Image Enhancement for Scanned Historical Documents in the Presence of Multiple Degradations

FAROUK SULEIMAN

July 2023

School of Computing, Science and Engineering University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy

Abstract

Historical documents are treasured sources of information but typically suffer from problems with quality and degradation. Scanned images of historical documents suffer from difficulties due to paper quality and poor image capture, producing images with low contrast, smeared ink, bleed-through and uneven illumination. This PhD thesis proposes a novel adaptative histogram matching method to remove these artefacts from scanned images of historical documents. The adaptive histogram matching is modelled to create an ideal histogram by dividing the histogram using its Otsu level and applying Gaussian distributions to each segment with iterative output refinement applied to individual images. The pre-processing techniques of contrast stretching, wiener filtering, and bilateral filtering are used before the proposed adaptive histogram matching approach to maximise the dynamic range and reduce noise. The goal is to better represent document images and improve readability and the source images for Optical Character Recognition (OCR). Unlike other enhancement methods designed for single artefacts, the proposed method enhances multiple (low-contrast, smeared-ink, bleed-through and uneven illumination). In addition to developing an algorithm for historical document enhancement, the research also contributes a new dataset of scanned historical newspapers (an annotated subset of the Europeana Newspaper - ENP – dataset) where the enhancement technique is tested, which can also be used for further research. Experimental results show that the proposed method significantly reduces background noise and improves image quality on multiple artefacts compared to other enhancement methods. Several performance criteria are utilised to evaluate the proposed method's efficiency. These include Signal to Noise Ratio (SNR), Mean opinion score (MOS), and visual document image quality assessment (VDIQA) metric called Visual Document Image Quality Assessment Metric (VDQAM). Additional assessment criteria to measure post-processing binarization quality are also discussed with enhanced results based on the Peak signal-to-noise ratio (PSNR), negative rate metric (NRM) and F-measure.

Keywords:

Image Enhancement, Historical Documents, OCR, Digitisation, Adaptive histogram matching

Table of Contents

LIST OF FIGURES

List of Tables

List of Abbreviations

ACMHE	Adaptive Contrast enhancement using Modified Histogram Equalization
AEPHE	Adaptive extended piecewise histogram equalization
BMEHHD	Bio-inspired Modelling for the Enhancement of Historical Handwritten Documents
BBTRSHD	Blind Bleed-through removal for Scanned Historical Documents
BMO	Barnacles Mating Optimizer
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolution Neural Network
CRF	Conditional Random Fields
DIBCO	Document Image Binarization Contest
DIQA	Document Image Quality Assessment
DRD	Distance Reciprocal Distortion
EDR	Edit Distance Rate
ENP	Europeana Newspapers Project
GT	Ground truth
HE	Histogram Equalization
HVS	Human Visual System
MSCN	Mean Contrast Normalised Coefficient
MPM	Misclassification Penalty Metric

MOS	Mean Opinion Score
MSE	Mean Square Error
NRM	Negative Rate Metric
OCR	Optical character recognition
PGA	Partition based Genetic Algorithm
PSF	Point spread function
PSNR	Peak Signal to Noise Ratio
SHO	Selfish Herd Optimizer
SSIM	Structural Similarity Index Measure
VDIQA	Visual Document Image Quality Assessment
VDQAM	Visual Document Image Quality Assessment Metric

Acknowledgements

I want to thank Allah for his guidance and blessings and for giving me the patience, perseverance, and fortitude to accomplish this research effectively.

This work is dedicated to my family, who have given me the support, love, and guidance to complete this goal, especially my father, Arc. Mohammed Sule and my mother, Dr Husseina Sule (PhD).

The thesis would not have succeeded without my primary supervisor Prof. Apostolos Antonacopoulos, and my second supervisor Stefan Pletschacher. I am grateful for their continuous support and guidance during my PhD programme.

I would also like to extend immense gratitude to my friends and colleagues within the University of Salford and the Manchester community. They have helped me with sound advice during my research work.

Finally, my immense appreciation goes to the Petroleum Technology Development Fund (PTDF) for sponsoring my PhD program. This accomplishment would not have been possible without them.

List of Publications

External Publications

Suleiman, F., Hughes, C. J., & Obio, E. B. (2021, December). Adaptive Enhancement for Scanned Historical Document Images. In 2021 IEEE 4th International Conference on Electronics and Communication Engineering (ICECE) (pp. 66-71). IEEE.

Internal Publications

Suleiman, F. and Antonacopoulos, A. (2022). Image Enhancement of Historical Documents. IPGRC. 2022: Resilience in Research and Practice conference, https://hub.salford.ac.uk/ipgrc-conference-2022/

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation of the study

Document collections such as books, journals, receipts, pictures, agreements, magazines, and medical records, would someday be considered valuable historical documents as they might be referred to for general accounts or improvement. However, a record number of these documents might be lost since they are either paper prints with a restricted lifetime or have a lot of noise because of their capture technique (Antonacopoulos & Downton, 2007). Historical documents hold essential information and making them accessible to readers worldwide is necessary for knowledge preservation. However, the resultant fragility of these documents restricts access to many researchers, making digitisation a more viable option (Jajware & Agnihotri, 2020). As a result, more digital libraries are created to expand access to scientific, educational, cultural, and historical documents and information. These Digital libraries provide excellent opportunities for educating, improving knowledge and providing the required historical background (Antonacopoulos & Karatzas, 2005).

In most instances, historical document images cannot be analysed visually or directly sent to an Optical Character Recognition (OCR) system because the original document manuscripts suffer from several kinds of degradation. Such degradation includes smeared ink, bleed-through, intensity variation etc., seen especially in historical documents (Kluzner et al., 2009). Historical document recognition is amongst the most challenging problems in image pre-processing. This is because historical document images suffer from numerous factors such as noise, low paper quality, poor typesetting, poor image capture, degradation, low contrast etc. (Hao et al., 2020). Therefore, to preserve these documents securely and safely in the library catalogue, it is recommended to digitise and archive them as a backup to the original manuscripts, providing an alternative in the event of age damage or loss (Gupta et al., 2007).

However, to make these documents to their highest standards digitally in libraries, it is vital to enhance clarity and visibility and generally advance the overall image quality (Li et al., 2020). Increased image quality can be achieved by developing suitable image enhancement methods to give the best quality document image results. Image Enhancement techniques are generally employed to revamp the clarity of images for viewing, eliminate blurring, decrease noise, and improve contrast. This uncovers more details that may have been concealed in the original image (Likforman-Sulem et al., 2011). The primary objective of enhancement is to improve an image so that the resultant image is a better representation for a specific application (Hong et al., 1998).

This PhD project intends to examine existing historical document image enhancement methods and create a novel image enhancement method that improves degraded historical documents by generally reducing noise artefacts, improving the objective and subjective quality of historical document libraries with little computational time. The research presents a method of image enhancement expected to perform well for most artefacts (multiple degradations) in the datasets and not just for one degradation as most in document enhancement literature showcase. This PhD report proposes an adaptive histogram matching technique that uses a reference image generated after denoising to match an ideal histogram image generated from state-of-the-art histogram manipulation. Both histograms are then matched to reveal an enhanced image. Creating the ideal histogram entails dividing the image histogram into two segments using the Otsu (Otsu, 1979) level as the boundary to separate two main distributions and computing the maximum of each distribution. The Otsu level is the threshold value where the sum of foreground and background pixels spreads at its minimum. Two Gaussian distributions are generated for each segment (foreground and background) and interpolated with the image's histogram to give importance weights to the distribution. This is used to ensure a better bimodal shape of the histogram using the multiplication of the Gaussian bell in each segment to give significance to the pixels following the parabolic curve. Both segments are combined to achieve the final reference histogram that is applied to traditional histogram matching. The novelty in this process is that for each image the adaptive histogram matching technique generates its reference histogram to be applied. Considering the large noise variance in different historical documents, some noise filtering techniques are also included in the process.

In addition to visual enhancement, the proposed technique is also designed to simultaneously improve binarization (significant in machine vision) outcomes. To carry out the experimental evaluation, this research considers newspaper images from Europe's significant libraries (Europeana Newspapers Project Dataset) available at the PRImA Research Lab. As expected with historical documents, a significant number of them have uneven font intensities/ blurred letters, blurred edges, and noise due to lack of proper enhancement and poor image scanning techniques. A subset of this dataset is created for easy experimentation and evaluation.

