ENHANCEMENT OF DEGRADED HISTORICAL DOCUMENT IMAGES FOR BINARIZATION

Yogish Naik G.R[1] , *Shashidhara B[2] , Vidyasagar K.B[3]

[1] Associate professor, Department of PG studies and research in computer science, Kuvempu University Shimoga, Inida. [2] Research scholar, Department of PG studies and research in computer science, Kuvempu University Shimoga, Inida. [3] Research scholar, Department of PG studies and research in computer science, Kuvempu University Shimoga, Inida
*shashidhara.sirmv@gmail.com, [1]ynaik.ku@gmail.com, [3]vkb2231@gmail.com

ABSTRACT:
This work introduces a technique for enhancing deteriorated images of historical documents, overcoming noise, uneven illumination, and texture degradation through five steps: preprocessing, texture enhancement, illumination correction, binarization, and post-processing. In the preprocessing phase, Wiener filtering is employed to effectively reduce noise and enhance overall image quality. Subsequently CLAHE is used to improve local contrast and detail, while Local Binary Patterns (LBP) further enhance texture patterns, contributing to the overall enhancement of image quality. To mitigate the effects of uneven illumination, Retinex-based algorithms are utilized following texture enhancement, ensuring preservation of texture while correcting illumination variations across the image. The binarization step employs Sauvola's method, segmenting the image into foreground and background regions based on local pixel intensity variations. Post-processing techniques, including morphological methods, are then applied to eliminate minor artifacts and refine text shapes in the binarized image, thereby improving readability. Experimental testing on a series of DIBCO datasets confirms the suggested methodology's capacity to considerably improve the visual quality and readability of historical document images. The suggested method outperforms existing methods in terms of metrics such as F-measure, PSNR, NRM and DRD, emphasizing its potential for advancing research in document restoration and analysis.

***Keywords**: Binarization, Degraded Historical document images, Thresholding, Noise reduction, Contrast enhancement, Texture refinement.*

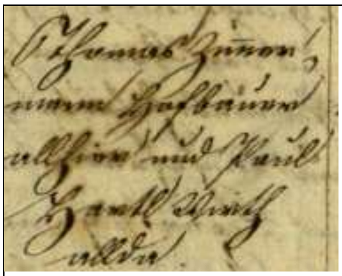
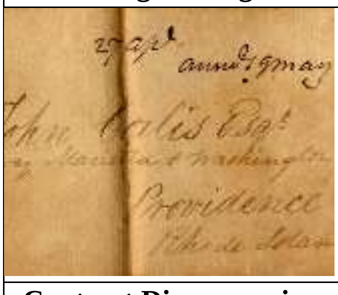
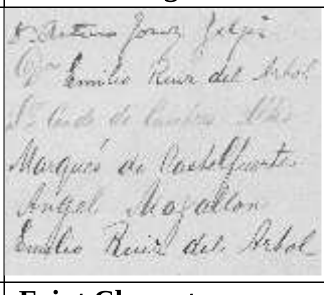
1. INTRODUCTION

This research presents a novel methodology for binarizing degraded document images, addressing the challenge of accurately separating foreground objects from background. The methodology is evaluated across various datasets and compared against ground truth images. The study aims to demonstrate the efficacy and robustness of the proposed technique, highlighting its strengths and potential areas for improvement. The comparison with ground truth images serves as a benchmark for evaluation.

Binarization converts grayscale images into binary ones, separating foreground and background objects for image analysis tasks like object detection and segmentation. Thresholding divides images into foreground and background regions based on pixel intensity values, comparing intensity values to a predefined threshold. Thresholding methods include global, local, adaptive, which apply a single threshold value to an entire image, compute threshold values for different regions, and adjust threshold values based on local characteristics **[11]**.

Historical documents are valuable artifacts, but degradation from time can threaten their preservation. Restoration requires sophisticated techniques to address issues like noise, contrast, texture, illumination, and binarization demands. Restoring degraded historical document images is a complex task that involves removing noise artifacts, enhancing contrast, restoring texture degradation, correcting uneven illumination, addressing handwriting variability, addressing ink fading and smudging, ensuring accurate reconstruction of crossings and overwriting, and accounting for irregularities in stroke thickness. These challenges are critical for maintaining the integrity of the document content, readability, and analysis. Restoring these documents requires a balance between preserving the original content and ensuring the legibility and authenticity of the original document **[16]**.

The impetus for this study arises from the importance of historical documents as valuable cultural assets and stores of communal knowledge. Handwritten papers, especially, provide distinctive perspectives on previous cultures, events, and personalities, rendering them priceless assets for historians, scholars, and the general public. Nevertheless, these papers frequently experience degradation over time, such as ink fading, paper deterioration, and other types of damage, which can compromise their readability and durability.

		
Bleeding Through Ink	**Texture Degradation**	**Thin Strokes**
		
Contrast Discrepancies	**Faint Characters**	**Ink Fading and Smudging**

Handwriting Variability	Varying Font Size	Noise Artifacts

Figure 1: displays examples of degraded historical document images.

This paper presents a comprehensive methodology for restoring Images of historical documents that have deteriorated or become damaged over time. using advanced techniques in computation of images. It includes noise reduction, contrast enhancement, texture refinement, illumination correction, binarization, and post-processing to ensure accessibility and legibility for future generations. Historical document analysis requires effective thresholding and binarization techniques to enhance legibility and extract valuable content from documents with degradation. These techniques isolate text from background noise, facilitating OCR and content analysis **[11][16]**.

This study explores binarization techniques for improving degraded historical document images, focusing on Wiener Filtering, CLAHE, LBP, Retinex-based techniques, Sauvola's Method, and morphological operations. Our methodology evaluated using the DIBCO series datasets, significantly enhances reconstructing deteriorated historical document images, facilitating archival, digitization, and analysis, and outperforms existing methodologies in various parameters **[17][19][20][32][33][34]**.

The structure of the work is as follows: Section 2 covers pertinent investigation in the domain of processing of historical document images. The method and layout of the suggested workflow for image processing are explained in Section 3. The outcomes of the experiment and the evaluation of performance are presented in Section 4. Section 5 provides concluding observations and a discussion of potential future directions to wrap up the article.

2. REVIEW OF LITERATURE

This literature review examines prior studies, methodologies, and advancements in document image binarization techniques. It aims to identify trends, challenges, and gaps, and provides insights into various methodologies, algorithms, and evaluation metrics. The review aims to establish a solid foundation for our proposed approach.

N. Otsu **[1]** in 1979, introduced a widely used automatic image thresholding strategy that Seeks to ascertain the most effective threshold value for dividing an image into distinct foreground and background regions by minimizing variance differences between foreground and background pixels.

W. Niblack **[2]** in 1986, introduced a local thresholding technique used for image binarization. It uses mean and standard deviation of pixel values to generate a threshold for each pixel, this technique effectively handling illumination and background intensity fluctuations in an image.

J. Bernsen's **[3]** in 1986, has proposed work on dynamic thresholding of gray-level images is a significant contribution for image processing. The work uses local thresholding for adaptive binarization, adjusting threshold values based on local image properties. This method is useful in document processing when global thresholding techniques cannot handle illumination, contrast or background noise.

Sauvola. J et al. **[4]** in 2000, has introduced a method on local thresholding technique used for image binarization and document analysis. It generates a pixel-specific local threshold calculated using the mean and standard deviation, the method demonstrating its effectiveness in handling illumination and background intensity fluctuations.

Wolf et al. **[5]** in 2002, has proposed binarization method is an adaptive thresholding technique used to convert grayscale images into binary images. The method dynamically adjusts the threshold based on local image features, making it useful for images with uneven illumination.

Kim et al. **[6]** in 2002, has introduced a method for document image binarization, aiming to accurately differentiate between foreground and background text. They used adaptive thresholding techniques to deal with illumination, noise, and background features, dynamically modifying threshold values to effectively binarize the image.

Gato et al. **[7]** in 2006, introduced a novel method for document binarization, transforming grayscale images into binary representations for easier text extraction and analysis. They combined local thresholding techniques with adaptive approaches, utilizing inherent properties of document images like text intensity distribution and local pixel patterns. Pretreatment stages improve image quality and performance, including noise reduction, contrast enhancement, and edge detection.

Su. B et al. **[8]** in 2010, has proposed an innovative method for binarizing documentation images, aiming to enhance optical character recognition and document analysis. They used a combination of global and local thresholding algorithms to dynamically adjust threshold settings, assessing local image properties like pixel intensities and gradients.

Gangamma et al. **[10]** in 2011, has proposed a technique to improve the quality of degraded Kannada documents. They propose image enhancing algorithms for Kannada script, addressing common issues like aging artifacts and uneven illumination. The approach aims to restore readability and make the records more accessible for analysis and preservation.

