An Efficient Approach For Improved Scene Text In Disturbed Document Images Using Different Binarization Methods

Kirti Tomar
Student
MITS Gwalior

Jaimala Jha
Asst. Prof.
MITS Gwalior

***Abstract*—The process of Image binarization is the parting of pixel values in 2 parts, white the same as background also black as the foreground. Thresholding acting as a major in the binarization of the images. Thresholding which can be classified by global as well as local. In the images by means of uniform contrast allocation of background as well as foreground such as document form of images, global form of thresholding is much better. In the despoiled document images, wherever significant noise / variation form background in disparity & exists the illumination, here present several pixels which can't easily be classified as the foreground or background. This paper analyses the five thresholding techniques such as Otsu Thresholding, Kittler and Illingworth thresholding, structural symmetric stroke, local adaptive thresholding, and niblack thresholding based on average filtering method for general gray images, normal and abnormal documentation image scans. The proposed method was being verified making use of a dataset in a different way illuminated document images being subjected to additional type of recognition of text. Experimental outcome, being expressed as the Precision & Recall, PSNR, F-Measure values for gained text strings, are being promising also agrees the worth of proposed method.**

Keywords*— *document images binarization; Otsu Thresholding; Kittler method ; Locally Adaptive Thresholding; niblack thresholding based on average filtering.

I. INTRODUCTION

Binarization is usually performed in the preprocessing stage. The aim of the binarization is to spilt the foreground text from the document background because text belongs to foreground and the motive is to recover text from images of historical document. Thresholding is important parameter denotes to gray-scale image conversion to a binary form of image. The thresholding being degraded by the documents is a key task. Primarily image binarization approach is a segregation of the pixel values in 2 different groups, white for background & the black for foreground.

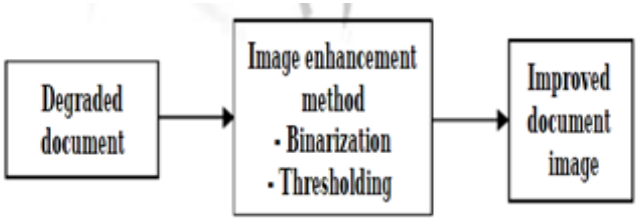

Fig.1. Binarization of degraded document image [1]

Some degraded document images are shown below, in which different kind of degradation are present. Every year the contest for image binarization methods is held by the experts that is known as DIBCO (Document Image Binarization Contest), images are taken from this series dataset only. Images (a) and (b) are taken from DIBCO 2009 & 2011 handwritten dataset respectively; image (c) is taken from H-DIBCO 2010 machine printed dataset. Binarization deals with this kind of degradation shown in the images and recover the images in its normal original form.[1]

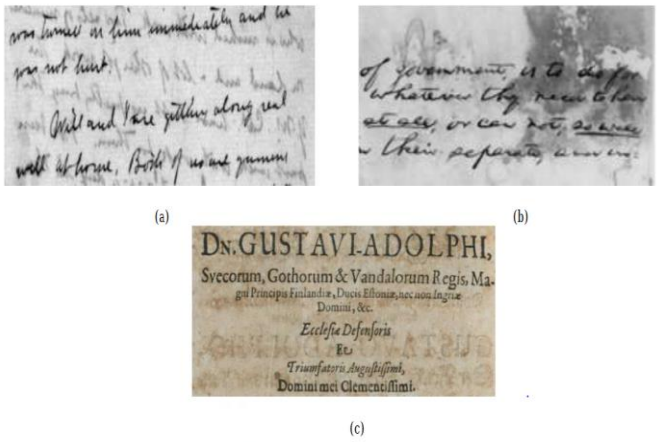

Fig.2. Degraded document image examples (a)–(c) are taken from DIBCO series datasets [1]

Document Image Binarization for the document analysis image also processing is pre-processing step. It enhances document processing methods such as OCR & the design analysis efficiency. Image Binarization is converting image of a document into a two-level image of a document. Pixels of the image are divided within dual pixel gathering, that is black and white. Document segmentation within the foreground text also the background is primary objective of picture binarization. Thresholding is the easiest approach to binarisation. In thresholding an ideal value of threshold is selected as well as by the comparison through this threshold value pixels are being categorized as foreground or background. But some papers suffer from degradation in actual life, such as irregular illumination, background noise, bleeding through, as well as lighting and contrast variation. It is a difficult job to select an ideal threshold value for such papers. False threshold value assessment results in the pixel being misclassified as a foreground or background. This impacts binarization outcomes and recognition of pattern applications precision. Binarization is generally classified as global & the local. Global type binarization is method used to binarize the entire picture with a single limit value. It is a quick method, but it fails for complicated background papers.