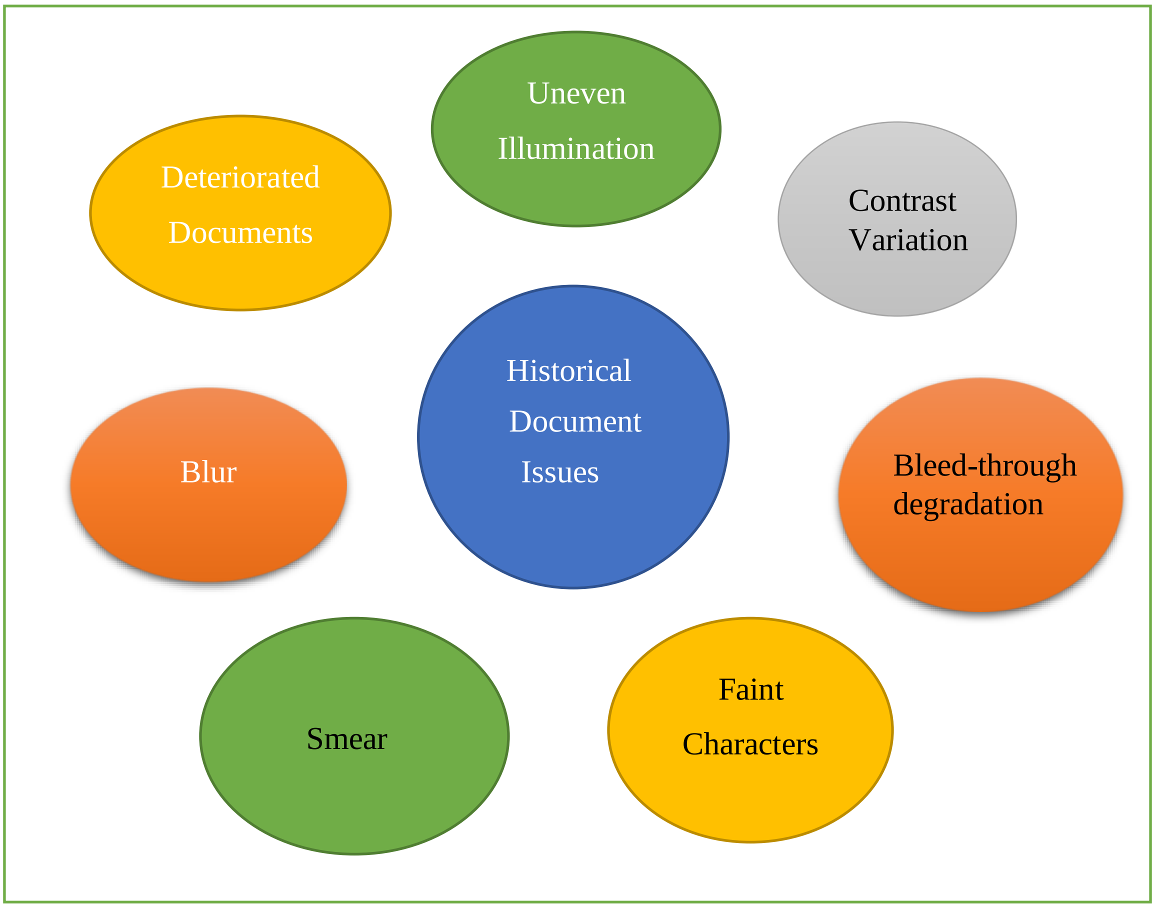


Figure 1- 1 Frequently seen degraded defects in Historical Documents

1.2 Problem Definition

Historical documents often deteriorate due to poor storage environments and a reduced contrast between foreground and background, which can be caused by humidity, paper decay, seeping of ink, etc. The resultant delicateness of these documents restricts access to many researchers, thereby making digitisation a more viable option (Kavallieratou & Antonopoulou, 2005). As such, before libraries showcase historical records, it is crucial to enhance the quality of the images for easy perception and processing.

Several challenges exist in literature because most documents are scanned or captured in hasty succession during the digitisation process. This factor and the degrading environmental factors present in some cases causes digital libraries to be of low quality. When these documents are passed to the OCR software or human visual system (HVS)/ human eyes for analysis, both systems face challenges due to poor document quality and issues with the digitisation process. From research carried out in the Europeana Newspapers (ENP) collections in the Pattern Recognition and Image Analysis Research lab (PRImA) database(Primaresearch, 2021), it was identified that a lot of the datasets have common undesired artefacts and degradations like uneven font intensities, uneven illumination, low-scan-contrast and faded ink, due to poor image acquisition and storage. The OCR performance on these documents compared to the corresponding ground truth shows a gap in what OCR recognises as compared to the original. The human visual system also can easily detect these artefacts and the resultant reduced legibility of the characters. (Papadopoulos et al., 2013).

Although several enhancement methods have been proposed in the literature, most are centred around specified problems such as bleed-through correction (He & Schomaker, 2019; Moghaddam & Cheriet, 2009; Sun et al., 2016; Yagoubi et al., 2015). This means that most enhancement methods are not designed to handle multiple degradations in historical documents. Deep Neural network, an unsupervised system with that can be trained to enhance multiple degradations, also has challenges and limitations in image enhancement, including the availability of labelled data for training. This can restrict the enhancement process for unknown artefacts that the neural network is not familiar with (Kang et al., 2019). Also, their computational complexity remains high, and the accuracy may degrade due to the non-uniformity of foreground and background intensity.

1.3 Aims and Objectives

1.3.1 Project Aim

To develop a method for enhancing historical documents with multiple degradation while improving readability and recognition results.

1.3.2 Project Objectives

The objectives of this project include:

i. Conduct a thorough analysis to produce a detailed report on historical document artefacts within the IMPACT and Europeana datasets, aiming to qualify and quantify issues present in the collections of national and major European libraries, focusing on holdings and digitization.

ii. Administer a human survey evaluation using a derived representative dataset, a smaller subset of the Europeana datasets, to gather quantifiable data on user perceptions and preferences related to historical documents.

iii. Develop a cutting-edge image enhancement approach for historical documents and assess its effectiveness in improving layout and Optical Character Recognition (OCR) outcomes.

iv. Perform a quantitative and qualitative evaluation using well-defined metrics, showcasing both subjective and objective performance improvements in layout and OCR.

v. Based on the analysis and report, propose new approaches for historical document enhancement, and evaluate their feasibility and effectiveness.

The above objectives are anticipated to be obtainable by first making critical reviews on existing historical document image enhancement literature.

1.4 Questions and Hypothesis

The main question of the study is: How to address the problem of undesired artefacts and degradations in historical documents by focussing on grayscale images and to what extent this is possible?

From the main question of the study, a few sub-questions can be derivative:

1. What are the historical document artefacts and problems and what are their prevalence?
2. What datasets contain the most significant problems and how are they represented?
3. What are the important approaches used for image enhancement of historical documents to discover the state-of-the-art in these studies?
4. How the impact of pre-processing enhancement affects the outcome of post-processing binarization, layout and OCR?
5. What are the new and state-of-the-art concepts that can be used for evaluation of the historical documents, as prescribed by the specific characteristics of the problem?

1.5 Contributions

The projected contributions for this research project are:

- Development of novel state-of-the-art adaptive document image enhancement method that enhances degradation according to artefact characteristics. The technique also ensures optimum improvement in OCR performances, visual appearance for human readers and improves the bimodality of grayscale document histograms with emphasis to the pixels related to foreground and background, while diminishing the middle pixel (noise) values. The middle pixel values are associated with undesired show-through artefacts and signal interference in the background, thus reducing this amount will diminish the undesired noise thereby ensuring a more desired document image for optimal enhancement of historical documents library.

- Report of qualified and quantified issues affecting OCR in historical document collections and the solutions.

1.6 Method of Study

The definitive goal of this research is to discover a solution to the research problem described earlier, which can be summarised in different types of degradation that can occur at several stages or exposures in a document lifecycle namely printing, storage, print reproduction and use. The research methodology will comprise four core stages which are: a theoretical study, measuring the performance of image enhancement methods, building a system for document image enhancement, and comparing and analysing this method with other state-of-the-art.

1.6.1 Theoretical Study

From a thorough literature review and study, it can be deduced that enhancement of historical documents includes reducing noise artefacts and refining the objective and subjective quality of images by improving their foreground and background qualities. This is important for better image representation and post-processing analysis like binarization and OCR.

1.6.2 Measuring the performance of image enhancement methods

Selecting an effective image enhancement method is a crucial step to take when looking at removing or reducing artefacts in historical document image output. Therefore, thorough research steps identify the state-of-the-art methods and record their performance to identify the gap for novel recommendation.

1.6.3 The proposed image enhancement method for historical documents

In this study, the researcher will focus on developing an image enhancement technique for historical documents. The proposed method for image enhancement of historical documents involves three components namely, contrast stretching, noise filtering, and the state-of-the-art adaptive histogram matching technique.

1.6.4 Performance comparison and discussion

The research will complete the result verification process where testing and evaluation are carried out to evaluate the image enhancement method. The experiment is conducted on several printed document images in Latin text with different types of artefacts present and compared with existing methods.

1.7 Thesis Structure

- Chapter one of this report presents a general overview of the research topic by specifying the background and motivation of the study, problem definition, the research aims and objectives, expected contributions, and the report's structure.

- Chapter two presents a literature review of historical document image processing issues and methods for enhancement, methods of document image quality assessment, image noise filtering, and review of postprocessing binarization methods.

- The third chapter presents the research methodology and tools used in the framework and the novelty of the experiments. The subset of the ENP dataset created for the experiment is also portrayed in this section.

- The fourth chapter discusses the implementation procedure to demonstrate the working model of the technique. The stages of the pipeline method are also defined depicting the working theory. Some preliminary performance evaluations to show the method's capability are also shown to validate it.

- In the fifth chapter, the performance analysis is shown. This involves the experimentation on the entire subset and the evaluation of its performance in terms of subjective and objective evaluation.