Ranganath et al. **[18]** in 2015, has proposed a method which combines various binarization approaches to address challenges like uneven illumination, noise, and aging artifacts in degraded document images. The hybrid approach can adaptively combine thresholding algorithms, maintaining significant information while minimizing artifacts.

Parashuram B et al. **[24]** in 2018, has proposed a Technique endeavors to enhance the caliber of deteriorated Kannada handwritten image documents through utilizing image enhancement techniques like noise reduction, contrast enhancement, and illumination correction. The authors also propose techniques specifically designed for the recognition and restoration of Kannada language script.

Xiong W et al. **[25]** in 2018 proposed a comprehensive approach for binarizing degraded historical document images, incorporating both preprocessing techniques and machine learning-based classification methods to achieve accurate results. The combination of local feature extraction, SVM classification, and adaptive thresholding demonstrates a systematic and a successful strategy for overcoming the difficulties involved with document binarization in the presence of degradation.

3. PROPOSED METHODOLOGY:

The proposed methodology entails a systematic amalgamation of image processing techniques. These encompass noise reduction, contrast enhancement, texture refinement, illumination correction, binarization, and morphological operations. Every facet of this process will be intricately crafted and fine-tuned to effectively mitigate prevalent degradation artifacts, all the while ensuring the preservation of the historical content's authenticity and integrity.

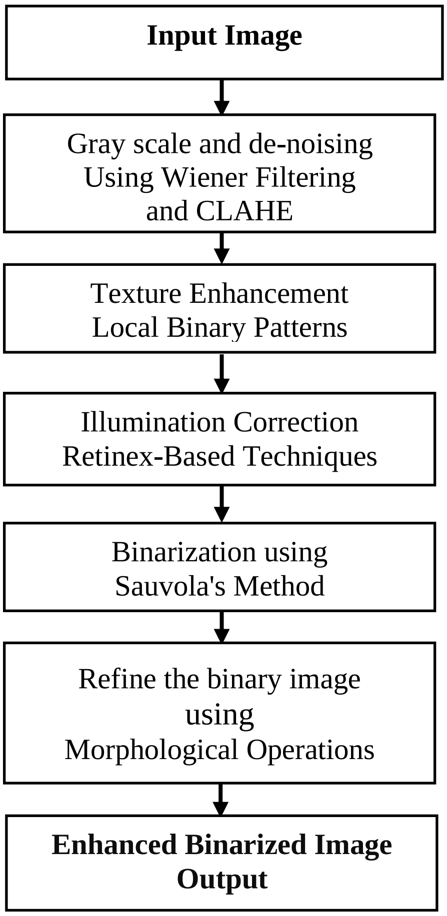


Figure 3: Flowchart Illustrating the Proposed Methodology

The proposed preprocessing methodology integrates Wiener Filtering and CLAHE as initial steps, strategically positioned before any other enhancement or correction techniques. These meticulously chosen processes are aimed at elevating the input image's quality and highlighting its intrinsic features. Subsequently, during the post-

processing phase, Morphological Operations come into play, executed following binarization. This pivotal step serves to refine the binary image further, focusing on tasks like noise reduction, gap filling, and enhancing structural features. Ultimately, these operations contribute significantly to optimizing the binary image's integrity and clarity, ensuring its suitability for subsequent analysis and interpretation **[11][16].**

3.1 Preprocessing

Preprocessing is vital for enhancing degraded historical document images, preparing them for subsequent enhancement steps. Converting these images to grayscale simplifies the processing workflow and preserves essential textual information. Initially, historical document images may be degraded due to factors like aging, fading, or damage, existing in either color or grayscale format. If in color, they are converted to grayscale, transforming them from a color space (e.g., RGB) to a single-channel representation where pixel values denote light intensity, ranging from black to white typically in an 8-bit grayscale image (0 for black, 255 for white) **[11].**

3.1.1 Noise reduction

Wiener filtering is a technique utilized specifically for image noise reduction. It operates by estimating the original signal within the image, striving in order to reduce the mean square error between this original and the resulting filtered image.

Steps in wiener filtering:

1. Local Block Division:

   For every pixel (x, y) in the source image and define a 5x5 local block centered at (x, y).

2. Calculating Mean and Variance for each local block B_{xy} and Calculate the mean μ and variance σ^2 using the formulas:

   $\mu = \frac{1}{25}$ * sum of pixel intensities in B_xy

   $\sigma^2 = \frac{1}{25}$ * sum of squared differences from μ for each pixel in B_{xy}

3. Average Variance of the Original Image compute the average variance Avg(σ) across all local blocks.

   $Avg(\sigma) = \frac{1}{M}\sum_{i=1}^{M} \sigma_i$, where M represents all of the local blocks.

4. Wiener Filter Transformation for each pixel (x, y) in the original image and get the mean μ and variance σ^2 of the local block centered at (x, y).

   Apply the Wiener filter transformation to obtain the filtered pixel value $\mathbf{I}_{\mathbf{wiener}}$ (x, y):

$$I_{wiener}(x,y) = \mu + \left(\frac{\sigma^2}{\sigma^2 + Avg(\sigma)}\right) * (G_o(x,y) - \mu) \qquad \text{----(1)}$$

   where I_{wiener}(x, y) represents the pixel-level grayscale value of the original image.

4. For every pixel (x, y) in the original image, repeat steps 2-4 to get the filtered image.

3.1.2 To enhancing the contrast of an image

CLAHE is a technique that enhances image contrast by dividing the image into smaller tiles and applying histogram equalization independently, preventing over-amplification of noise.

Steps involved in CLAHE:

1. Divide the input image into non-overlapping tiles.
   tile T_{ij} represents a sub-image region defined by the i-th row and j-th column of tiles.
2. Compute the histogram of pixel intensities within each tile.
   For each tile T_{ij}, compute the histogram H_{ij} of pixel intensities. The histogram should span the entire intensity range (e.g., 0 to 255 for an 8-bit image) and count the number of occurrences of each intensity level within the tile.
3. Apply histogram equalization independently to each tile.
   For each tile T_{ij}, compute the cumulative distribution function (CDF) CDF_{ij} from the histogram H_{ij}

   $$CDF_{ij}(k) = \sum_{p=0}^{k} H_{ij}(p) \quad \text{------------(2)}$$

   Where: H_{ij} (p) is the histogram of pixel intensity p within the tile.
   The transformation function $T_{i,j}(k)$ for each tile is then computed as:

   $$T_{i,j}(k) = clip\left(\frac{CDF_{i,j}(k) - CDF_{min}}{N\ x\ M - CDF_{min}} * (L-1)\right) \quad \text{-----------(3)}$$

   Where: L is the number of intensity levels (typically 256 for an 8-bit image).
4. Clip the histogram to limit the contrast enhancement.
   After histogram equalization, clip the histogram of each tile to limit the contrast enhancement.

   clipping the histogram is:

   $$H_{clipped}(k) = \begin{cases} H(k) & if\ CDF(k) \leq L \\ \frac{H(k) - CDF(k)}{1-L} & if\ CDF(k) > L \end{cases} \quad \text{------------(4)}$$

   where:

   $H_{clipped}(k)$ represents the clipped count of pixels with intensity level k in the histogram. $H(k)$ is the count of pixels with intensity level k in the original histogram. $CDF(k)$ is the cumulative distribution function value for intensity level k. L is the specified limit for the cumulative distribution function (CDF).
5. Combine the equalized tiles to form the final enhanced image.

   $$I_{clahe}(x, y) = \sum_{i=1}^{n} H_k * T_{ij}(x, y) \quad \text{----------(5)}$$

3.2 To enhance texture patterns and image quality

LBP is a texture descriptor in image processing that assigns binary codes to pixels based on their intensity. These codes are used to create a histogram, representing the distribution of patterns within a local neighborhood, where $I(x_c,y_c)$ represents the intensity value of the center pixel. The LBP code for the center pixel is computed as follows:

$$LBP(x_c, y_c) = \sum_{i=0}^{P-1} S(i). 2^i \quad \text{-----------(6)}$$

Where:

- P is the number of neighboring pixels considered.

S(i) is a binary function defined as:

$$S(i) = \begin{cases} 1, & if\ I(x_i, y_i) \geq I(x_c, y_c) \\ 0, & otherwise \end{cases} \quad \text{------------(7)}$$

using LBP for texture enhancement: Enhancement function f to the LBP image is given by,

$$I_{enh}\,(x, y) = f_{ad\_g}(I_{LBP}(x, y), L\_mean(x, y), L\_var(x, y)) \quad \text{----------(8)}$$

Where,

- f_{ad_g} is the adaptive Gaussian filtering function tailored for enhancing texture features in the LBP image. $I_{LBP}(x, y)$ represents the LBP image with LBP codes computed for each pixel. $I_{enh}(x, y)$ denotes the enhanced image obtained by applying the enhancement function f to the LBP image.