A distinct threshold value is selected for each pixel in the Local binarization method in place of a threshold for entire picture. Depending on the neighborhood pixels, the threshold is selected. Local thresholding does not yield excellent outcomes for background noise records. Binarization methods can be classified according to the threshold value selection criteria. Some of the threshold selection techniques are dependent on histogram, clustering, local techniques of adaptation. Other techniques use decision tree, combining different binarization techniques with iterative methods. [2]

Image handling methods have a broad range in today's aera. In many fields such as astronomy, distant sensing, microscopy or tomography, cryptography, the picture can be helpful. There are many pictures of the ancient books that are very helpful to us but these pictures are unreadable owing to non-maintenance. Captured pictures are blurred, noisy and of inadequate spatial or temporal resolution due to measuring device imperfections and observed scene instability. These images have become degraded images, and although they are very useful to us, we cannot use them. Sometimes some pictures are manually degraded owing to which the image quality is reduced and that helpful picture becomes a waste picture for us. The degraded pictures are combined in front and background format. This background can be separated from the foreground text. The suggested method provides an effective way to distinguish the text from the loud pixels of the background. The picture is carried through the various techniques that generate the readable picture output. [3]

Fig.3. Example of binarization [3]

II. DIFFERENT METHODS

A. Otsu Thresholding

Otsu's thresholding technique includes iterating by various threshold values also calculating a spread measure for each side of the limit for pixel concentrations, i.e. pixels falling in foreground or in background. The key is to search value of threshold at minimum of the sum of the foreground as well as background spreads. Based on the histogram form, OTSU's thresholding technique is used in image processing for automatic binarization level choice. The algorithm assumes that there are two fundamental classes in the picture: Foreground and Background. It then calculates an ideal threshold value which minimizes the weighted within these two classes ' class variances. [4] Mathematically, minimizing the variance within the class is the same as maximizing the variance between classes. [5]

B. Kittler method

Kittler & Illingworth's thresholding techniques also called as minimum error thresholding (MET) method [14]. The algorithm is dependent on Bayesian form of classification rule. This ways are initially computes bi-model histogram of gray level image which is simply distributed. Then calculates the priori form of probability of gray level of the histogram and find the mean of the overall probability. Its initial form of threshold for offered image also divides image in foreground & background form of classes the probability, the mean as well as standard form of deviation are calculated by the following equations. [6]

$$p_i^{(T)} = \sum_{g}^{h} h(g)$$

$$p_i(T) = \sum_{g=a}^{h} h(g)$$

$$\mu(T) = \frac{1}{p_i(T)} \sum_{g=a}^{h} h(g)$$

$$\sigma^2(T) = \frac{1}{p_i(T)} \sum_{g=a}^{b} \left(g - \mu_i(T)\right)^2 h(g)$$

$$where, a = \begin{cases} 0 & i = 1 \\ T+1 & i = 2 \end{cases} \text{ and } b = \begin{cases} T & i = 1 \\ n & i = 2 \end{cases}$$

The criterion function is,

$$J(T) = 1 + 2\left[p_1(T) log\, \sigma_{bg}(T) + p_2(T) log\, \sigma_{fg}(T)\right] - 2\left[P_1[T] + log p_2(T)\right]$$