1.8 Research Task

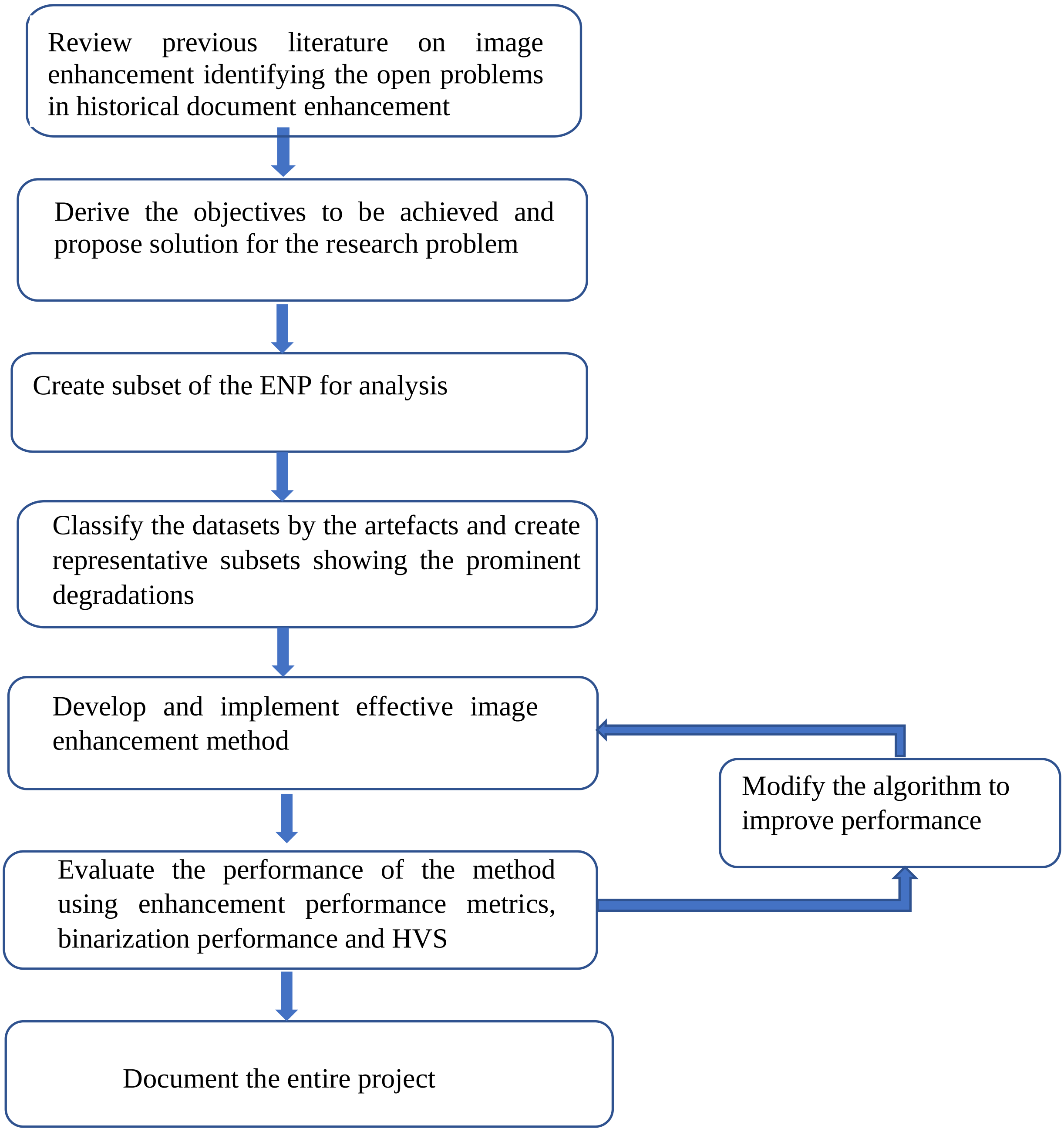


Figure 1- 2 Research plan flow chart

CHAPTER 2

LITERATURE REVIEW

As one of the objectives of this research is to propose a historical document image enhancement technique that can improve the quality of historical document images. This chapter encompasses the observing and analysing of various knowledge sources, procedures and algorithms involved in historical document image enhancement, and how the amassed knowledge can be exploited to improve image enhancement.

The goal of this literature review is to discover and explore the gains of image enhancement algorithms and likewise the shortcomings in the current existing algorithms and methods. The literature review also intends to find the gaps in the existing research, techniques applied and possible solutions to overcome these drawbacks. The chapter presents an overview and critical review of the ideas, techniques, and contributions provided by researchers from works of literature on document image enhancement. The first segment of the chapter discusses different document images and how they can be improved. In the last segment of the chapter, the methods and challenges involved in image enhancement of historical document images are reviewed.

Image enhancement methods are procedures initiated to improve degraded image quality. It is a necessity because follow-up analysis phases such as human evaluation, page layout analysis or OCR require a prespecified level of quality to attain optimum performance. Most image enhancement processes generally start with data set collection, image analysis procedure for noise removal, digital enhancement application, test, and evaluation.

2.1 Document Image Quality Assessment (DIQA)

The factors that can influence degradation in documents can be classified into two groups. Factors attributable to bad preservation state, such as stains, humidity, oxidation, rips, holes, etc. The other degrading factors are derived from the method of creation used, for example, carbon-copied documents, typewritten documents, photocopy prints, faxes etc. The degree of effect by these factors on documents can be noticed in the varied background and foreground intensity, shadowing, character smear, skewing, low contrast, noise, fragmented and linking fonts, blurred font/character strokes(Gatos et al., 2006). Modern office document handling challenges come mostly from the digitisation of several-generation photocopies where in each generation, the character strokes grow increasingly thinned and start to break down at feeble points. While dealing with historical or archived documents, the initiating original copy might have over time become subject to ageing, therefore lightening text density. Inappropriate methods of digitisation may result in watered-down text and broken or connected strokes. This can have a negative influence on visual comprehension and inaccurate OCR performance on degraded documents and restoring stroke width is of prime importance to achieving better document appearance(Shi & Govindaraju, 2004). Document image enhancement techniques aim to correct and reduce the effects on degradation towards visualisation and recognition. The form of image enhancement includes noise reduction, side enhancement and distinction enhancement. Enhancement could also be described as the method of improving electrically saved image prevalence(Bankman, 2008).

In the following subsections, several works of literature on historical document datasets and their challenges are discussed.

2.1.1 Historical Document Digital Libraries - Challenges

Antonacopoulos, A. (2010) discussed an overview of the general challenges of large-scale digitisation amongst which were the document characteristics through its lifespan. It was deduced that the major challenges in Image Analysis were Scanning methods, Image Compression, Image enhancement Layout Analysis and Character recognition. These areas have substantial room for improvement in Binarization applied techniques for effective segmentation to have greater quality historical libraries as the digitisation of historical documents only came into large scale enterprise a little over ten years ago.

Antonacopoulos et al. (2011) discussed a means to improve historical document libraries by comparative evaluation of layout analysis methods intended for historical documents that are scanned. The competition aimed to evaluate and test layout analysis on new historical image datasets to see how they perform. The evaluation was directed at checking the ability to adequately segment regions based on text and background. It was also emphasised from the results that better-enhanced images performed better and made easy recognition for the segments.

According to (Ye & Doermann, 2013), (Shahkolaei et al., 2019) document image quality assessment is a challenging problem in Historical document enhancement as there are not reliable methods for estimating the level of degradation. Previous work on the assessment of image quality was more focused on natural scenery images, however, document properties do not permit their application. The objective of their paper is to present a survey on the topic as a standard for future work since there have been relatively minute publishing on the DIQA problem. They also discussed objective measures and subjective experiments that can be used to assess document image quality and estimate the level of degradations. The objective measure deals with measuring image quality when the image vision is received machine by computational algorithms e.g., OCR software, the document image quality may be presented as a measure of the OCR accuracy and DIQA metrics. The subjective on the other hand was described in their paper as an image quality test based on human perception. It was carried out by making surveys in which participants can grade the image performance and quality on an ordinal scale from bad to excellent. From the various test expressed, they were able to show that an improvement in enhancement for a subjective consumer might not mean an improvement in enhancement for an objective user as machines tend to read images differently from the Human Vision System.

According to (Sulaiman et al., 2019) who also discussed and reviewed the issues, challenges, techniques and future directions on degraded historical document binarization. This includes classifying some of the most frequently degraded defects in historical documents like uneven illumination, contrast variation, smeared ink, faded ink, blur and bleed-through. They also described how these defects have a significant impact on document recognition after binarization due to improper segmentation because of the interference of the defects. The performance measure for document binarization reviewed and suggested by (Pratikakis et al., 2013) and (Pratikakis et al., 2017) includes F-measure, Pseudo F-Measure (FMp), Peak Signal-to-Noise (PSNR), Negative Rate Metric (NRM), Misclassification Penalty Metric (MPM), Average Quality Score, Distance Reciprocal Distortion (DRD). However, these techniques are only traditionally applied to binarized images and not to measure enhancement.

According to (Shahkolaei et al., 2018), the importance of enhancement measures in historical document images as the severity of degradation before imaging is usually unknown. The assessment of degradation is important to help with tuning parameters, selecting the proper algorithm etc. The paper proposed a Visual Document image Quality Assessment Metric (VDQAM) using Visual Document Image Quality Assessment (VDIQA) through human visual system to score the quality of historical documents instead of OCR performance, due to OCR working best when an image is segmented through binarization, thus making OCR working best as a binarization metric and not as conventional image enhancement. The paper also proposed an objective no-reference quality metric based on mean contrast normalised coefficients (MSCN). The image would be segmented in four layers on log-Gabor filters on the assumption of the sensory of the human visual system (HVS) as it relates to the area of text and non-text as shown in Figure 2-1. The research made a distinction that a document with physical noise close to the text and far from the text may not equally contribute to the visual quality of document images. It also established that text and non-text parts do not have an equal impact on HVS judgements. Shahkolaei et al (2019) improved in the earlier proposal by segmenting tested images into two layers using the log-Gabor filters and MSCN coefficients. The robust improvement stems from taking the locally weighted mean phase angle extracted from the two layers. These spatial statistics are utilised for quality assessments. They also proposed a degradation classification model based on each proposed metric to measure the possibility of the different artefacts. SVM classifiers are used to categorise the degraded images into four degradation categories. The paper suggests that the reason why human judgments are used in the introduced dataset instead of OCR accuracy is that OCR engines are not perfect, especially for some of the languages, old writing styles and fonts. The other reason is that a higher OCR accuracy does not automatically mean that a document's image is of high quality and may just mean that the text region is not degraded. This method is ideal for the part of the objective evaluation of the proposed method and result.

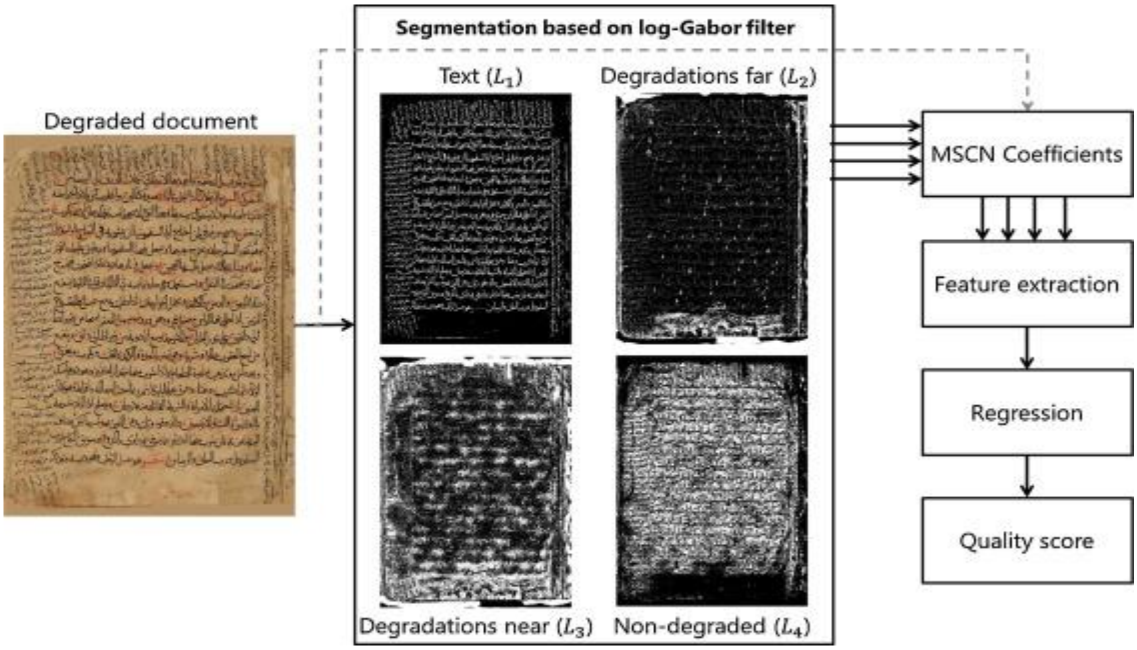


Figure 2- 1 Proposed VDQAM metric illustration (Shahkolaei et al., 2018)

