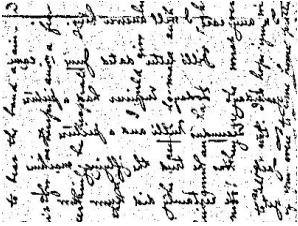
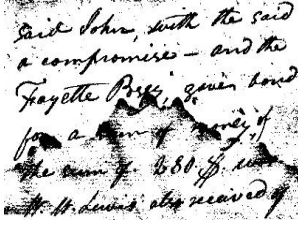
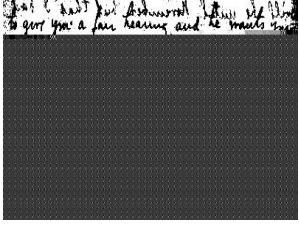
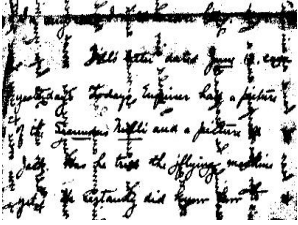
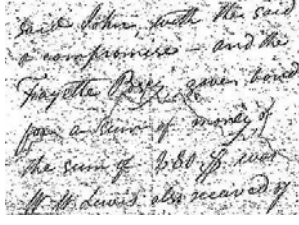


$$T = \frac{\max + \min}{2} \quad (7)$$

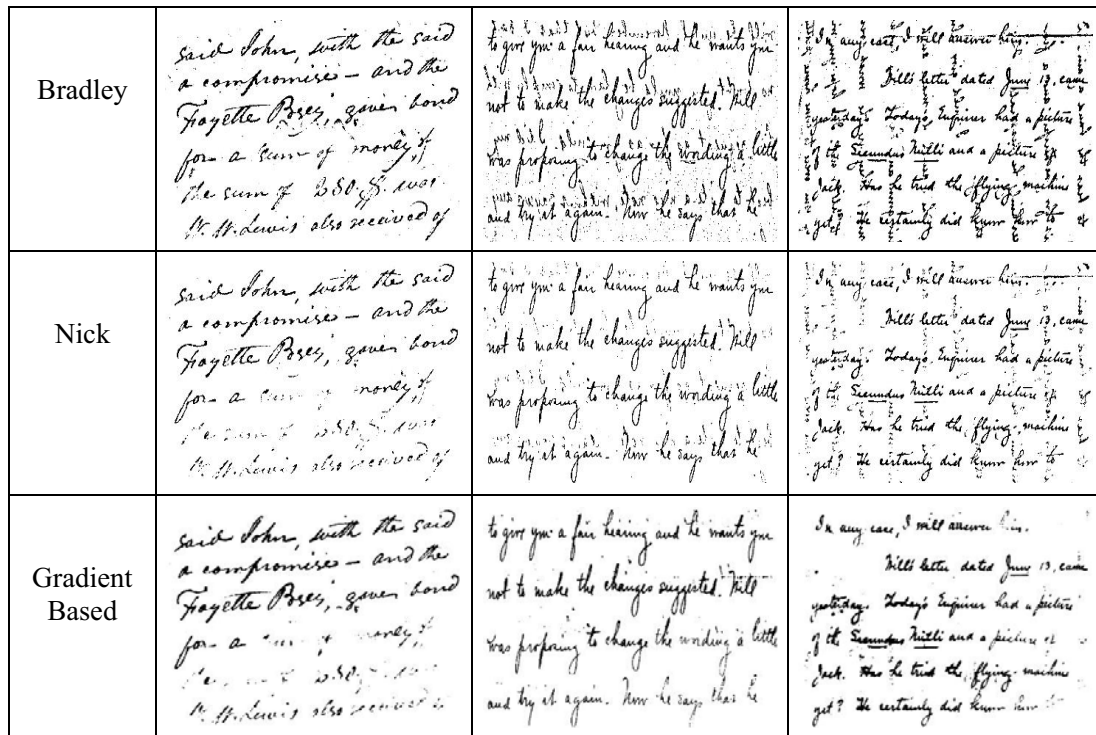
## 3. Result

In this paper, 14 document images from Handwritten Document Image Binarization Contest (H-DIBCO) [24] dataset experimented. The images contain various degradations such as shadows, non-uniform illumination, stains, smudges, bleed-through and faint characters [1]. All the processed images are in greyscale images and the size of each image is 400 × 400 pixels, 72 dpi, and 8-bit depth. All the programs were written in C programming and ran using Ubuntu with a Linux 3.5 from an Asus laptop with AMD Athlon™ II P320 Dual-Core Processor 2.10GHz and 3.00GB RAM. The seven (7) types of binarization were tested and shown in figure 2. Three examples of document images were illustrated in figure 2. The



first row shows the original image with degradation and shadowing problem and follows by binarization methods such as Otsu, Local Adaptive, Niblack, Bernsen, Bradley, Nick, and Gradient Based method. Based on observation, the resulting image from Bradley and Gradient Based is satisfied compared to other methods.