The criterion function J(T) can easily computed along with the finding of its minimum error threshold is comparatively simple task and finds the threshold T_2 as given below:

$$T_2 = arg_{1 \le t \le n} min J(T)$$

C. Structural Symmetric Stroke

At the stroke stage, the text's symmetry rests on the stroke boundary's gradient orientation and magnitude. This estate was studied in the job of SWT. In this document we use a greater amount of the symmetry property. The main observation is that a text region is generally highly self-similar and contrasts strongly with its local context in terms of low-level picture signals such as gradient and texture.[7]

D. Locally Adaptive Thresholding

A threshold T(x,y) is a value such that;

$$b(x,y) = \begin{cases} 0 & if\ I(x,y) \le T(x,y) \\ 1 & otherwise \end{cases}$$

Where the binarized image is b(x, y) & pixel intensity is I(x, y)[0,1] at the location (x, y) of the image I. A threshold is being calculated for every form of pixel in local adaptive method, depending on few of the local statistics like neighborhood pixel range, variance, or surface-fitting parameters. It is approached in different methods like background subtraction, model of water flow, mean & standard pixel value derivation, and contrasting local picture. few disadvantages of local thresholding methods dependent on region size, individual picture features, also consuming the time. few scientists therefore make use of hybrid strategy which applies worldwide & local techniques of thresholding as well as few make use of morphological operators. Niblack, Sauvola as well as Pietaksinen using method of local variance when Bernsen utilizes mid-range throughout the local block.[8]

E. Niblack's Method

Niblack is a technique for local thresholding. For each and every pixel, a distinct threshold value is calculated in local thresholding techniques. It utilizes local image statistics for calculating the limit, such as variance, range. A rectangular window being slided by the gray scale picture in the Niblack technique to estimate the pixel limit. To estimate the limit, it utilizes the mean local statistics and normal window deviation. Estimates of threshold T(i, j) as given below.

$$T(i,j) = \mu + k * \sigma$$

In Eq. μ is mean of window as well as σ is normal window deviation. The k is constant, defining binarization size & quality. Since this technique depends on the local characteristics of the picture, it is influenced by the blank fields of the picture and is not effective for background noise images.[9]

III. LITERATURE SURVEY

Mehmet Yasin AKPINAR, et al [2018] The conversion of image-based documents into digital and processible forms can be accomplished quite successfully with optical character recognition (OCR) tools. However, there are still problems with preserving the format on the original document. An important one of these problems is the reading of the tabular data. In this paper, a method is proposed in which the tabular data contents of hard-copy documents is extracted from the text and character positions which are obtained from an OCR tool and transferred to digital forms. The performance of the method is measured by the number of detected rows and columns and presented with the results of other commercial products.[10]

Harneet Singh et al [2018] Since the information accumulation age began, data entry has been a hectic task. Apparently, that's why data entry has a full-scale career. The technique of data entry differs according to the demands. Punch cards were used in the early days of pcs and gradually keyboards and mouse came into the image that we are still using. Touchscreens did not take much time to substitute physical keyboard inputs and now the age of optical input is here as a consequence of human intelligence and innovation. Where the user does not even have to worry about data entry, but can simply use an optical reader or scanner for data entry. The individual components such as text, pictures, and unique characters can be differentiated using the computing energy. The job is done by OCR-Optical Character Recognizer. When it comes to character recognition, OCR operates comparable to humans as it retains a character database and compares all the components scanned with the database, making it really easy to comprehend. This article describes an OCR's work in its various phases. That research enables in discovering the standard system's multiple disadvantages. The document also discusses how to eliminate these shortcomings and how to achieve a better OCR that is prepared for the future.[11]

N Prameela et al [2017] The Aim Recognizing offline Handwritten Telugu characters making use of Optical character recognition, OCR is the very common also difficult pattern recognition topics This article proposes a three-stage OCR scheme for the Telugu papers, namely pre-processing, extraction of features as well as classification. We used median filtering on characters of input in the pre-processing phase and applied technique of normalization and skeletonization over characters to extract border border pixel points. In extraction stage of the feature, the individual characters are initially divided into 33 grids & centroid of the corresponding is evaluated for all nine zones. This allows us to identify personalities of various types. Thereafter, we drew the symmetric horizontal as well as vertical projection angel to image closest pixel known as Binary External Symmetry Axis Constellation for the handwritten text unconstrained form. Place where calculated Euclidean horizontal as well as vertical distance from centroid of each area for the same closest pixel. we Then calculated Euclidean mean range as areas mean angular values. This is regarded as our suggested system's main function values. Finally, as a classifier, both SVM , Quadratic Discriminate Classifier (QDA) were needed independently. [12]