2.2 Image Noise Filtering

Noise is generally classified as any irrelevant or irregular information within the textual information of a document image that makes the image distorted. Examples of image noise include stray marks, marginal noise, ink blobs and salt-and-pepper noise. Noise removal is considered the most important task towards image enhancement and the technique is largely known as Noise Filtering. Noise Filtering is a well-known technique for removing errors that might have occurred in image acquisition. They are usually classified into linear and non-linear. Examples of noise are Amplifier noise (Gaussian noise), Salt-and-pepper noise (Impulse noise), Shot noise, Quantization noise (uniform noise), Film grain, anisotropic noise, Speckle noise (Multiplicative noise) and Periodic noise (Mafi et al., 2019).

For instance, during the image acquisition stage, the photoelectric sensor introduces the white Gaussian noise because of the thermal motion of electrons. As a result of the unstable network transfer, impulse noise is added into the image (Boonserm, 2015).

2.2.1 Noise Filtering Method

Nguyen et al. (2010) proposed a spatial denoising algorithm to process grayscale images tainted by Gaussian noise. The process involved using local weighted mean, local weighted activity, and local maxima for noise detection and a spatially additive Gaussian filter is used to counter the additive noise. The filter satisfactorily deals with the level of local smoothness without overcompensating due to its local statistics consideration. The proposed method prioritises computational cost, over smoothness and detection error when removing the noise in grayscale images. This method, therefore, would not be efficient due to the disadvantage of loss of detailed image preservation thereby reducing OCR performance when applied to documents in the subset collection

Qiu et al. (2011) improved the above method by developing a model to estimate the noise. The method entailed combining block-based method and filter-based method to achieve noise standard deviation with low computing demand. The improved standard deviation is used in the noise estimation and the adaptive noising method removes the noise. The performance metrics showed it was self-determined, better than the original method and adapted to image contents. However, the seemingly improved Gaussian filter would not be efficient in edge preservation.

Zhu and Huang (2012) proposed an enhanced median filtering algorithm for image noise reduction. The algorithm was configured to adaptively resize the mask according to noise level and retain image details more efficiently. The results from the experiment showed that the performance of noise reduction was desirable for live images but would not give good results in document images when applied on its own.

Hambal et al. (2017) from their research on the existing techniques of noise filtering on historical documents referenced Qiu et al (2011) and determined that the method reduced Gaussian noise but caused an excessive blurring of edges. They proposed using median filtering for historical documents as it is particularly applicable to removing salt-and-pepper noise and causes relatively low edge blur and can be used in computer vision applications. Median filtering is identical to averaging filter in the sense that the output pixel is put as an average pixel value in the regions of the corresponding input pixel. Since the median is less sensitive than the mean to extreme values, median filtering is consequently able to remove aberrations without diminishing image sharpness. The experimental survey showed that the Median Filter removes impulse, Gaussian noise and preserves edges when applied to historical documents.

The median filter gives the best result with an impulse noise of less than 0.1% which is common in Historical documents. It may not be effective for high impulse noise found in X-Ray images.

Wang et al. (2018) from their research on blurred image restoration using knife-edge function and optimal wiener filtering were able to demonstrate the impact of wiener filtering on restoring motion blur images. The experiment is modelled as the convolution of a point spread function (PSF) and the original image represented as pixel intensities. They used the knife-edge function as a system degrade function to obtain and simulate the blur. The experiment starts with the Prewitt edge detection operator and autocorrelation function is used to calculate the direction and scale of the motion-blur. Subsequently, they added the optimal window to the edge extension image before the detected knife-edge function is used to attain the system degradation function. The Wiener filter is then applied to obtain the restored image and truncate the border. This experiment shows the importance of the Wiener filter in restoring blur degradation that is triggered in document images by scanner/ camera movements during capture.

Gavaskar and Chaudhury (2018) from their research titled fast adaptive bilateral filtering were able to demonstrate the effect of the bilateral filter on edge preservation of characters in an image while smoothing away noise. They improved the conventional bilateral filtering (Aurich

& Weule, 1995; Smith & Brady, 1997; Tomasi & Manduchi, 1998) which describes the bilateral filter as an edge-preserving smoothening technique that uses a range kernel and a spatial kernel (Gaussian kernels). The input to the range kernel is denoted as the difference between the specified pixel and its neighbour. When the classic algorithm detects a large difference in the edge of a pixel, then it assigns a small weight to the neighbouring pixel and essentially excludes it from aggregation. This technique, therefore, ensures the avoidance of mixing of large intensity pixels guaranteeing the preservation of edges. The new fast adaptive bilateral filtering technique proposed combined the process with a histogram approximation technique using polynomials and achieved better results with fewer convolutions than earlier methods. The effectiveness for sharpening, deblocking and texture filtering was also illustrated. This research proves the effectiveness of bilateral filtering when applied to documents and natural images.

2.3 Enhancement Approaches Based on Contrast

Image histograms are frequently normalised by the total number of pixels in an image and represents the intensity levels of pixels in an image. Suppose an image is predominantly dark, then its histogram would be skewed towards the lower end of the grayscale, and when predominantly bright, its histogram is skewed towards the lower end. Where 0 is black and 255 is white on the grayscale. Digital image histogram is a discrete function in the range of pixel values [0, L-1](Rao, 2020). For an 8-bit image the range will be [0, 255].

$$h(r_k) = n_k$$

r_k= k^{th}intensity value, n_k = Number of pixels in the image with intensity r_k.

Assuming an $M * N$ image, a normalized histogram

$p(r_k) = \frac{n_k}{MN}, K = 0,1 \ldots . L - 1$

Is related to probability of occurrence of r_k in the image

This section shows a survey on methods based on pixel manipulation of images.

2.3.1 Image Pixel Manipulation Approach

Kim (1997) has discussed that brightness and scene intensity in images can be changed by histogram equalisation, due to its general flattening quality. Y.-T. Kim (1997) proposed a method of histogram equalisation referred to as bi-histogram equalisation dominated the drawback of the histogram equalisation. The purpose of the proposed algorithm is to preserve the average intensity of a degraded image whilst the contrast is enhanced. The specified methodology starts by breaking the input image into two sub-images based on the mean of the input image. One of the sub-images is the set of specimens that are less than or equal to the mean whereas the other one is the set of specimens greater than the mean. Histogram equalisation is a technique for adjusting image intensities to enhance contrast. It is used to improve poor contrast distribution in images by stretching the intensity range of the image. This allows for areas of lower local contrast to gain a higher contrast thus improving the overall quality of an image. Histogram Equalisation is achieved by remapping grey levels of the image based on the probability distribution of the input grey image levels. It levels and stretches the dynamic range of the image's histogram thereby resulting in overall contrast enhancement (Kaur et al., 2011). Histogram equalisation is commonly used in enhancing contrast due to its simple function and effectiveness.

Kim et al. (1998) can be viewed as an extension to this technique as they deliberated on a block-overlapped histogram equalisation system for refining the contrast of image sequences using numerous applications. The regular histogram-based contrast enhancement technique and the above bi-histogram equalisation application is inadequate when applied to a real-time application. This is because of a large computational and storage prerequisite. It also exhibits quality degradation caused possibly by loss of infrequently distributed pixel intensities, which may result in terrible loss of vital information. The proposed system was able to enhance local contrast while suppressing undesired noise amplification. However, this method had a drawback in preserving the original brightness of the image when applied to Historical documents datasets. However, Histogram Equalisation can lead to over enhancement when applied to certain document images if the algorithm is not properly adapted for the dataset.

Sengee and Choi (2008) proposed an improved image enhancement method called Brightness Preserving Weight Clustering Histogram Equalisation. The algorithm was structured in such a way that it could simultaneously preserve the original image brightness and enhance visualisation of the original image. The method functions by assigning each non-zero bit of the original image's histogram to a different cluster and calculating each cluster's weight. To reduce cluster numbers, three criteria are used namely, the weight of cluster, weight ratio and the widths of two neighbouring clusters to merge with pairs of neighbouring clusters. The clusters attain equal partitions as their image histogram result. Finally, transformation functions for each cluster's sub-histogram are computed, and the sub-histogram's grey levels are mapped to the resultant image by an equivalent transformation function. However, this method is not effective for enriched contrast as it tends to give a washed-out effect when applied to datasets. Moghaddam and Cheriet (2009) proposed a bleed-through correction technique using a variational approach. The variational approach is modelled using an estimated background based on the availability of the verso (rear) side of the document image. The model also utilises an advanced global control flow field using wavelet shrinkage depending on complexity and non-linearity. The system is effective for double-sided document images by using reverse diffusion between the two sides of the document. The flow field classifies the information on the image based on its relation to the edges and boundaries. The method is robust to noise and complex background and can be applied to document images and other fields of image processing.

Yang and Wu (2010) suggested an image contrast enrichment which is especially suitable for multiple-peak images. The given method was used to remove two distinct drawbacks of Histogram Equalisation algorithm by foremostly convolving the input image by Gaussian filter with optimum parameters. In the next step, the original histogram was divided into different areas using the valley values of the image histogram to diminish the washed-out effect. This method was found to outperform earlier applications in aspects of simplicity and flexibility. The result establishes that the proposed algorithm has good performance in image enrichment but not of substantial enhancement in images with dark shadows which are some characteristics of the datasets.