Original			
Otsu			
Local Adaptive			
Niblack			
Bernsen			



**Figure 2.** The comparison of binarization method on document images.

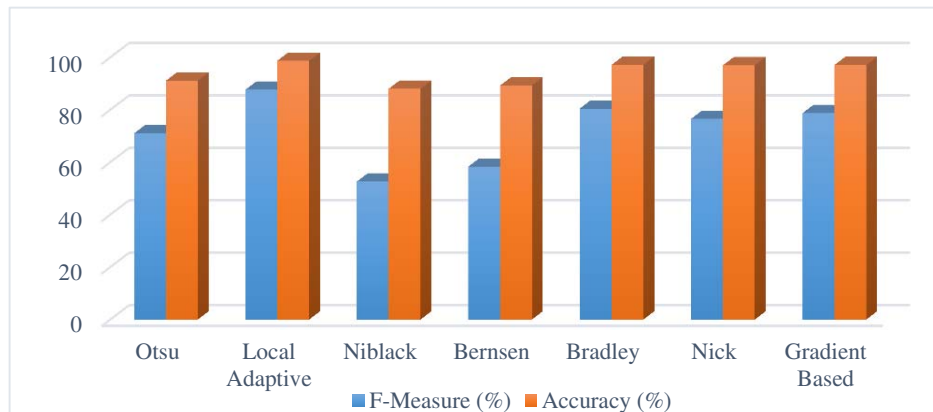
Next, a few image quality assessment (IQA) was calculated to find the best binarization methods. In this paper, the evaluation based on ME (Misclassification Error), PSNR (Peak Signal to Noise Ratio) and MPM (Misclassification Penalty Metric) was obtained. All the assessment equation can be referred on H-DIBCO [25], [26]. The lower of ME and MPM, while higher of PSNR value represent the good binarization result. Table 1 provides the comparison result of 7 binarization methods. The lowest value of ME indicates a good performance in case of binarization as displays by the Bradley method (ME = 0.0285) and follow by Gradient Based method (ME = 0.0292) respectively. Opposite, the higher value in term of PSNR present the best binarization was come from the Nick method (PSNR = 24.8349) compared to others. Besides, comparisons based on MPM show that the performance of the Local Adaptive method obtained 0.2376, which is lower than the other methods. The highest value of MPM came from the Otsu method (24.1048) since it produces a degraded and non-uniform image.