Thulasi Kishna N.P et al [2017] This document reflects an appealing technique for converting the picture into an editable text as written by OCR. It demonstrates a computer's capacity to obtain and acknowledge handwritten input with this offline handwritten character recognition technique. Computers may find it difficult to decipher with distinct fonts and styles the precise handwritten characters. This paper focuses on handwritten Malayalam recognition(a Language of South India) characters. So cursive form of Malayalam characters been known by Hidden Markov Model (HMM). The classification performed with Artificial Neural Network (ANN). Recognition of Handwritten character through the high form of accuracy also the effective form of method so as to know the cursive letters are of much involved in proposed form of system.[13]

Michael D. Kim et al [2017] This article provides a quantitative assessment using an optical character recognition (OCR) technology of the dynamics-based de-blurring technique. Despite the study of different image de-blurring algorithms, there was no normal performance metric; de-blurred pictures were often assessed in a qualitative way. In this research, using a robotic vision system, blurry pictures comprising alphanumeric characters were acquired during fast movement. The blurry pictures acquired were restored using the de-blurring technique based on dynamics. OCR rates from the deblurred pictures were calculated by the dynamicsbased technique and contrasted with those by other wellknown techniques for a quantitative assessment. Results o f theexperiment show that the technique based on dynamics ha s the highest quantitative results.[14]

Hui Zho et al [2016] An significant region of pattern recognition is optical character recognition (OCR). Automatic OCR form of system is a type of smart system that is much essential for present automation applications like ATMs. A RMB OCR (ROCR) scheme is introduced in this document using FPGA, consisting of the module Character Segmentation (CS) as well as OCR. The application of the CME M7 was being required. 83.30% Logic Elements (LEs) of M7 FPGA was being needed for implementation of system. The entire system operates at a peak frequency of the 100MHz also is capable of one picture processing per 6ms at a good 98.45 percent identification level. A full ROCR system solution on FPGA is provided in this article. The solution involves the CS and OCR phases algorithms and implementations. The CME M7 chip was needed to carry out the suggested job and test it. The entire device consumes only 83.3 percent of an M7's on-chip LE, operates at a peak frequency of the 100 MHz as well as it is capable of one picture processing in 6ms with a good 98.45 percent identification rate. [15]

IV. PROPOSED METHODOLOGY

Problem statement:

1. The main problem with local adaptive thresholding technique is region size dependant, individual image characteristics, and time consuming.

2. some researchers use a hybrid approach that applies both global and local thresholding methods. and some use morphological operators. The major problem with thresholding is that we consider only the intensity, not any relationships between the pixels. There is no guarantee that the pixels identified by the thresholding process are contiguous.

Propose work

This is an implementation of a traditional Niblack local image thresholding with an integral image method, which guarantees constant computation time regardless of the neighborhood size. The integral image method is used only in `average filter`. The integral image is calculated with:
% Matrix 't' is the sum of numbers on the left and above the current cell.It seems to me the image must be converted to an integral image first
The value of the weight 'k' is used to control and adjust the effect of standard deviation due to objects features. Niblack suggests the value of 'k' to be -0.2. However, the algorithm is sensitive to the 'k' value and the optimal value varies from image to image. Common values are between -0.2 and -0.1 but it makes sense to try values from -0.3 to 0.3
.

Propose Algorithm:

Step.1. First we browse image from dataset.
Step.2. Apply Otsu Thresholding method on this original image.
Step.3. Apply Kittler method on this original image.
Step.4. Use structural symmetric stroke(sss) on this original image.
Step.5. Use Thresholding method on this original image.
Step.6. Use local adaptive thresholding method on this original image.
Step.7. Use niblack thresholding based on average filtering method on this original image.
Step.8. Calculate parameter PSNR, precision, recall and F-measures.
Step.9. Exit.