Ling et al. (2015) proposed an Adaptive Extended Piecewise Histogram Equalization (AEPHE) algorithm to enhance dark images with a wide dynamic range. The procedure entailed creating a unique Histogram which is then converted into an assembly of lengthened structured histograms. Then an Adaptive Histogram Equalization AHE which alters contrast and power preservation is further created and separately linked to these established piecewise histograms. The resulting histogram showing image upgrade is made by a subjective blend of these evened out histograms. The experiment showed that AEPHE is an improvement over various previous advanced algorithms in improving dark images. The methodology involves two novel measures, first for intensity preservation measure and the second for contrast boosting, to characterise the geometric features of the original histogram. Towards balancing the intensity preservation and contrast adaptively, they further develop a novel adaptive HE based on these two measures to evade unforeseen over-enhancement or under-enhancement. Finally, all equalised piecewise histograms are fused by a weighting function to efficiently merge the effect in the overlapping parts. The experimental results demonstrate that AEPHE significantly enhances dark regions without introducing excessive enhancement or unnatural artefacts. However, this method reduces the sharpness of the image therefore further deteriorating font intensities when applied to certain degraded images in the dataset like uneven illumination and show-through.

Santhi and Banu (2015) in their paper Adaptive Contrast Enhancement using Modified Histogram Equalization (ACMHE) proposed an adapted enhancement using adjusted histogram to minimise the problems of over enhancement, saturation artefacts and change in mean brightness usually associated with conventional histogram equalization. The histogram of the input image is divided into four sub-histograms based on the median. A clipping process based on the input image mean is then applied. Each partitioned histogram is equalized independently, and a contrast enhancement rate is formulated to achieve the varying contrast for the output images. From their experiment, the proposed algorithm achieved better enhanced images than contemporary techniques in terms of contrast per pixel and structural similarity index. This method is effective for low contrast images; however, it may be appropriate for artefacts like show through and smeared ink which the ENP contains.

Sun et al. (2016) in their paper blind-through removal for scanned historical document images proposed a method to remove bleed-through from historical documents. The procedure requires only the scanned image side as an improvement to Moghaddam and Cheriet (2009). The method presents a new Conditional Random Field (CRF) for one side of the scanned image, referred to as the blind method. The scanned historical document is composed of three components namely, foreground, bleed-through and background. The method uses logistic and Gaussian distributions to approximate the three components using conditional probability distribution models of the foreground, bleed-through and background. The K-means algorithm is used to generate a coarse labelling and parameters of the component-wise model are computed accordingly. The factors of the component condition probability distribution (CPD) is established based on an initial segmentation of the input image. Then a conditional random field-based method is implemented to capture the relation between observed pixels in the scanned image and its spatial relation. Finally, a random filling algorithm was implemented to in-paint the bleed-through region. The experiment showed that the method preserves the foreground and removes the bleed-through region but will not be suitable for other degradations like uneven ink and uneven illumination which the ENP dataset contains.

Pratikakis et al. (2017) in their paper Bio-inspired modeling for the enhancement of historical handwritten documents proposed a pre-processing step aiming to enhance historical document images and improve subsequent binarization. The algorithm is based on the OFF-centre ganglion cells of the Human Vision System (HVS). They described how the HVS does not detect pixel intensity values from 0 to 255 as computer vision systems do. In HVS, brightness and darkness are different stimulations rather than a single value of variable. The ON-centre and OFF-centre ganglion cells are the antagonistic responses responsible for bright and dark perception. The method tackles the enhancement problem of historical handwritten documents by modeling the OFF-ganglion center-surround cells that exist in the retina. It uses a region and heuristics-based algorithm to support the matched center-surround cells. This is achieved by combining the Perona and Malik (1990) diffusion filter with Niblack binarization, and the stroke width of each pixel calculated before the BIO-inspired model.. Figure 2-2 shows the enhancement performance after binarization of the pre-enhanced image compared to the enhanced. The resultant images indicate an improvement in binarization after the pre-processing step. However, this method is not ideal for faded ink, smeared ink, and broken character degradation which the ENP dataset contains.

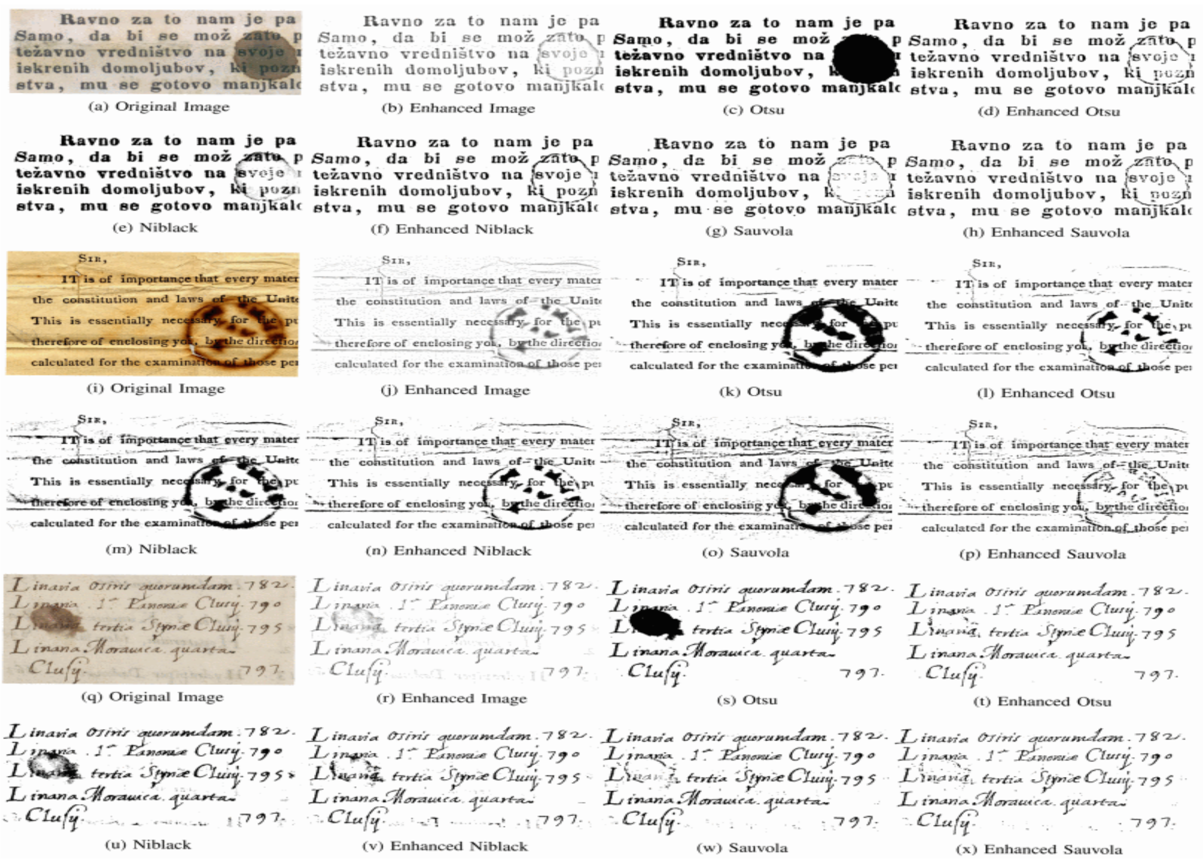


Figure 2- 2 Image 00675211 with broken characters

Ghosh et al. (2019) in their paper Contrast Enhancement of Degraded Document Image using Partitioning based Genetic Algorithm proposed a method to enhance poor quality documents. As restoration of documents in digital form improves accuracy in text recognition, the paper demonstrates an optimization approach named Partition based Genetic Algorithm (PGA) to enhance the contrast of documents with low illumination. The method uses a recursive partitioning to divide an image to sub-images with lower intensity variations. The GA is applied to each sub-image to maintain most of the text-pixels for an improved contrast. The technique is performed on grey level images. In their work, uniform mutation and uniform chromosomes are used. The operations are performed on a set of chromosomes (Xi) which form a population. The chromosomes are vectors of integer values and built from the initial image. The unique integer values present in an image form the chromosome. The method was able to handle noise optimally on use of PGA pre-binarization. The method was tested extensively on the DIBCO 2013 and H-DIBCO 2016 datasets to validate the enhancement method performance. Subsequent binarization using Otsu's method in comparison with the ground-truth was carried out using 6 metrics, namely, Precision, Recall, F-Measure, Accuracy, Peak Signal to Noise Ratio (PSNR) and Distance Reciprocal Distortion (DRD). The recorded mean results showed that the PGA improves the image quality of Otsu's binarization. Precision improvement is minimal, Recall and F-measure is quite remarkable. The authors also claimed decent improvement in terms of Accuracy, PSNR and DRD. The limitation of this research is in cases where the sub-images does not include text, the algorithm magnifies the noise to increment the edges. This leads to a drop in precision. This method would not be ideal for degradations with images and portions with no text which the ENP datasets contain.

(Ahmed et al., 2020) in their paper Gray Level Enhancement using Barnacles Mating Optimizer proposed a meta-heuristic algorithm using BMO. Grey level mapping technique is utilised to convert an image to an optimised version. This is carried out by mapping grey levels of source images into a new set of grey level values. The technique starts by converting the input image into a vector, and an operation denoted to it. The image is represented by an ordered vector of D integers in the interval [0,255] of grey values. A fitness function is then incorporated to evaluate the quality of the agents in the algorithm and a mapping to calculate the fitness value as maximum and expected output image should have a greater number of edges with higher intensity and higher contrast. The technique is applied to benchmark datasets Kodak, MITAdobeFiveK H-DIBCO and H-DIBCO 2018 datasets and evaluated using PSNR, Structural Similarity Index Measure (SSIM) and Visual Information Fidelity. The obtained result recorded was quite significant in the enhancement performance from the output image and histogram respectively. The technique was also evaluated with binarization to indicate the robustness and pixel level clarity of the method post enhancement. From the binarization results, the method also performed significantly in F-Measure, Pseudo-F-Measure, PSNR and Distance Reciprocal. However, the technique may not be suitable for very low contrast images which the ENP contains.