3.3 To adjust uneven illumination across the image

Retinex-based illumination correction removes uneven illumination while preserving texture information by separating an image into illumination and reflectance components using a logarithmic operation and convolution process.

Steps involve in Retinex-based illumination correction:

1. Logarithmic Operation:
   The input image I(x,y) is transformed using a logarithmic function to improve the dynamic range and make it easier to separate illumination and reflectance components. This step can be represented as:

   $$Log\ (I_{Log}(x, y)) = log\ (I(x, y)) \quad \text{---------------(9)}$$

2. Convolution Operation:
   The log-transformed image $I_{log}(x,y)$ is convolved with a Gaussian filter to estimate the illumination component. The convolution operation can be represented as:

   $$I_{illumination}(x, y) = I_{log}(x, y) * G(x, y) \quad \text{----------------(10)}$$

   Where:

   - $I_{illumination}(x,y)$ represents the estimated illumination component of the image. denotes the convolution operation. G(x,y) is the Gaussian filter kernel.

   The Gaussian filter kernel G(x,y) is defined as:

   $$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \text{----------------(11)}$$

   Where:

   - The spatial coordinates relative to the kernel center are indicated by x and y.
   - The amount of smoothing is determined by the standard deviation (σ) of the Gaussian distribution.
   - A larger σ value results in more smoothing, while a smaller σ value preserves finer features.
   - The base of the natural logarithm, denoted by e, is approximately equal to 2.71828.

3. Reflectance Estimation:
   Once the illumination component is estimated, the reflectance component can be obtained by dividing the original image by the estimated illumination:

$$I_{reflectance}(x,y) = \frac{I(x,y)}{I_{illumination}(x,y)}$$ --------------(12)

4. Correction and Enhancement:
   The illumination-corrected image is obtained by multiplying the estimated reflectance with the original image:

$$I_{corrected}(x,y) = I_{reflectance}(x,y) * I(x,y)$$ ---------------(13)

3.4 Binarization

Retinex-based illumination correction is used for binarization, segmenting images into foreground and background regions. Sauvola's method [4] computes a local threshold for individual pixels, dividing intensities above the threshold into foreground and background regions. Pixels with higher intensities are set to foreground, while those below is background.

Threshold calculation for each pixel is given by

$$T(x,y) = \mu(x,y) * (1 + k * \frac{\sigma(x,y)}{R} - 1)$$ ------------(14)

Where:

- T(x,y) is the threshold value for the pixel at coordinates (x,y). μ(x,y) is the mean intensity of the pixel's neighborhood. σ(x,y) is the standard deviation of the pixel's neighborhood. k is a user-defined constant (typically between 0.2 and 0.5) that controls the sensitivity to local variations. R – max σ of the image.

formula for obtaining the final binary image:

$$B(x,y) = \begin{cases} 1 & if\ I(x,y) > T(x,y) \\ 0 & otherwise \end{cases}$$ --------------(15)

Where:

- B(x, y) signifies the binary value at the coordinates (x, y) within the resultant binary image.
- I(x, y) represents the intensity value of the pixel located at (x, y) within the grayscale image.
- T(x, y) indicates the threshold value assigned to the pixel positioned at (x, y).

3.5 To refine the binary image

Morphological operations are image processing techniques that analyze and manipulate images based on object shape and structure. They eliminate noise, detect edges, and extract features. Experimentation with different morphological operations and parameters can enhance binarized images. Combining multiple operations or applying them iteratively can yield better results. Opening, a combination of erosion and dilation, removes small-scale features while maintaining larger text shape.

In the erosion process, a structuring element (kernel) traverses the input image. At each kernel position, if all pixels underneath in the input image are white (255) or nonzero, the central pixel of the kernel remains white; otherwise, it becomes black (0). This process contracts or erodes the boundaries of foreground text in the image.

$$(A \ominus B)(x) = \begin{cases} 1, & if\ (B \oplus x) \subseteq A \\ 0, & otherwise \end{cases} \quad \text{---------(16)}$$

Dilation involves a scanning process similar to erosion, where if at least one pixel under the kernel in the input image is white, the center pixel of the kernel turns white; otherwise, it remains black. This expands or dilates the boundaries of foreground text in the image.

$$(A \oplus B)(x) = \begin{cases} 1, & (B \oplus x) \cap A \\ 0, & otherwise \end{cases} \quad \text{--------------------(17)}$$

Opening is a morphological image processing operation involving erosion followed by dilation. It's beneficial for eliminating small elements, noise, or slender structures from the foreground, while retaining larger structures and the overall text shape. The opening operation $(A \ominus B) \oplus B$ is defined as:

$$(A \ominus B) \oplus B = \{x \mid (B \oplus x) \subseteq A\} \quad \text{----------------(18)}$$

where:

- A is the input binary image, B is the structuring element (kernel), $\ominus$ denotes erosion, $\oplus$ denotes dilation, x represents the translation of the kernel.

4. EXPERIMENTAL FINDINGS AND PERFORMANCE EVALUATION

The experiment aims to develop and assess a comprehensive workflow tailored for enhancing degraded historical document images and accurately binarizing them. It addresses challenges associated with such images by proposing a systematic approach that combines preprocessing, texture enhancement, illumination correction, and advanced binarization techniques. Utilizing datasets like DIBCO 2011, DIBCO 2013, DIBCO 2016, and DIBCO 2018, which contain various types of deterioration in both handwritten and printed materials, the experiment ensures a robust evaluation. By comparing against other algorithms using these datasets, both qualitative and quantitative assessments are conducted to validate the effectiveness of the proposed method **[9][13][21][23].**

F-measure: The restoration images' completeness and accuracy in relation to the original images is gauged by the F-measure. It combines recall (the proportion of relevant examples that have been retrieved over the entire number of relevant instances) with accuracy (the fraction of relevant instances among the retrieved instances). The harmonic mean of precision and recall is known as the F-measure.

F-m$= 2 * \frac{(Rc * Pr)}{(Rc + Pr)}$ $\qquad$ $\boldsymbol{Rc} = \frac{TP}{TP+FN}$ $\qquad$ $\boldsymbol{Pr} = \frac{TP}{TP+FP}$ ---(19)

Where,

Precision(Pr), Recall(Rc), True Positive (TP), False Positive (FP) and False Negative (FN).

One often used statistic to evaluate the quality of a recovered image to the original image is the PSNR (Peak Signal-to-Noise Ratio). It calculates the ratio of a signal's full potential power to the power of noise that tampers with the representation of the signal, reducing its accuracy. Better image quality is indicated by greater PSNR values, which are measured in decibels (dB).

NRM (Noise-to-Reconstruction Ratio): NRM quantifies the ratio of noise present in the restored image to the amount of noise present in the original image. It helps assess the effectiveness of noise reduction techniques in the restoration process, with lower NRM values indicating better noise reduction.

$$\boldsymbol{PSNR} = 10 * Log_{10}\left(\frac{MAX^2}{MSE}\right) \quad \text{---(20)} \qquad \boldsymbol{NRM} = \frac{Mean\ Squared\ Error\ (MSE)}{Mean\ Squared\ Signal\ (MSS)} \text{---(21)}$$

$$\boldsymbol{MSE} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(I_{ori}(i,j) - I_{pro}(i,j))^2}{M\ x\ N} \qquad \boldsymbol{MSS} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ori}(i,j)^2}{M\ x\ N}$$

Where:

- MSE is the Mean Squared Error,
- Iori(i,j) is the pixel value of the original image at position (i,j),
- Ipro(i,j) is the pixel value of the improved or reconstructed image at position (i,j),
- M and N are the dimensions of the images.

DRD (Distance Reciprocal Distortion): DRD compare the difference between the original and restored images, The DRD formula calculates the distortion of each pixel and then averages them. It can be expressed as:

$$DRD = \frac{1}{NUBN}\sum_{k=1}^{NUBN} DRD_k \qquad \text{----------(22)}$$

Where:

DRD_k represents the distortion of the k-th pixel. NUBN is the total number of non-uniform pixels in the window size.

The distortion of each pixel, denoted as DRD_k, is determined through the utilization of a weighted matrix. This value can be computed as follows:

$$\boldsymbol{DRD_k} = \sum_{i=-2}^{2}\sum_{j=-2}^{2}|\boldsymbol{GT_k(i,j)} - \boldsymbol{B_k(x,y)}| \times \boldsymbol{W_{Nm}(i,j)} \qquad \textbf{---(23)}$$

Where:

GT_k (i,j) denotes the pixel value of the reference image at the position (i,j). B_k (x,y) denotes the pixel value of the enhanced image at the position (x,y). W_Nm (i,j) stands for the weighted matrix.