Flow Diagram

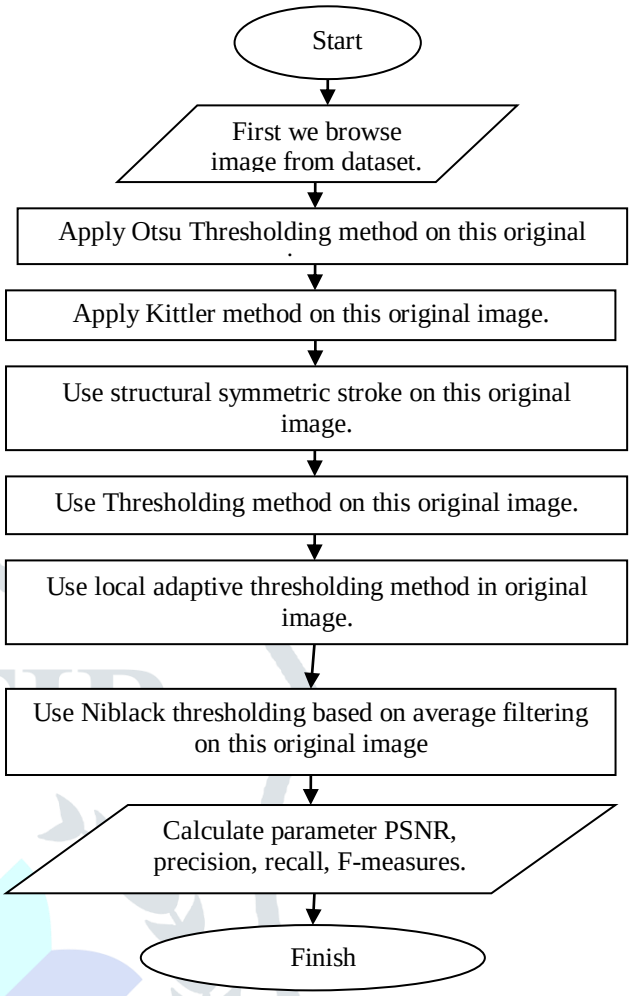

Fig. 4. Flow diagram of the proposed methodology.

V. PERFORMANCE METRICES

The metrics those appropriate for performance comparison of various algorithms of binarization being as PSNR, FMeasure, Precision & Recall.

A. PSNR

PSNR is required to make sure same among 2 form of images. It's required for an images containing noise. PSNR is shown by Eq.1.

$$PSNR = 10log_{10}\frac{C * C}{MSE} \quad (1)$$

Where C is the constant & MSE (mean square error) shown in the divergence among the distorted image as well as real image. The PSNR value should be much more for improved outcome. [9]

B. PRECISION

Precision is the true form of pixels being mined split through sum form of pixels being mined. Precision is shown in the Eq.2.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

TP shows true positive which is pixels those are the foreground in together ground truth as well as binarized form of image. FP shows the false positive that is pixels being recognized as foreground within the binarized form of image other than are in fact background into the ground of the truth image.

C. RECALL

Recall is true form of pixels being mined split through total no. of true pixels. Recall is given by Eq. 3

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

TP shows true positive that is pixels those are foreground within together ground truth & binarized image. FN shows false negative that is pixels being identified as the background in binarized image as they are normally foreground in the ground type of truth image.

D. F-MEASURE

F-Measure is illustrated as in Eq. 4

$$F - Measure = \frac{(2 * Precision * Recall)}{Precision + Recall} \quad (4)$$

F-Measure is harmonic mean of Precision as well as Recall. It's the value as high for much improved outcome.

VI. EXPERIMENT RESULT ANALYSIS

We collect several document images which are often used in the comparison experiments. With these images, we evaluate the performance of our proposed method and compare with several existing methods. In our experiment, after the Otsu Thresholding, documented image restores method works based on Kittler method, structural symmetric stroke local adaptive thresholding and average filtering. In order to compare the results of the above methods quantitatively, we apply PSNR, Precision, Recall, F-measure and time is applied to evaluate the binarization. The algorithm being designed on the MATLABR18 (A) making use of the Image Processing Toolbox. In an implementation, given algorithm is compared to various form of algorithms. As we have viewed in the experimental results. To compare the results of the above-mentioned methods.