Guha et al. (2022) in their paper titled Enhancement of image contrast using Selfish Herd Optimizer, proposed a pre-processing image enhancement method to optimise the pixel intensity values of an input image to obtain a contrast enhanced version of the same. The process is implemented by a customisation of nature-inspired optimisation algorithm called Selfish Herd Optimizer (SHO). The optimisation challenge is resolved by two different solution representations: pixel wise optimisation (SHO(direct)) and transformation function-based optimisation (SHO(transformation)). The method was experimented over the popular Kodak image dataset, and according to the authors, was observed to outperform many existing methods published recently. The paper further investigated the quality of the SHO(direct) approach by applying it to enhance the degraded document images dataset of H-DIBCO 2018 and compared it with their corresponding GT images. Four evaluation metrics are used to check the binarization performance, F-Measure, pseudo-F Measure, PSNR and Distance Reciprocal. The results showed that the technique enhances uneven illumination and background variations but is not ideal for faded ink and bleed-through degradations which comprise part of the ENP dataset. The advantage of this method is that it significantly improves background illumination for both scene and document images.

2.4 Enhancement Based on Hybrid Methods

This section reviews techniques on the combination of existing methods to achieve desirable results.

2.4.1 Non-Linear Hybrid approach

Hossain et al. (2010) proposed an image enhancement method for medical images applicable to general document images with dark shadows. The process was an improvement to Yang and Wu (2010) and was based on combining transform domain with non-linear histogram equalization. The method’s performance was compared to histogram equalisation and showed significant improvement over it. The algorithm was set up by mixing the non-linear technique and the logarithmic transform coefficient histogram equalisation. Logarithmic transform histogram matching is based on the concept that the relationship between stimulus and perception is logarithmic. This technique enhances the visual quality of images that contain dark shadows due to the limited dynamic range of imaging like x-ray images and poorly captured document images. However, this method does not preserve the edges of the document due to the tendency of over-enhancement as criticised by Cheng and Zhang (2012).

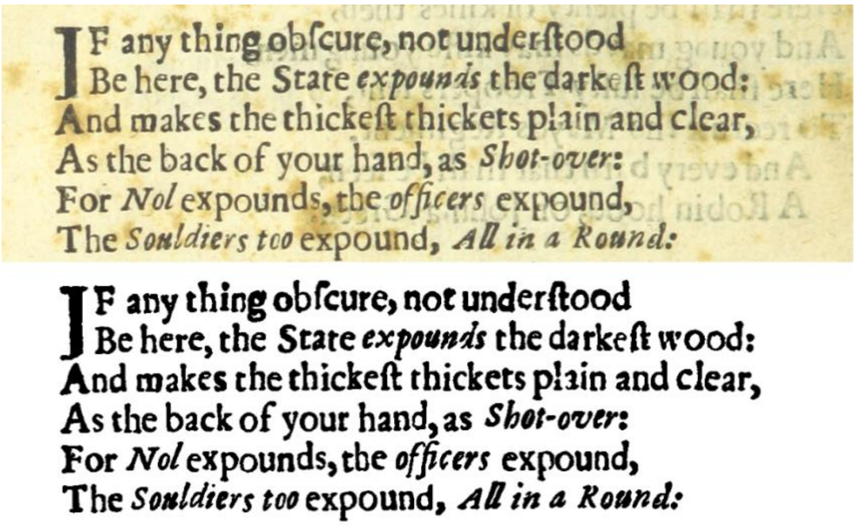
Cheng and Zhang (2012) were able to propose a technique to detect and hence prevent over-enhancement. The causes for over-enhancement were investigated in detail and a quite efficient criterion was proposed. They deduced from the experiments that their method can effectively locate the over-enhanced areas accurately and effectively and provide a quantitative criterion to assess the over-improvement levels well. The given method will be useful for vigorously monitoring the quality of the improved image and optimising the parameter settings of the contrast improvement algorithms. However, the above and former methods even when applied still have drawbacks of noise.

Rani et al. (2014) proposed a method for enhancing underwater grayscale images using a hybrid approach of stretching and filtering. Their research entailed improving underwater images affected by scattering effects due to the light absorption which can contribute to image blur. The approach showed significant improvement in subjective and objective parameters compared to proposed and other methods especially against previous methods of applying contrast and filtering separately. The proposed enhancement approach includes transforming the image through the enhancement pipeline of contrast equaliszation then applying homorphic filtering, wavelet denoising, bilateral filtering and finally contrast stretching. From the experiment, this method would not be suitable for certain artefacts in the dataset like show-through.

Boudraa et al. (2019) were able to propose a hybrid enhancement method using a combination of techniques to improve degraded historical documents. The paper describes a multiphase system that hybridises several effective image thresholding methods to achieve the best binarization output. The binarized image has the Contrast Limited Adaptive Histogram Equalization CLAHE algorithm applied to improve the contrast in particularly defective images with uneven disparity. Finally, a special transformation is inputted for purpose of removing scattered noise and correcting character form. From the experiment, it was denoted that the framework performed better than earlier methods however CLAHE has a reputation for over-enhancement which may not be suitable for certain degradations in the dataset collection.

2.5 Binarization (Post Processing)

In the field of image processing, binarization is known as any technique that converts an image of several bits' depth into only two bits of depth. In other words, it turns a grayscale or colour image into a black and white one. This approach is applied when there is a need to separate the background from some objects of interest. It is a widespread technique used as a prior step before further processing. For example, in document image analysis, binarization takes place when it is required to separate the background from the characters (Bonny & Uddin, 2019). It has been broadly proven in (Pratikakis et al., 2017) (He & Schomaker, 2019) how efficient this procedure is before character recognition. Most OCR systems require document binarization as the pivotal first step. When it comes to degraded historical documents, image binarization performance can indicate image quality. Figure 2-3 shows a typical example of degraded document image binarization.


IF any thing obſcure, not underſtood
Be here, the State *expounds* the darkeſt wood:
And makes the thickeſt thickets plain and clear,
As the back of your hand, as *Shot-over:*
For *Nol* expounds, the *officers* expound,
The *Souldiers too* expound, *All in a Round:*

IF any thing obſcure, not underſtood
Be here, the State *expounds* the darkeſt wood:
And makes the thickeſt thickets plain and clear,
As the back of your hand, as *Shot-over:*
For *Nol* expounds, the *officers* expound,
The *Souldiers too* expound, *All in a Round:*

Figure 2- 3 Binarization example of a degraded document image. Original image (up), after binarization (down). Taken from DIBCO 2011 dataset

In Figure 2-4 the basic architecture of document binarization is shown. As shown in the diagram, it is divided into two major steps, the pre-processing and post-processing step. The pre-processing step covers a variety of commonly used enhancement techniques which varies from contrast stretching, histogram equalization etc. The post-processing step consists of the main procedure of binarization called thresholding (core technique that converts the image into binary format). However, enhancement techniques are not always a part of all binarization methods as some methods do not include a prior enhancement because they work with different approaches to the problem.

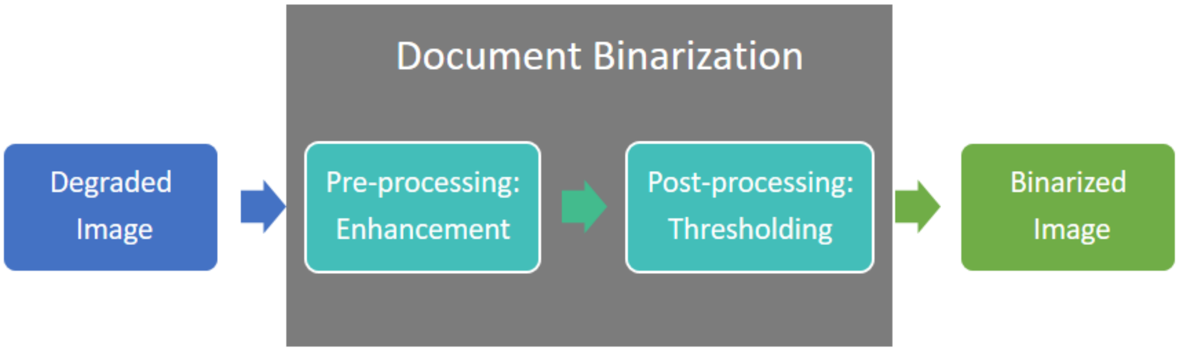


Figure 2- 4 Architecture of document binarization

2.5.1 Review of Binarization methods

A wide range of methods has been proposed in the past years to overcome the degradations and artefacts in the binarization process of historical documents. They can be classified depending on the nature of their techniques. Methods based on thresholding techniques are categorised in global, local and hybrid approaches. The global thresholding is when a single threshold value is used to separate the grayscales values of the whole image in two logical values and therefore obtain the binarized image. On the other hand, local thresholding is based on the calculation of threshold either from each pixel or a set of pixels, which usually depends on certain pixel limits in each object of an image. An image with multiple objects of similar pixels can be categorised into different classes and a set of thresholds can be calculated for each group of pixels locally in the image(Bonny & Uddin, 2019). Also, hybrid thresholding methods are those that attempt to unify the advantages of global and local techniques. Moreover, other recent advances in document binarization use machine learning techniques in a variety of ways to segment the foreground text from the background. This section reviews some well cited state-of-the-art binarization methods within the literature.

Otsu

Otsu’s method (Otsu, 1979) is a landmark of global binarization thresholding technique. It assumes that the image has a bimodal histogram and therefore executes an algorithm to compute an automatic global threshold to separate the image in two classes. The algorithm's goal consists of finding the threshold attributed to minimising intra-class intensity variance, or equivalently, by maximising inter-class variance. On that direction, the final threshold computed is an optimised grayscale value that separates the image in two groups of pixels as shown in Figure 2-5.