These metrics provide quantitative measures of image quality and are useful for comparing different restoration algorithms or parameter settings. They help evaluate how well a restoration algorithm preserves important image features and minimizes artifacts or distortions. Improved performance is shown by lower values for NRM and DRD, but by higher values for F-measure and PSNR. The objective is to optimize these metrics according to the images restoration task's specific requirements.

Experimental Setup and process:

We utilized a dataset consisting of deteriorated historical document images acquired from different DIBCO databases. This collection showcases documents that display various types of degradation, such as fading, stains, ink seepage, and paper deterioration.

In preprocessing, Wiener Filtering is a noise reduction technique applied to enhance the quality of input images. Equation (1) computes a weighted average between the original pixel value and the mean intensity value of the local block, adjusting the filtering degree based on local image characteristics and overall texture variability. Experimental results validate the effectiveness of Wiener filtering in enhancing degraded historical document images, reducing noise and enhancing clarity. A kernel size of 5x5 is typically used, and noise variance is empirically determined based on image characteristics and noise type. CLAHE is utilized to enhance local contrast and detail visibility in images. Equation (2) measures the distribution of pixel intensities within each tile, guiding subsequent histogram equalization steps. Equation (3) adjusts pixel values based on the cumulative distribution function, while equation (4) performs clipping to balance contrast enhancement and noise suppression, ensuring improved image quality without losing important details. Equation (5) integrates enhanced tiles to produce the final image, benefiting from CLAHE's application. CLAHE employs an 8x8 tile size to balance local contrast enhancement and noise amplification, with a clip limit of 0.03 controlling contrast enhancement extent.

For texture enhancement, Local Binary Patterns (LBP) algorithms are employed to improve texture details and enhance text and patterns in images. Equation (6) computes the LBP code, serving as a texture descriptor by capturing local texture patterns around each pixel. Equation (7) utilizes a binary function to encode texture patterns based on intensity comparisons between the center pixel and its neighbors. Equation (8) enhances texture features while preserving image details by combining information from the LBP image, local mean, and variance. Adaptive Gaussian filtering adjusts pixel values based on local characteristics, resulting in improved texture-enhanced images. LBP is configured with 8 points and a radius of 1 to effectively capture texture information in historical document images.

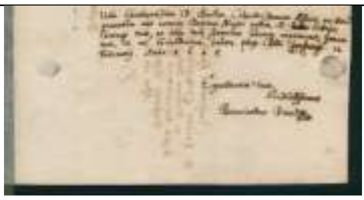
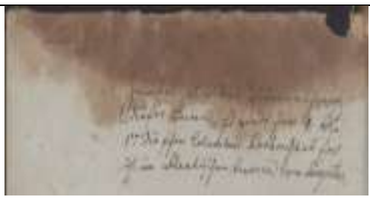
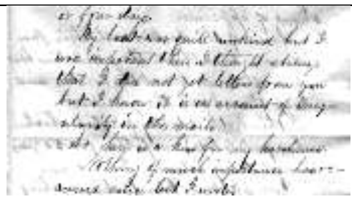
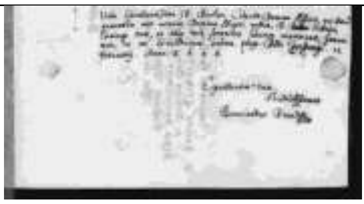
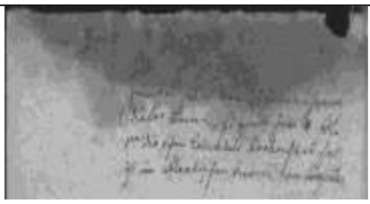
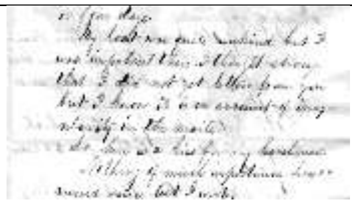
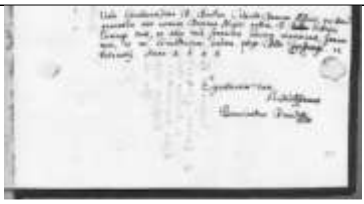
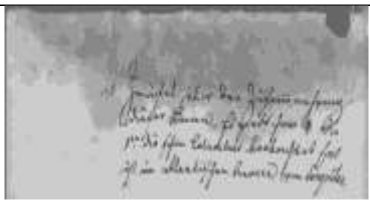
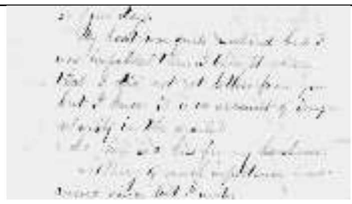
For uneven illumination, Retinex-based techniques are employed to rectify uneven lighting and enhance overall brightness and contrast in images. Equation (9) involves a logarithmic operation that enhances the image's dynamic range, improving visibility of details and textures by compressing high-intensity values and expanding low-intensity values. Equation (10) applies a Gaussian filter kernel through convolution to smooth the log-transformed image and estimate the illumination component, effectively removing uneven lighting effects while preserving texture information. Equation (12) separates the image into intrinsic reflectance and illumination components, enabling selective correction and enhancement. Equation (13) completes the Retinex-based illumination correction process, resulting in an image that removes uneven illumination while preserving texture details, enhancing visual quality. Multi-scale Retinex algorithm is utilized for its ability to enhance both local and global contrast while preserving important details.

To achieve binarization, Sauvola's adaptive binarization method is employed to distinguish the foreground (text) from the background in the enhanced images. Equation (14) determines the local threshold using the mean and standard deviation of the pixel's neighborhood, adjusted by a factor involving the constant k and the relative standard deviation. Equation (15) generates the final binary image by comparing pixel intensities of the original grayscale image against the calculated local threshold, effectively segmenting the input image into foreground and background areas. Each pixel is assigned a binary value of 1 or 0, denoting foreground and background respectively, based on a binary decision rule. Sauvola's method uses a 15x15 window size to accommodate typical text regions in historical documents, ensuring robustness against variations in layout and text size. The k-value of -0.2 dynamically adjusts the threshold according to local image characteristics, influencing the binarization outcome.

To refine the binarized image, we employed Morphological operations, including erosion, dilation, and opening, are employed for specific image processing tasks. Equation (16) illustrates erosion's role in shrinking foreground object boundaries by assigning values based on the fit of the structuring element within the set. Conversely, equation (17) depicts dilation's function in expanding foreground object boundaries by evaluating overlap with the set. Equation (18) outlines the outcome of opening, effectively removing small objects, noise, or thin structures while preserving larger ones. Dilation fills gaps between objects, enhancing their thickness. Morphological operations utilize a 3x3 kernel size for opening, with one iteration to remove noise and refine the binarized image.

Each technique is systematically applied in sequence to the input degraded historical document images, where the output of each step becomes the input for the next. This methodical approach ensures a progressive enhancement and accurate binarization of the images, ultimately leading to improved readability and analysis of the historical documents. Experimental observations are derived from two sets of experiments: one relying on visual perception and the other on evaluation using established measures such as F-measure, PSNR, and NRM. A performance comparison is conducted between our methodology and four widely recognized binarization approaches, including Otsu's**[1],** Niblack's**[2],** Gato's**[7],** Bernsen's**[3]** and wolf's**[5]** techniques.