**Table 1.** A few evaluation techniques

Method	ME	PSNR	MPM ( $10^{-3}$ )
Otsu	0.0886	16.5185	24.1048
Local Adaptive	0.0390	16.0604	0.2376
Niblack	0.1196	9.7405	7.2171
Bernsen	0.1073	12.1823	3.2518
Bradley	0.0285	19.3170	0.9831
Nick	0.0301	24.8349	1.7177
Gradient Based	0.0292	19.6308	1.5587
Method	ME	PSNR	MPM ( $10^{-3}$ )
Otsu	0.0886	16.5185	24.1048
Local Adaptive	0.0390	16.0604	0.2376
Niblack	0.1196	9.7405	7.2171
Bernsen	0.1073	12.1823	3.2518
Bradley	0.0285	19.3170	0.9831
Nick	0.0301	24.8349	1.7177
Gradient Based	0.0292	19.6308	1.5587
Method	ME	PSNR	MPM ( $10^{-3}$ )
Otsu	0.0886	16.5185	24.1048
Local Adaptive	0.0390	16.0604	0.2376
Niblack	0.1196	9.7405	7.2171
Bernsen	0.1073	12.1823	3.2518
Bradley	0.0285	19.3170	0.9831
Nick	0.0301	24.8349	1.7177
Gradient Based	0.0292	19.6308	1.5587



In order to prove the sensitivity of each binarization methods, the evaluation based on the F-measure and the Accuracy was obtained. Normally, the higher on F- measure and Accuracy shows the effectiveness of binarization method [7], [27]. The result is presented in figure 3. As shown in the figure, Local Adaptive method achieved the highest result which is (F-measure = 87.678) and (Accuracy = 98.729) compared to other methods. While, the lowest value in term of F-measure and Accuracy came from Niblack method (F-measure = 52.657) and (Accuracy = 88.037).



**Figure 3.** Comparison of seven types binarization methods based on F-measure and Accuracy

#### 4. Gaps in Literature

Many techniques have been proposed so far for document binarization as shown in literature survey. It has been concluded from the existing research is that no technique is perfect for every case. Therefore still some research is required in this field of image binarization. Following are the main limitations of this research work:-

1. Many researchers have used image filters to reduce the noise from the image, but the use of the guided filter (best edge preserving filter) is not found. It may increase the accuracy of the available binarization methods
2. In the most of the techniques, the contrast enhancement is either done by traditional methods or not done. So, adaptive contrast enhancement is required.
3. Most of the methods have neglected the use of edge map which has the ability to map the exact character in an efficient manner.

#### 5. Conclusion

This paper has focused on the seven (7) types of binarization method and experimented on degraded document image. Document binarization is an important application of vision processing. The main objective of this paper is to evaluating the shortcomings of algorithms for degraded image binarization. It has been found that each technique has its own benefits and limitations; no technique is best for every case. The main limitations of existing workers are found to be noisy and low intensity images. In near future, we will propose a new algorithm which will use the more reliable methodology to enhance the work. We will propose a new algorithm which will use nonlinear enhancement as a pre-processing technique to improve the results further.

#### 6. References

- [1] 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, Jan. 2014.
- [2] W. A. Mustafa, "A proposed optimum threshold level for document image binarization," *J. Adv. Res.*