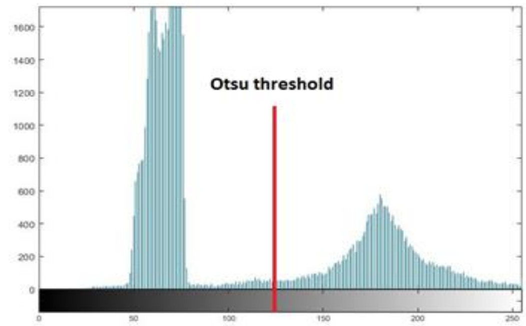


Figure 2- 5 Otsu global threshold example

Niblack's method

Another commonly used method of Binarization is Niblack's presented in (Niblack, 2003). This is a local adaptive binarization algorithm that uses a rectangular sliding window in which different thresholds are computed based on local mean and standard deviation. The threshold T for pixel $f(x, y)$ is the center pixel in a rectangular shifting window and is calculated in equation (2.1):

$$T(x, y) = m(x, y) + k * s(x, y) \qquad (2.\ 1)$$

Where $m(x. y)$ and $s(x, y)$ are the average and standard deviation respectively inside the rectangular region. The value of k manages the number of text areas. There is a trade-off for the value of k that prioritises local details for small windows but cannot be too small because it will not cover relevant objects. Usually, some authors recommend using a window's size 15x15 and $k = -0.2$. However, Niblack's method tends to fail when the background has a light texture.

Sauvola

Sauvola's method (Sauvola & Pietikäinen, 2000) is another standard among local thresholding technique. It consists of a very similar approach to Nicblack's but differs in how the local threshold is calculated. He proposes the equation (2.2) in which R is the dynamic range of standard deviation. Common values of k = 0.5 and $R = 128$. Experiments have shown remarkable results using Sauvola for document binarization owing to its capacity to manage variation in illumination, resolution, variation, and noise. Nevertheless, it tends to perform badly in very light and notably dark background.

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{s(x,y)}{R} - 1\right)\right] \tag{2. 2}$$

Bernsen's method

Bernsen's method (Bernsen, 1986) is another local thresholding algorithm that uses the mean and contrast information for computing the threshold over a region. The mean is computed using the highest and the lowest grey levels Z_{high} and Z_{low} respectively in equation (2.3).

$$T(x,y) = \frac{Z_{high} + Z_{low}}{2} \tag{2. 3}$$

Also, it measures the local contrast $C(x,y)$in the area as defined in equation (2.4):

$$C(x,y) = Z_{high} - Z_{low} \tag{2. 4}$$

If this local contrast is less than a predefined value, then the region is a single class, foreground, or background. Large text areas may take place in the image document, in that case, the pixel is taken as background. This method uses a window size of $t = 15$. The main disadvantage of this method consists of the production of a huge amount of noise in degraded historical documents.

Bradley's

Bradley's method (Bradley & Roth, 2007) is a local adaptative technique which is very simple but with high robustness against strong illumination variations. The idea behind this technique lies in its capacity to compute a unique threshold for each pixel in the image. The author claims that this method is similar to Wellner's algorithm (Wellner, 1993) and outperforms it. It is based on the computation of integral images and first-order statistics in a neighbourhood, making the method suitable for real-time applications on a live video stream.

First, the integral image also known as a summed-area table is computed. At each location of the integral image $I(x, y)$ is stored as the sum of all $f(x, y)$ terms to the left and above the pixel (x,y). This representation allows computing effortlessly, the sum of any rectangular array in the image given the upper left corner and the lower right corner.

Secondly, the average is computed in a window of pixels centred around each pixel using the integral image and then carrying out a comparison. The comparison consists of reducing the value of the current pixel (if it is less than the average) to black, otherwise, it is set to white.

In the literature review, there are not sufficient experiments with Bradley's method in historical document binarization. Therefore, it is considered in the tests as a remarkable approach in state-of-the-art experiments review.

Dilu

Dilu's method (Lu et al., 2018) is a recent technique that applies different contrast enhancement for areas with different contrasts before applying local thresholding binarization. First, the contrast image is computed to serve as a basis for dividing areas. The whole image is divided into non-significant areas, significant areas, and comparatively significant areas through two division levels as shown in Figure 3-6. The first level division is called coarse region division and consists of dividing the image into four regions A, B, C and D. Afterwards, each region is classified into non-significant or significant according to certain criteria based on contrast variance. For non-significant areas, the grey values of pixels within this area are set to 255 corresponding to the white background. Later, the second level division or fine region division is applied. This consists of four-divisions over the remaining regions in which other criteria are used to categorise the regions to non-significant again or significant or comparatively significant. In this division, if for example region A was not classified as a non-significant area then it is divided into AA, AB, AC, and AD regions as the figure shows. For those regions classified as significant areas and comparatively significant areas have a weak contrast enhancement and strong contrast enhancement respectively. Therefore, a contrast enhancement application is performed depending on the contrast variation in each region. Finally, a local threshold is applied to each area. The criteria that decide how to classify each region depend on two main parameters and which are later explained how were chosen.

According to the original paper, this method can effectively adjust the pixel grey values of an image with non-uniform illumination, bleed-through, and variable background. As a result, these three issues in image binarization can be solved.

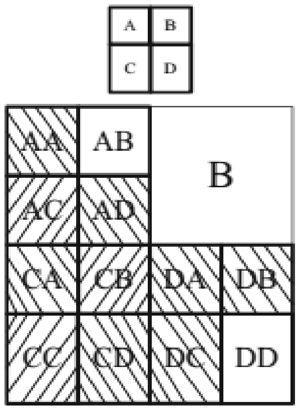


Figure 2- 6**:** Dilu's diagram division Empty box Background, box with right side stripe Areas with significant grayscale contrast, and box with left side stripe Areas with comparatively significant grayscale contrast

Howe

Howe's method was first presented in (Howe, 2011) and later improved in (Howe, 2013) This method first labels image pixels as background or foreground by minimising a global energy function inspired by Markov random field model. Secondly, in formulating the data-fidelity term of this energy it relies on the Laplacian of the image intensity to distinguish ink from the background. This detail allows an important invariance to differences in contrast and overall intensity. Thirdly, it adds edge discontinuities in the smoothness term of the global energy function, altering ink boundaries to align with edges and letting a stronger smoothness incentive over the rest of the image. Finally, this method explores in a very creative way a combination of techniques used previously in image binarization as motivation e.g., Markov random fields, Laplacian filter and edge detection based on Canny's detector. Also, a greatly advantageous feature consists of the automatic selection of their parameters and therefore, avoids the common unpleasant parameter tuning carried out by users. This method outweighs classical methods in which parameters must be selected.

Convolution neural networks

Convolution neural networks attain effective performances on various applications including document analysis. For instance, the front runner of the DIBCO 2017 event (Pratikakis et al., 2017) utilizes the U-Net convolutional network architecture for improved classification of pixels. In Tensmeyer's method (Tensmeyer & Martinez, 2017), entirely convolutional neural network is implemented at multiple image scales. In (Calvo-Zaragoza & Gallego, 2019; Peng et al., 2017), a deep encoder-decoder architecture is utilised for document image binarization. In (Vo et al., 2018), a hierarchical deep supervised network is implemented for document binarization, which attains state-of-the-art performance on numerous benchmark datasets. Grid Long Short-Term Memory (Grid LSTM) network is used for binarization in (Westphal et al., 2018) but exhibits a lower performance compared to Vo's method (Vo et al., 2018). In (He & Schomaker, 2019), the basic U-net neural network was used to learn degradations in document images from the Monk system as well as other benchmark datasets, they proposed a novel method for document enhancement and binarization based on iterative deep learning. The method uses a sample patch from an image to repetitively predict the uniform image through recurrent refinement or stacked refinement.

Jemni et al. (2022) in their paper titled Enhance to read better: a multi-task adversarial network for handwritten document image enhancement, proposed an architecture for handwritten document binarization based on Generative Adversarial Networks. The method recovers the degraded images while conserving their readability by integrating a Handwritten Text Recognition to evaluate the enhanced image in addition to the discriminator.

The challenge from reviewing the above-stated CNN methods is that although they can improve binarization substantially, the visual enhancement underperforms with errors that occur due to a lack of effective training data. Some of these anomalies from the results include the dominant background noises and the missing thin or weak strokes. Some other challenges include overfitting of the images causing the model to perform well on training data but generalize poorly due to new unseen data. Some biases also inherited in the training data have also impacted the results.

2.6 Problem/ Limitations for the Existing Literature

In summary of the literature review, there are many enhancements and binarization methods that can perform well with a particular type of degradation but enhancement/ binarization method that can work for multiple degradation are either limited or left for future work.

While some methods excel in specific aspects of document enhancement, there is a need for comprehensive approaches that address multiple facets simultaneously. Achieving a balance between image quality, OCR accuracy, and layout optimization remains a challenge. Image enhancement and subsequent binarization are important step towards the development of a system for document image recognition and it has a wider application towards digitization. The futurist approach should focus on the time and accuracy along with the global method to deal with various kinds of issues associated with prehistoric documents. One of such recommendation can be to improve the quality of degraded document prior to performing binarization and in doing so, an image enhancement method that works with multiple degradation can certainly help in the future(Sulaiman et al., 2019; Xiong et al., 2021).
This gap in the literature inspires this research to facilitate and develop an enhancement method that works for various degradations and achieves a balance between image quality, OCR accuracy and layout optimisation.

2.7 Existing Methods Selection

From the literature review, 2 classical and 4 recent methods of document image enhancement are selected to compare with the proposed enhancement method. The classical methods chosen are CLAHE(Reza, 2004), and Histogram Equalization (HE). The classical methods were chosen due to their pixel distribution qualities in the spatial domain and histogram equalisation and manipulation which this work is related to and the availability of source code.

The modern state of the art methods chosen for comparative analysis are Adaptive Contrast Enhancement using Modified Histogram Equalization (ACMHE)(Santhi & Banu, 2015), Sun et al. (2016) method for Blind bleed-through removal for scanned historical document image with conditional random fields, Bio-inspired Modelling for the Enhancement of Historical Handwritten Documents (BMEHHD) (Zagoris & Pratikakis, 2017), Gray Level Enhancement using Barnacles Mating Optimizer proposed a meta-heuristic algorithm using BMO (Ahmed et al., 2020), Enhancement of image contrast using Selfish Herd Optimiser(Guha et al., 2022), and Contrast Enhancement of Degraded Document Image using Partitioning based Genetic Algorithm(Ghosh et al., 2019) which are not based on HE method. The selection of these methods was based on how relatable they are with the proposed method in terms of foreground and background intensity manipulation and the pre-processing enhancement methodology to improve subsequent binarization. Regrettably, among the selected methods for comparison, only the methods ACMHE (Santhi & Banu, 2015), Sun et al. (2016) blind bleed-through removal for scanned historical document image with conditional random fields (BBTRSHD) and Gray Level Enhancement using Barnacles Mating Optimizer (BMO)(Ahmed et al., 2020) had their source codes for implementation available. Other efforts to contact the authors of all other relevant literature for the implementation codes were unsuccessful at the time of writing up the research.