There are 9 steps in this menu bar:

Fig. 5. Browse original image from dataset.

Fig. 6. Apply Otsu on this original image.

Fig. 7. Apply Kittler's method on this original image.

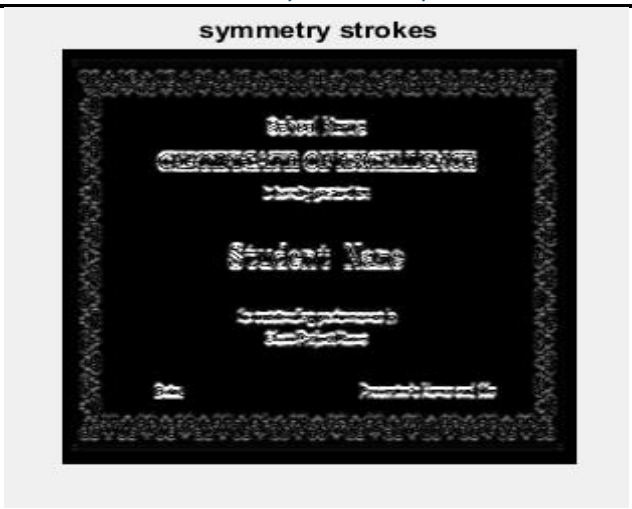

Fig. 8. Use structural symmetric stroke on this original image.

Fig. 11. Niblack thresholding based on average filtering.

Fig. 9. Use Thresholding method on this original image.

Fig. 10. Use local adaptive thresholding in original image.

TABLE I. COMPARISON TABLE OF BASE RESULTS ON DIFFERENT PARAMETERS.

Image name	PSNR	Precision	Recall	F-measure	time
Images	7.3650	0.65085	0.56210	60.32283	1.273313
certificate	9.0129	0.56953	0.88518	69.31065	1.236021

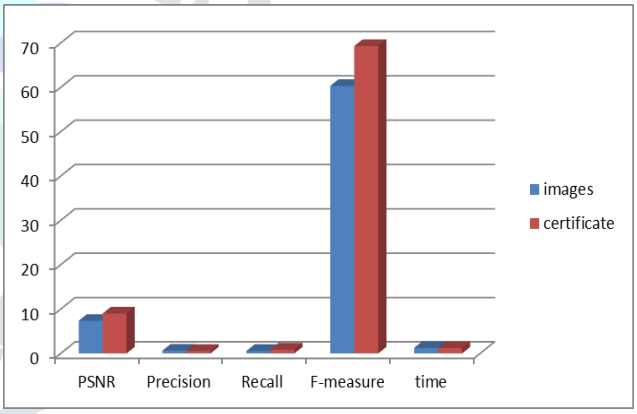

Fig. 12. Graph comparison of Base results of different parameters.

TABLE II. COMPARISON TABLE OF PROPOSES RESULTS ON DIFFERENT PARAMETERS.

Image name	PSNR	Precision	Recall	F-measure	time
Images	12.8534	0.95148	0.97919	96.51372	0.312452
Certificate	13.2199	0.92873	0.94743	93.79845	0.364395

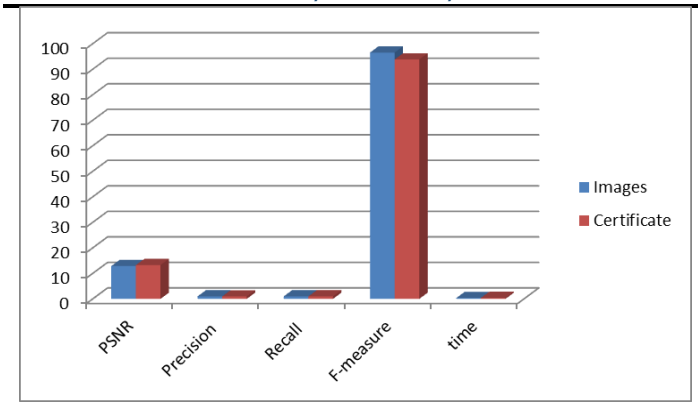

Fig. 13. Graph comparison of Propose results of different parameters.