Qualitative Results -Visual representation of the experiment conducted:

	Image-1 of DIBCO-2018	**Image-2 of DIBCO-2018**	**Image-HW06 of DIBCO-2013**
a)			
b)			
c)			

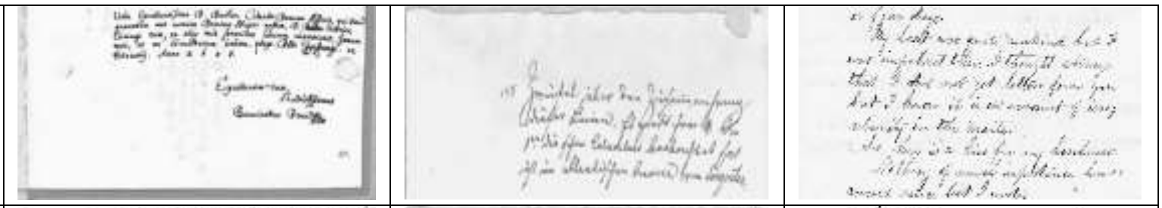
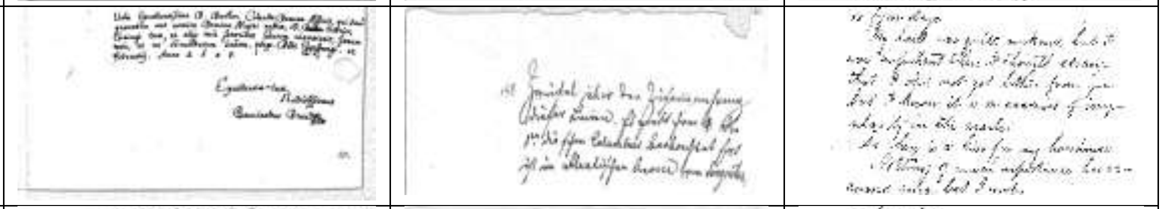
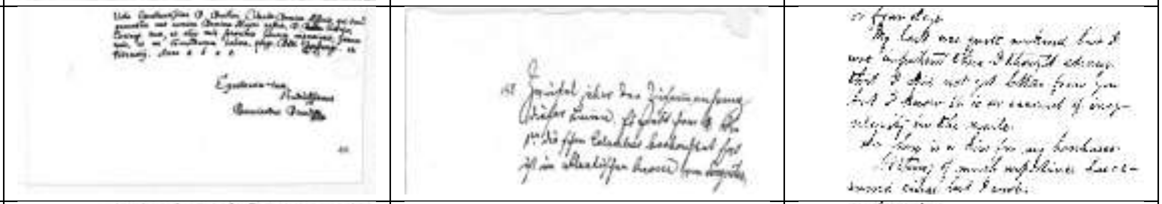
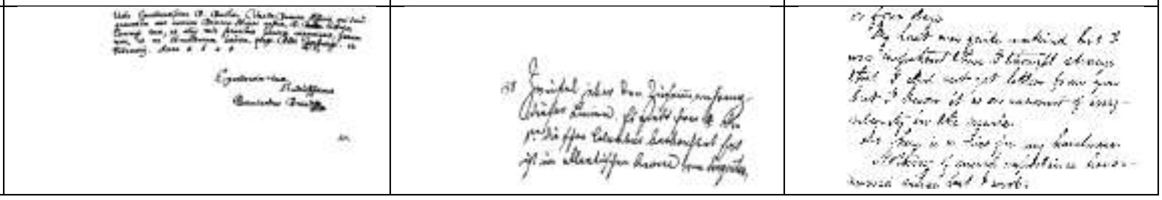
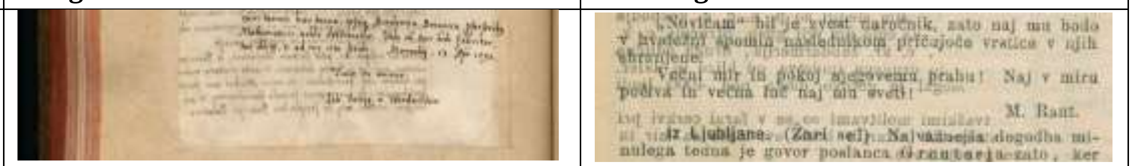
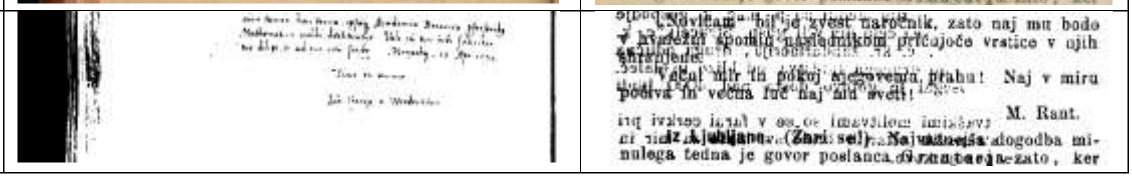
d)	
e)	
f)	
g)	

Figure 3: illustrates the results of proposed method in each stages: a) The input image. b) The output after preprocessing. c) The image after texture enhancement. d) The result of illumination correction. e) The outcome of Sauvola. f) The final result after post-processing using morphological opening operation. g) ground truth image.

In the domain of image binarization, numerous methods have emerged to automate the selection of thresholds, with the goal of precisely distinguishing foreground elements from the background. In this study, we assess the efficacy of our proposed approach against five established techniques commonly utilized in this domain. Each of these methods offers distinct benefits and considerations, tailored to address diverse demands and contexts encountered in image binarization endeavors.

	Image - 4 of DIBCO 2018 dataset / **Image –PR06 of DIBCO 2013 dataset**
a)	
b)	

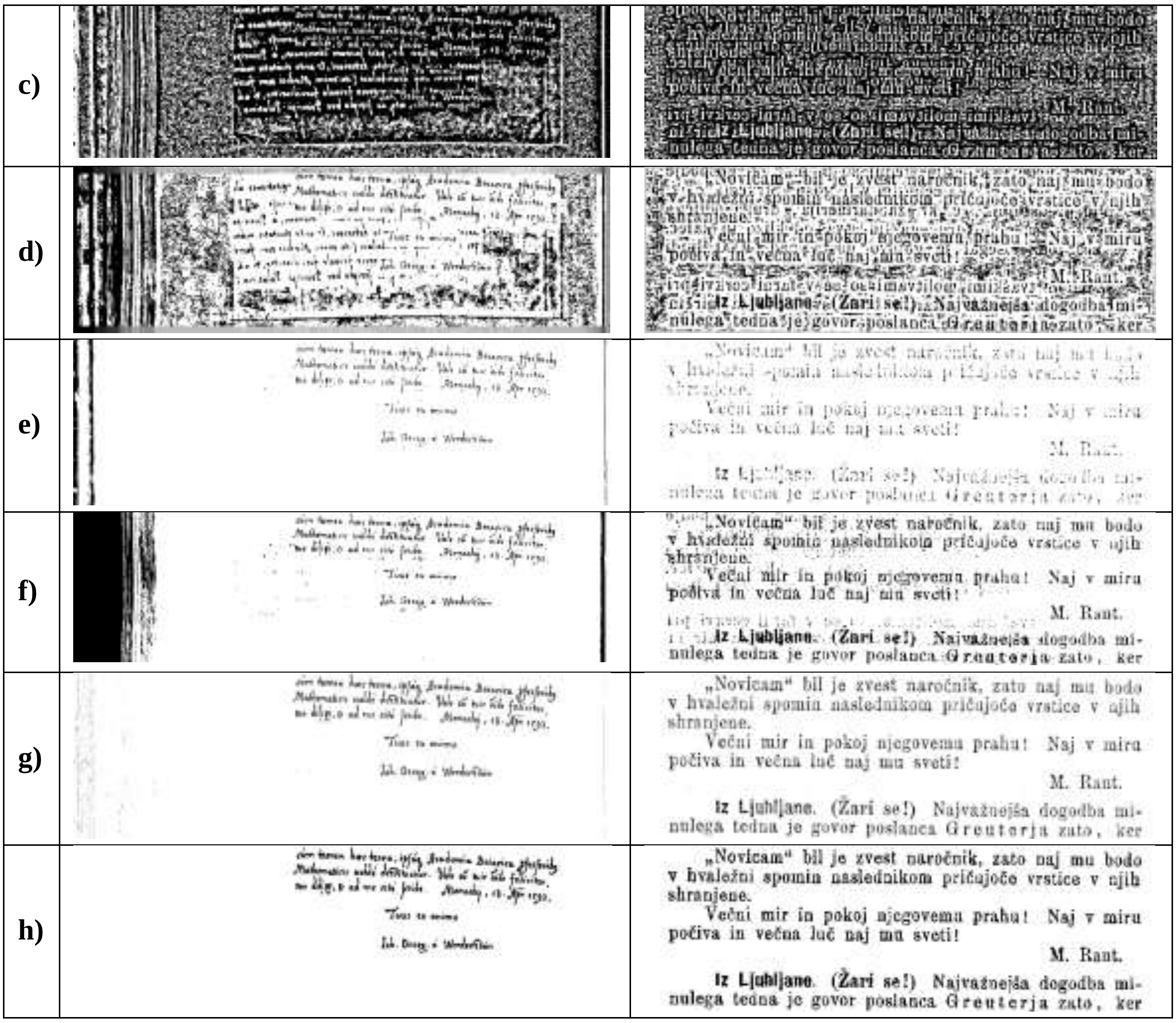


Figure 4: Evaluation of results from our proposed method compared to other methods: a) Input image, b) Otsu's result, c) Niblack's result, d) Bernsen's result, e) Wolf's result, f) Gatos's result, g) Proposed method's result, h) Ground truth image.