- Comput. Appl.*, vol. 1, no. 1, pp. 8–14, 2017.
- [3] W. A. Mustafa and H. Yazid, “Illumination and Contrast Correction Strategy using Bilateral Filtering and Binarization Comparison,” *J. Telecommun. Electron. Comput. Eng.*, vol. 8, no. 1, pp. 67–73, 2016.
  - [4] W. A. Mustafa and H. Yazid, “Background Correction using Average Filtering and Gradient Based Thresholding,” *J. Telecommun. Electron. Comput. Eng.*, vol. 8, no. 5, pp. 81–88, 2016.
  - [5] N. Ntogas and V. Dimitrios, “A Binarization Algorithm for Historical Manuscripts,” in *International Conference on Communications*, 2008, pp. 41–51.
  - [6] B. Gatos, I. Pratikakis, and S. J. Perantonis, “Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information,” *19th Int. Conf. Pattern Recognit.*, pp. 1–4, Dec. 2008.
  - [7] S. Lu, B. Su, and C. L. Tan, “Document image binarization using background estimation and stroke edges,” *Int. J. Doc. Anal. Recognit.*, vol. 13, no. 4, pp. 303–314, Oct. 2010.
  - [8] J. Sauvola and M. Pietikäinen, “Adaptive document image binarization,” *Pattern Recognition*, vol. 33, pp. 225–236, 2000.
  - [9] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” in *IEEE Transactions On Systems, Man, And Cybernetics*, 1979, vol. 20, no. 1, pp. 62–66.
  - [10] R. F. Moghaddam and M. Cheriet, “Low quality document image modeling and enhancement,” *Int. J. Doc. Anal. Recognit.*, vol. 11, no. 4, pp. 183–201, Feb. 2009.
  - [11] O. Nina, B. Morse, and W. Barrett, “A Recursive Otsu Thresholding Method for Scanned Document Binarization,” in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2010, pp. 307–314.
  - [12] Y. Zhang and L. Wu, “Fast Document Image Binarization Based on an Improved Adaptive Otsu ’ s Method and Destination Word Accumulation,” *J. Comput. Inf. Syst.*, vol. 6, no. 7, pp. 1886–1892, 2011.
  - [13] Reza Farrahi Moghaddamn Mohamed Cheriet, “AdOtsu: An adaptive and parameterless generalization of Otsu’s method for document image binarization,” *Pattern Recognit.*, vol. 45, pp. 2419–2431, Jun. 2012.
  - [14] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Third. Upper Saddle River, NJ, USA: Prentice Hall, 2008.
  - [15] M. D. Vlachos and E. S. Dermatas, “Non-uniform illumination correction in infrared images based on a modified fuzzy c-means algorithm,” *J. Biomed. Graph. Comput.*, vol. 3, no. 1, pp. 6–19, Nov. 2012.
  - [16] M. A. Bakhshali, “Segmentation and enhancement of brain MR images using fuzzy clustering based on information theory,” *Soft Comput.*, vol. 20, pp. 1–8, 2016.
  - [17] H. Yazid and H. Arof, “Gradient based adaptive thresholding,” *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 926–936, Oct. 2013.
  - [18] W. Niblack, *An Introduction to Digital Image Processing*. Englewood Cliffs: Prentice-Hall, 1986.
  - [19] J. Shi, N. Ray, and H. Zhang, “Shape based local thresholding for binarization of document images,” *Pattern Recognit. Lett.*, vol. 33, no. 1, pp. 24–32, Jan. 2012.
  - [20] K. Khurshid, I. Siddiqi, C. Faure, and N. Vincent, “Comparison of Niblack inspired Binarization methods for ancient documents,” *Proc. SPIE-IS&T Electron. Imaging*, vol. 7247, pp. 1–9, 2009.
  - [21] Wellner, “Adaptive thresholding for the digital desk,” *Tech. Rep. Eur.*, pp. 93–110, 1993.
  - [22] D. Bradley and G. Roth, “Adaptive Thresholding Using the Integral Image,” *J. Graph. GPU, Game Tools*, vol. 12, no. 2, pp. 13–21, 2011.
  - [23] J. Bernsen, “Dynamic thresholding of grey-level images,” in *Proceedings of the Eighth International Conference on Pattern Recognition*, 1986, pp. 1251–1255.
  - [24] K. Zagoris, “H-DIBCO ’16,” 2016. [Online]. Available: <http://vc.ee.duth.gr/h-dibco2016/>. [Accessed: 01-Jan-2016].
  - [25] W. A. Mustafa, H. Yazid, and S. Yaacob, “Illumination Normalization of Non-Uniform Images Based on Double Mean Filtering,” in *IEEE International Conference on Control Systems, Computing and Engineering*, 2014, pp. 366–371.
  - [26] W. A. Mustafa, H. Yazid, M. Jaafar, M. Zainal, A. S. Abdul-, and N. Mazlan, “A Review of Image Quality Assessment (IQA): SNR, GCF, AD, NAE, PSNR, ME,” *J. Adv. Res. Comput. Appl.*, vol. 7, no. 1, pp. 1–7, 2017.
  - [27] S. Thilagamani and N. Shanthi, “Gaussian and Gabor Filter Approach for Object Segmentation,” *J. Comput. Inf. Sci. Eng.*, vol. 14, pp. 1–7, 2014.