Conclusion

Binarization is key step for the document image of the analysis with the recognition system. Generally, a huge no: of binarization ways are being proposed which is to address various form of binarization problems. The main goal of binarization is to diverse text from the degraded document background. This paper analyse the five thresholding techniques such as Otsu Thresholding, Kittler and Illingworth thresholding, structural symmetric stroke, local adaptive thresholding, and ni black thresholding based on average filtering method for general gray images, normal and abnormal documentation image scans. Niblack thresholding based on average filtering thresholding search to generate much enhanced outcome when compared to different ways at time of considering document images scans.

References

[1] Amoli Panchal, Chintan Panchal, Bhargav Shah, "Image binarization techniques for degraded document images: A review", international journal of engineering development and research, 2017IJEDR, volume 5,pp.1191-1195.

[2] N. Chaki, S. H. Shaikh, and K. Saeed, "Exploring Image Binarization Techniques", Studies in Computational Intelligence, vol. 560, pp. 5–16, 2014.

[3] Yogita Kakad1 , Savita Bhosale, "Survey on Document Image Binarization for Degraded Document Images" , International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064, Volume 4 Issue 1, January 2015.

[4] Sunil L. Bangare, Amruta Dubal, Pallavi S. Bangare, and Dr. S. T. Patil, "Reviewing Otsu's Method For Image Thresholding", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 9 (2015) pp. 21777-21783.

[5] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," in IEEE Transactions On Systrems, Man, And Cybernetics, 1979, vol. 20, no. 1, pp. 62–66.

[6] Pavlos Stathis, Ergina Kavallieratou, and Nikos Papamarkos, "An Evaluation Technique for Binarization Algorithms", Journal of Universal Computer Science, vol. 14, no. 18 (2008), pp. 3011-3030.

[7] Zheng Zhang, Wei Shen, Cong Yao, and Xiang Bai, "Symmetry-Based Text Line Detection in Natural Scenes", IEEE explore, 2015, pp.2558-2567.

[8] T.Romen Singh, Sudipta Roy, O.Imocha Singh, Tejmani Sinam, Kh.Manglem Singh, "A New Local Adaptive Thresholding Technique in Binarization", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 2, November 2011,pp.271-277.

[9] Jyotsna1 , Shivani Chauhan2 , Ekta Sharma3 , Amit Doegar4, "Binarization Techniques for Degraded Document Images - A Review". 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), 978-1-5090-1489-7/16/$31.00 ©2016 IEEE.

[10] Mehmet Yasin AKPINAR , Erdem EMEKL˙IG˙IL, Seçil ARSLAN, "Extracting Table Data from Images Using Optical Character Recognition Text", 978-1-5386-1501-0/18/$31.00 c 2018 IEEE.

[11] Harneet Singh, Anmol Sachan, "A proposed approach for character recognition using Document Analysis with OCR", Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018), 978-1-5386-2842-3/18/$31.00 ©2018 IEEE.

[12] N Prameela1 , P Anjusha, R Karthik, "Off-line Telugu Handwritten Characters Recognition using optical character recognition" , International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.

[13] Thulasi Kishna N.P, Seenia Francis, "Intelligent Tool For Malayalam Cursive Handwritten Character Recognition Using Artificial Neural Network And Hidden Markov Model" , Proceedings of the International Conference on Inventive Computing and Informatics (ICICI 2017) IEEE Xplore Compliant - Part Number: CFP17L34-ART, ISBN: 978-1-5386-4031-9

[14] Michael D. Kim∗ and Jun Ueda, "Dynamics-based motion de-blurring improves the performance of optical character recognition during fast scanning of a robotic eye" , 1083-4435 (c) 2017 IEEE.

[15] Hui Zho, GuoJun Zhu, Yunfeng Peng, "A RMB Optical Character Recognition System Using FPGA", 978-1-5090-2377-6/16/$31.00 ©2016 IEEE.