Table 1: Metrics Comparison for Image Enhancement Techniques with proposed approach

Images	Metrics	Otsu [1]	Niblack [2]	Bernsen [3]	Wolf [5]	Gatos [7]	Proposed
Image - 4 of DIBCO 2018 dataset	**F-measure**	71.57	31.54	37.29	74.17	68.56	**89.24**
	PSNR	14.52	04.85	06.87	16.26	12.21	**18.62**
	NRM($\times 10^{-2}$)	08.04	18.45	14.58	07.25	10.65	**05.04**
	DRD	13.54	112.26	98.58	12.24	15.14	**3.756**
Image –PR06 of DIBCO 2013 dataset	**F-measure**	69.36	27.15	31.74	71.55	67.86	**91.28**
	PSNR	11.28	03.28	05.86	14.32	11.25	**18.54**
	NRM($\times 10^{-2}$)	11.45	19.68	16.14	06.28	11.27	**05.28**

DRD	17.56	115.85	106.28	12.34	16.84	**3.985**

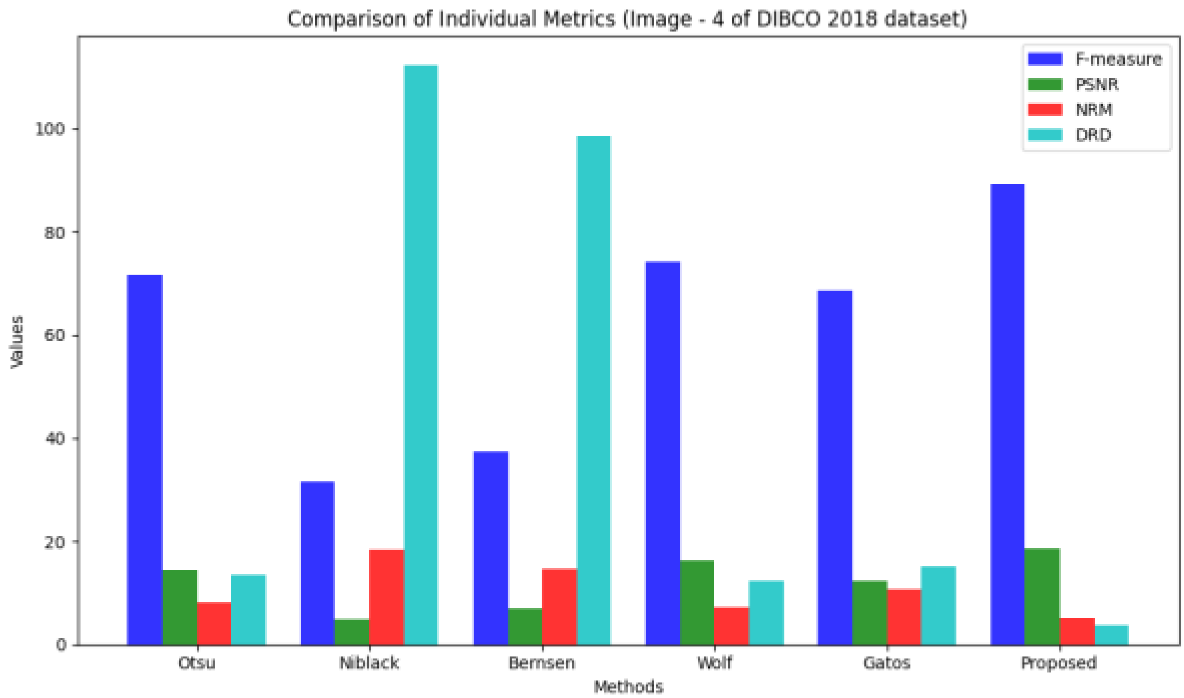


Figure 5: graph plot for Table-1 of Image - 4 of DIBCO 2018 dataset (figure 4)

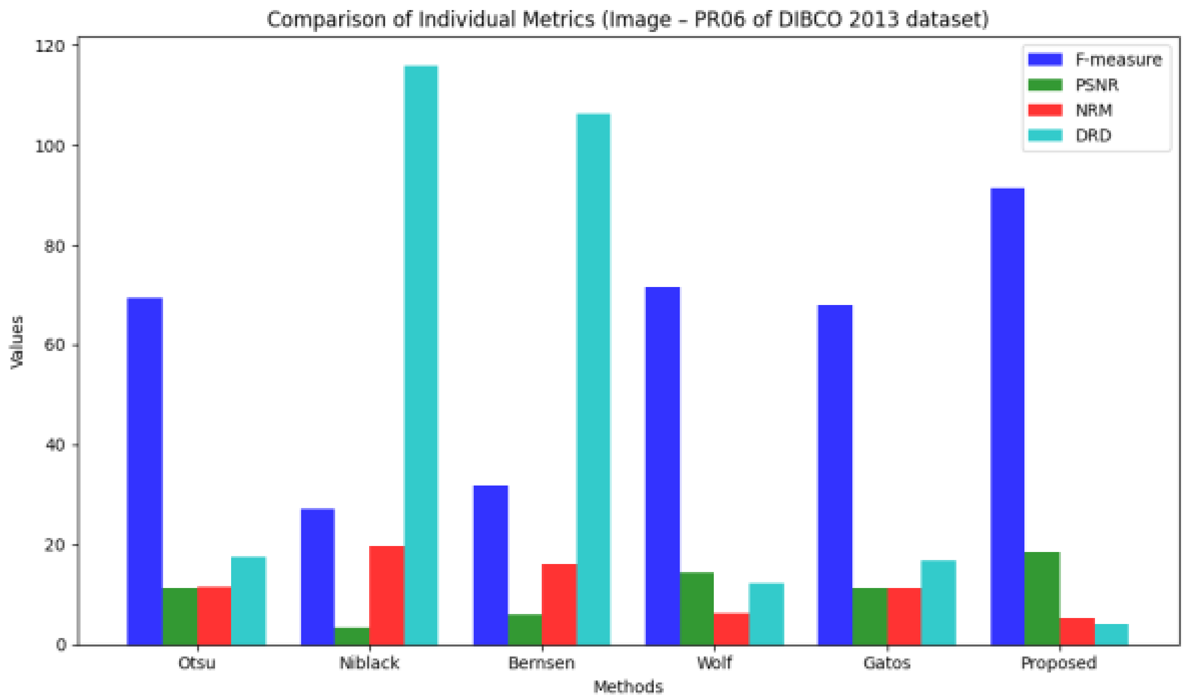


Figure 6: graph plot for Table-1 of Image – PR06 of DIBCO 2013 dataset (figure 4)

Results and Discussion for the result of table-1:

The table provided compares five popular image binarization methods across different evaluation metrics for two different images from the DIBCO datasets (DIBCO 2018 dataset and DIBCO 2013 dataset). The proposed method's performance is highlighted for analysis of its efficiency compared to other methods.

Image - 4 of DIBCO 2018 Dataset:

- F-measure: The proposed method achieved the highest F-measure of 89.24, indicating superior performance in terms of precision and recall compared to other methods.
- PSNR: The proposed method also attained the highest PSNR value of 18.62, suggesting better preservation of image quality during binarization.
- NRM: The proposed method recorded the lowest NRM value of 05.04, implying effective noise reduction compared to other methods.
- DRD: The proposed method recorded a DRD value of 3.756, significantly lower than other methods, indicating better preservation of image distortion.

Image – PR06 of DIBCO 2013 Dataset:

- F-measure: The proposed method outperformed other methods with the highest F-measure of 91.28, showcasing its consistency across different datasets.
- PSNR: The proposed method achieved the highest PSNR value of 18.54, indicating superior image quality preservation.
- NRM: The proposed method recorded the lowest NRM value of 05.28, suggesting effective noise reduction.
- DRD: The proposed method recorded a DRD value of 3.985, significantly lower than other methods, indicating better preservation of image distortion.

The proposed method consistently outperformed other methods in all evaluation metrics for both datasets, demonstrating its effectiveness in various scenarios. It achieves a balance between precision and recall, preserves image quality, reduces noise, and preserves image distortion.

we present a comprehensive evaluation of our proposed methodology across a diverse range of DIBCO series of dataset images. These images were meticulously processed using our proposed approach and subsequently compared against their respective ground truth images for thorough assessment. Through this comparative analysis, we aim to demonstrate the efficacy and robustness of our proposed technique in accurately binarizing degraded document images. Each dataset image underwent meticulous processing steps as outlined in our methodology, ensuring consistency and reliability in the evaluation process. The comparison with ground truth images serves as a benchmark for quantitatively and qualitatively assessing the performance of our proposed method, shedding light on its strengths and potential areas for improvement.

Original images of DIBCO 2018 dataset [23]	**Binarized images obtained from our proposed method**	**Ground truth image**

Figure 7: proposed method result obtained for DIBCO 2018 dataset images

Original images of DIBCO 2016 dataset [21]	Binarized images obtained from our proposed method	Ground truth image

Figure 8: proposed method result obtained for DIBCO 2016 dataset images

Original images of DIBCO 2013 dataset [13]	Binarized images obtained from our proposed method	Ground truth image

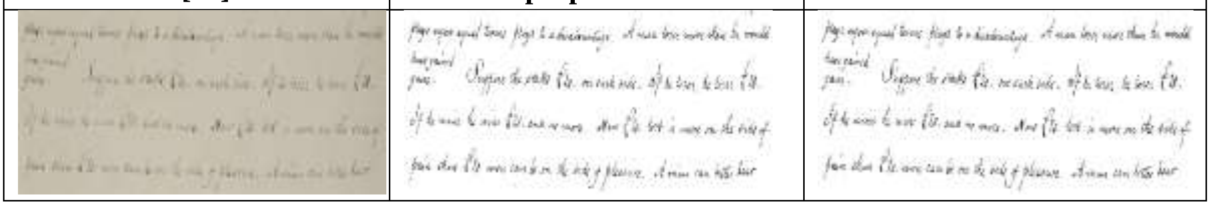


Figure 9: proposed method result obtained for DIBCO 2013 dataset images

Original images of DIBCO 2011 dataset [9]	Binarized images obtained from our proposed method	Ground truth image


Receipt from Miss Butterfield
for 2 beds, desks & benches
March 24 - 1920.
27 apl. ans'd 9 may
John Corlis Esq.
Providence
Rhode Islan
will sagen: dem blind-geflügleten Kinde
verstehe auch dises nicht.
r Liebes Gottheit.
Man lese:
Dn. GUSTAVI ADOLPHI,
Svecorum, Gothorum & Vandalorum Regis, Ma-
Ecclesiae Defensoris
Et
Domini mei Clementissimi.

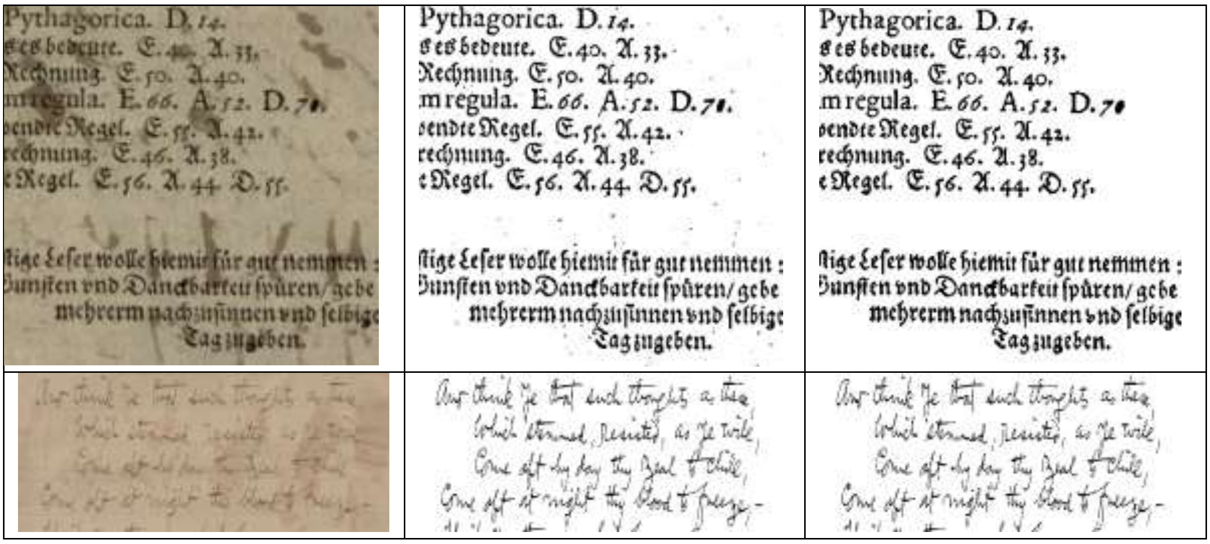


Figure 10: proposed method result obtained for DIBCO 2011 dataset images

Our research has shown the proposed method to be a superior choice over traditional methods in document image analysis. It effectively addresses challenges like bleed-through, ink spots, and spill-through artifacts, and maintains stroke connectivity. It produces high-quality binarized images, accurately distinguishing foreground text from background noise. However, it faces limitations in binarizing images with similar backgrounds and text colors, and with images with dark-colored degradations or similar degradation colors.

In our study, we sought to assess the efficacy of our proposed method by conducting experiments on images sourced from the DIBCO dataset. The objective was to thoroughly evaluate the performance of our method in comparison to other novel approaches. Through meticulous experimentation and analysis, we aimed to ascertain the efficiency and robustness of our proposed method in addressing the challenges inherent in document image binarization. The table provided below encapsulates the average results obtained from the evaluation of each DIBCO dataset image using our proposed method. This comparative analysis serves to validate and reinforce the findings presented in the table-2, thereby offering a comprehensive understanding of the performance of our proposed method in relation to existing state-of-the-art approaches.

Table 2: Analysis of the proposed and existing state-of-the-art methods' comparative performance on images from the DIBCO dataset

Dataset	Metrics	Sehad [30]	Michalak [31]	Khitas [29]	Mustafa [27]	Mustafa [28]	Proposed
DIBCO 2018	**F-measure**	87.36	86.48	88.59	89.28	88.21	**89.56**
	PSNR	17.65	16.58	18.12	17.54	17.28	**18.52**
	NRM($\times 10^{-2}$)	05.67	05.74	05.38	05.53	05.38	**05.24**
	DRD	04.28	04.34	04.25	04.39	04.15	**04.09**

DIBCO 2016	**F-measure**	85.26	88.24	87.25	87.45	89.12	**90.27**
	PSNR	16.54	17.62	16.23	16.89	17.84	**18.47**
	NRM($\times 10^{-2}$)	05.47	05.64	05.49	05.58	05.48	**05.14**
	DRD	04.25	04.28	04.33	04.27	04.32	**04.12**
DIBCO 2013	**F-measure**	87.14	86.54	89.78	88.24	87.45	**90.58**
	PSNR	18.12	17.93	17.96	18.21	18.45	**18.76**
	NRM($\times 10^{-2}$)	05.45	05.67	05.78	05.74	05.65	**05.27**
	DRD	04.37	04.33	04.52	04.41	04.55	**04.16**
DIBCO 2011	**F-measure**	88.43	88.45	89.22	88.75	89.34	**89.67**
	PSNR	17.56	16.92	16.28	17.48	18.02	**18.34**
	NRM($\times 10^{-2}$)	05.29	05.34	05.37	05.24	05.28	**05.13**
	DRD	04.54	04.38	04.62	04.53	04.32	**04.29**

Results and Discussion for the result of table-2:

Table-2 presents a comparative analysis of different methods' performance metrics across DIBCO 2018, 2016, 2013, and 2011, focusing on key evaluation measures like F-measure, PSNR, NRM, and DRD.

- Our proposed method consistently outperforms other methods on DIBCO 2018 dataset, achieving highest F-measure (89.56) and PSNR (18.52), lowest NRM (05.34) and DRD (03.57), and competitive performance in F-measure and DRD.
- Our proposed method outperforms the DIBCO 2016 dataset, achieving the highest F-measure (90.27), PSNR (18.47), and lowest DRD value (03.79). Furthermore, Method 5 has commendable performance, particularly in terms of F-measure.
- Our proposed method maintains its dominance on the DIBCO 2013 dataset, with the highest F-measure (90.58), PSNR (18.86), and the lowest DRD value (04.28). Furthermore, Method 5 demonstrates competitive performance, especially in the F-measure and PSNR measures.
- Our suggested solution outperforms the DIBCO 2011 dataset, with the highest F-measure (89.67), PSNR (18.34), and lowest DRD value (04.12). Furthermore, Method 5 shows competitive performance across a variety of criteria.

These findings demonstrate our proposed method's consistent superiority across several DIBCO datasets and reinforce its usefulness in producing high-quality outcomes in document image binarization tasks. The proposed method effectively binarizes degraded historical document images, demonstrating superior performance across all datasets and metrics. Its high F-measure scores and high PSNR values suggest robustness in separating foreground from background, while lower NRM and DRD values indicate reduced errors and distortion.

5. CONCLUSION:

Our proposed method significantly advances the field of document image binarization, surpassing current state-of-the-art techniques across various assessment measures and datasets. Extensive experimentation and analysis have shown that our method produces high-quality binarized images with enhanced precision, recall, signal-to-noise ratio, noise reduction, and distortion preservation. The consistent superior performance of our method across multiple DIBCO datasets underscores its versatility and applicability in real-world scenarios, demonstrating its effectiveness in addressing challenges in document processing and analysis.

The success of our method is greatly influenced by the careful selection of parameters, including the window size for background estimation, the number of iterations required for accurate background approximation, and the appropriate size ranges to eliminate noise while preserving text integrity. Typically, bold text requires more iterations to achieve precise background estimation, whereas faint text needs a larger window size to effectively distinguish it from the background. The text size range can be adjusted based on font thickness and pixel size within the image, allowing for optimal parameter customization to enhance performance.

In future work, we aim to address the challenges in image with similar backgrounds and text colors, and images with dark-colored degradations or similar degradation colors. and improving stroke width connectivity. Further we plan to implement the novel technique to enhance its practical usability for large-scale document processing tasks.

REFERENCES:

[1] Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62-66.

[2] W. Niblack, “An Introduction to Digital Image Processing”, Prentice-Hall, Englewood Cliffs, NJ, 1986 pp. 115–116.

[3] J. Bernsen, “Dynamic thresholding of grey-level images,” in International Conference on Pattern Recognition, 1986, pp. 1251–1255.

[4] Sauvola, J., & Pietikäinen, M. (2000). Adaptive document image binarization. Pattern Recognition, 33(2), 225-236.

[5] C. Wolf, J.. -M. Jolion and F. Chassaing, "Text localization, enhancement and binarization in multimedia documents," 2002 International Conference on Pattern Recognition, Quebec City, QC, Canada, 2002, pp. 1037-1040 vol.2, doi: 10.1109/ICPR.2002.1048482.

[6] Kim, I., Jung, D., & Park, R. (2002). Document image binarization based on topographic analysis using a water flow model. Pattern Recognition, 35(2), 265-277.

[7] Gatos B, Pratikakis I, Perantonis SJ (2006) Adaptive degraded document image binarization. Pattern Recognition 39:317–327.

[8] Su, B., Lu, S., & Tan, C. L. (2010). Binarization of historical document images using the local maximum and minimum. In Proceedings of the 2010 International Workshop on Document Analysis Systems (pp. 65-69). ACM.

[9] I. Pratikakis, B. Gatos, and K. Ntirogiannis, “ICDAR 2011 Document Image Binarization Contest (DIBCO 2011),” in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, 2011.

[10] Gangamma, B., & Murthy, S. K. (2011). Enhancement of degraded historical Kannada documents. International Journal of Computer Applications, 29(11), Volume 29– No.11.

[11] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, “Digital Image Processing Using MATLAB”, Third Edition Tata Mc Graw Hill Pvt. Ltd., 2011.

[12] Howe, N.R. Document binarization with automatic parameter tuning. IJDAR **16**, 247–258 (2013). https://doi.org/10.1007/s10032-012-0192-x.

[13] Pratikakis, I., Gatos, B., & Ntirogiannis, K. (2013). ICDAR 2013 Document Image Binarization Contest (DIBCO 2013). 2013 12th International Conference on Document Analysis and Recognition, 1471-1476. doi: 10.1109/ICDAR.2013.219.

[14] Sehad, A., Chibani, Y., Cheriet, M., & Yaddaden, Y. (2013). Ancient degraded document image binarization based on texture features. In 8th International Symposium on Image and Signal Processing and Analysis (ISPA 2013) (pp. 630-635). Trieste, Italy.

[15] Singh, B. M., & CSIT, M. (November 2014). Efficient binarization technique for severely degraded document images. CSIT, 2(3), 153–161. DOI: 10.1007/s40012-014-0045-5.

[16] Gatos, B.G. (2014). Imaging Techniques in Document Analysis Processes. In: Doermann, D., Tombre, K. (eds) Handbook of Document Image Processing and Recognition. Springer, London. https://doi.org/10.1007/978-0-85729-859-1_4.

[17] Wagdy, M., Faye, I., & Rohaya, D. (2014). Degradation Enhancement for the Captured Document Image Using Retinex Theory. In 2014 International Conference on Information Technology and Multimedia (ICIMU) (pp. 363). Putrajaya, Malaysia: IEEE. DOI: 10.1109/ICIMU.2014.7066715.

[18] Ranganatha D and G. Holi, "Hybrid binarization technique for degraded document images," 2015 IEEE International Advance Computing Conference (IACC), Banglore, India, 2015, pp. 893-898, doi: 10.1109/IADCC.2015.7154834.

[19] Sehad, A., Chibani, Y., Hedjam, R., & Cheriet, M. (2015). LBP-Based Degraded Document Image Binarization. In Image Processing Theory, Tools and Applications (pp. 978-1-4799-8637-8/15/$31.00). IEEE.

[20] Wagdy, M., Faye, I., & Rohaya, D. (2015). Document Image Binarization Using Retinex and Global Thresholding. Electronic Letters on Computer Vision and Image Analysis, 14(1), 61-73.

[21] I. Pratikakis, K. Zagoris, G. Barlas, and B. Gatos, “ICFHR 2016 handwritten document image binarization contest (H-DIBCO 2016),” in Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, 2016.

[22] D. Lu, X. Huang, C. Liu, X. Lin, H. Zhang and J. Yan, "Binarization of degraded document image based on contrast enhancement," 2016 35th Chinese Control Conference (CCC), Chengdu, China, 2016, pp. 4894-4899, doi: 10.1109/ChiCC.2016.7554113.

[23] I. Pratikakis, K. Zagori, P. Kaddas, and B. Gatos, “ICFHR 2018 competition on handwritten document image binarization (H-DIBCO 2018),” in Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, 2018.

[24] Bannigidad, P., Gudada, C. (2018). Restoration of Degraded Historical Kannada Handwritten Document Images Using Image Enhancement Techniques. Advances in Intelligent Systems and Computing, vol 614. Springer, Cham. https://doi.org/10.1007/978-3-319-60618-7_49.

[25] Xiong W, Jingjing X, Xiong Z, Wang J, Liu M, (2018), Degraded historical document image binarization using local features and support vector machine (SVM), Pages 218-223, ISSN 0030-4026,https://doi.org/10.1016/j.ijleo.2018.02.072.

[26] Baidoo, E., & Kontoh, A. K. (2018). Implementation of Gray Level Image Transformation Techniques. I.J. Modern Education and Computer Science, 5, 44-53. DOI: 10.5815/ijmecs.2018.05.06.

[27] Mustafa, W. A., Yazid, H., & Jaafar, M. (2018). An improved Sauvola approach on document images binarization. Journal of Telecommunication, Electronic and Computer Engineering, 10, 43-50. https://api.semanticscholar.org/ CorpusID:108313164.

[28] Mustafa, W. A., & Abdul Kader, M. M. M. (2018). Binarization of document image using optimum threshold modification. IOP Conference Series: Journal of Physics: Conference Series, 1019(1), 012022. doi:10.1088/1742-6596/1019/1/012022.

[29] Khitas, M., Ziet, L., & Bouguezel, S. (2018). Improved degraded document image binarization using median filter for background estimation. Elektronika ir Elektrotechnika, 24(3), 81-85. https://doi.org/10.5755/j01.eie.24.3.20982.

[30] Sehad, A., Chibani, Y., Hedjam, R., & Cheriet, M. (2019). Gabor filter-based texture for ancient degraded document image binarization. Pattern Analysis and Applications, 22(1-22). https://doi.org/10.1007/s10044-018-0747-7.

[31] Michalak, H., & Okarma, K. (2019). Improvement of image binarization methods using image preprocessing with local entropy filtering for alphanumerical character recognition purposes. Entropy, 21(6), 589. https://doi.org/10.3390/e21060589.

[32] Samosir, R. S. (2020). Enhancement of Old Document Image after Restoration with Morphological Approach. International Journal of Computer Theory and Engineering,Vol 12(1),feb-2020.

[33] Yang, J., Xu, Y., Yue, H., Jiang, Z., & Li, K. (2020). Low-light image enhancement based on Retinex decomposition and adaptive gamma correction. IET Image Processing. DOI: 10.1049/ipr2.12097.

[34] Venkatesh, S., John De Britto, C., Subhashini, P., & Somasundaram, K. (2022). Image Enhancement and Implementation of CLAHE Algorithm and Bilinear Interpolation. Cybernetics and Systems. DOI: 10.1080/01969722.2022.2147128.

[35] Singh, P., Vats, E., & Hast, A. (2017). Learning Surrogate Models of Document Image Quality Metrics for Automated Document Image Processing. arXiv preprint arXiv:1712.03738v1 [cs.CV]. Retrieved from https://arxiv.org/abs/1712.03738v1.