2.8 Summary

From the literature survey carried out, it was deduced that existent enhancement methods have a range of little to significant drawbacks when applied to multiple degradations present in historical documents datasets as several of them are designed for specific problems. However, from the different methodologies studied in the surveys, a novel combination of contrast stretching, noise filtering and a state-of-the-art **adaptive histogram matching** technique which is novel and hasn't been used in previous literature at the time of writing this report is proposed. The algorithm was structured to adapt optimally to the prevalent historical document degradations like faded ink, bleed-through, uneven illumination etc. with diminutive drawbacks. The anticipated advantage of this method over existent literature is its ability to perform on several drawbacks like bleed-through, uneven ink, faded ink etc., and not just a specific artefact.

CHAPTER 3

RESEARCH METHODOLOGY AND TOOLS

3.1 Overview

This chapter describes the research approach that has been adopted. It explains the steps that are taken to achieve the research aims and objectives. The research is structured around quantitative and qualitative research aimed at algorithmic experiments, hypotheses, proven research theories and computational evaluation. To ensure a much more robust argument, qualitative evaluation using human opinion score is also carried out. The first part of this chapter explains the background of the ENP dataset showing a brief review of the challenges in terms of degradation/ artefacts. The second part deals with the methodology and subsequently, the background of the tools utilised in the enhancement method.

3.2 Dataset

The Document Image Binarization Contest (DIBCO) (DIBCO, 2019) datasets are a well-known collection of several historical document images used in a series of international binarization contests. These datasets comprise a small subset of document images machine-printed and handwritten per year with associated binarized ground truth. However, this benchmarking dataset has drawbacks such as lack of artefact classification, number of images and text ground-truth for OCR engine recognition. For this reason, the PRImA (Primaresearch, 2020) library catalogue containing the Europeana Newspaper (ENP) dataset collection can be a valuable option for researchers. This selection is primarily due to the research experimental need of having better algorithm performance analysis due to the availability of text ground truth for OCR experiments.

The Europeana Newspapers Project (ENP), in which PRImA participated, involved the collation and enhancement of historical newspapers assembled from the European Library and Europeana. The project was EU funded to achieve a vast amount of searchable historical newspapers through the most impactful European cultural heritage websites, and the Europeana Library. The project was embarked upon to improve the search and presentation process for readers and users. The project had the goal of creating quality estimation toolkit to support future digitisation and OCR developments in the decision-making process(Clausner et al., 2016). The resulting collection from this comprehensive project is numbering over 11 million scanned newspaper pages accompanied with OCR text. All page images in the dataset are either 300dpi or 400 dpi quality and are a wide-ranging distribution of grayscale, bitonal and colour pages. The images are saved as TIFF files with lossless compression to enable the perfect reconstruction of compressed data irrespective of the source files. The newspapers are of great value and interest to researchers and the general public interested in European heritages dated since World War I, they are also in easy-to-understand text not requiring advanced knowledge(Clausner et al., 2015). Considering the project's requirement for assessing performance and ensuring quality, and considering the accessible data and resources, the authors aimed to establish a lasting, representative, and extensive dataset. This dataset is intended to extend beyond the project's duration and serve as a valuable resource for researchers involved in document analysis and recognition. The ENP derived dataset in the PRImA database is a structured subset comprising of 528 images organised by century, colour depth (bitonal, greyscale and colour), language and popular artefact keywords and groundtruth. The ENP dataset was created in the view of the need of the project for performance evaluation and quality assurance for researchers in document analysis and recognition. This qualities of the ENP dataset are the primary motivation of choice for this research.

However, like most historical document collection sets, there are certain degradation and quality challenges that have been documented in previous works in the literature. However, classifying the various problems, ranging from poor image capture to poorly preserved documents themselves, is where the PRIMA datasets (including the one from the Europeana Newspapers Project) stand out from other datasets.

This is the first time in which several state-of-the-art methods are going to be tested in the PRIMA ENP dataset to see how they affect/ improve the resultant images. This can be a challenge due to variations and high-quality large file sizes. The suggested solution in this research is a smaller subset with cut out patches of whole images and their ground-truth identifying the challenges for easier representation, experimentation and faster processing.

3.2.1 Challenges in PRImA Dataset images

Historical documents are the basis of obtaining meaningful cultural and scientific knowledge that can be used for information retrieval. These documents are available in reference centres of several government departments or libraries with high valuable importance regarding cultural, scientific, and legal aspects. Nowadays, there is a growing interest in transforming these documents into digital format (Clausner et al., 2017) as it makes them more readable and accessible.

Nevertheless, in general, historical documents have a variety of artefacts referred to as document degradations. These degradations can be corrected with the help of document enhancement to improve binarization results further, enabling better extraction of accurate information using the OCR engine tool (Clausner et al., 2020). In particular, the PRImA dataset containing the Europeana Newspaper collections has a wide range of degradations making this dataset a very challenging collection. However, a crucial feature of a search tool inside the PRImA dataset allows the selection and organisation of documents by a variety of characteristics including degradation types and artefacts. Consequently, it is possible to test algorithms' performance for each type of degradation. This option significantly enables advancement in the recent state of the art in contrast to the well-known DIBCO dataset (DIBCO, 2019). In the following section, the typical degradation examples in the PRImA dataset are depicted and explained.

3.2.1.1 Typical degradation and artefacts

In the ENP collections, there are a wide variety of degradations or artefacts in images presented at the time of selection. In this research, the artefacts suggested in the literature(Sulaiman et al., 2019) from the PRImA dataset are considered. These artefacts are smeared ink, low scan contrast, faded ink, uneven illumination, broken characters and show through. In the following paragraph, examples of these challenges of low-image quality documents are discussed.

- **Uneven Illumination**

The undesired effect of uneven illumination occurs when illumination appears distributed in a non-uniform way in the final digitised image, with some areas brighter than others due to a wrong acquisition image system or detrimental illumination conditions. This leads to difficulties in documents image analysis (Van Kempen et al., 1997), especially when it comes to Human Visual System and OCR results. Without robustness in front of uneven illumination, the HVS and OCR cannot extract text properly for character recognition, Figure 2-1 shows some examples seen from the ENP.

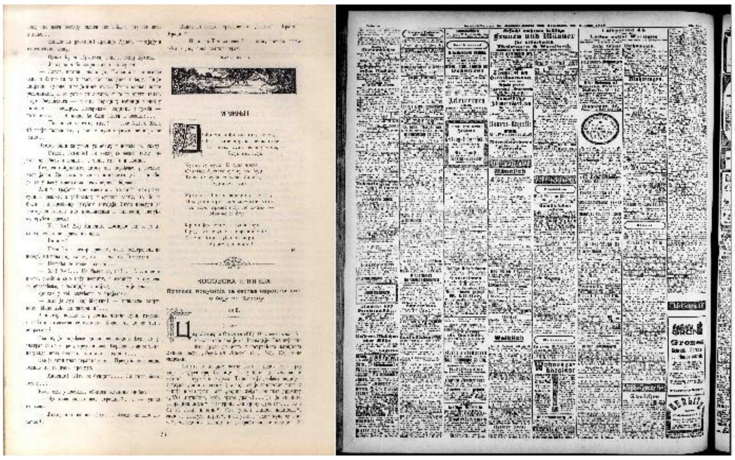
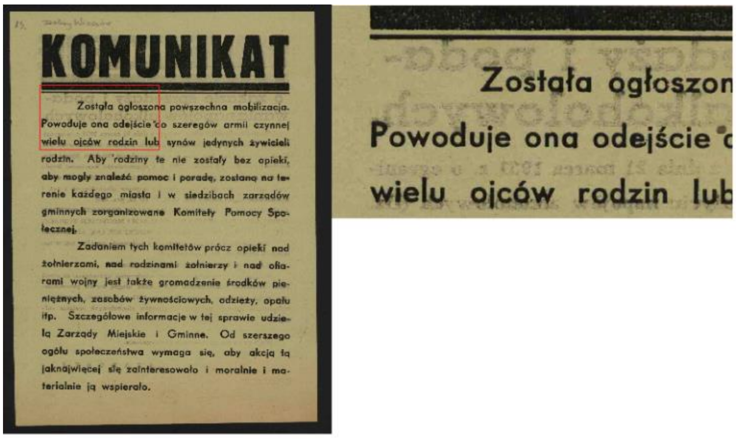


Figure 3- 1 Uneven illumination examples from ENP dataset a) 00675813 image b) 00674388 image

The figure 3-1 above is a typical example of the datasets found in the ENP dataset. The far-left side as shown in image (a) of the figure has a lower font intensity than the right side thereby obscuring OCR systems and the Human Visual system. The effect means the whole section on the document becomes unreadable.

- **Show-through**

Show-through artefact is also known as ink-bleed degradation. This artefact occurs when the ink from one side starts to appear on the other side of the document and depicts the words as if written on two sides of the page. This is because in the past some documents were written on both sides and show through with time (Xu et al., 2012). In document enhancement and binarization, this is a great challenge as the HVS, and OCR could misclassify ink from the other side as if it were written on the main side as shown in Figure 3-2.


KOMUNIKAT

Została ogłoszona powszechna mobilizacja. Powoduje ona odejście do szeregów armii czynnej wielu ojców rodzin lub synów jedynych żywicieli rodzin. Aby rodziny te nie zostały bez opieki, aby mogły znaleźć pomoc i poradę, zostaną na terenie każdego miasta i w siedzibach zarządów gminnych zorganizowane Komitety Pomocy Społecznej.

Zadaniem tych komitetów prócz opieki nad żołnierzami, nad rodzinami żołnierzy i nad ofiarami wojny jest także gromadzenie środków pieniężnych, zasobów żywnościowych, odzieży, opału itp. Szczegółowe informacje w tej sprawie udzielą Zarządy Miejskie i Gminne. Od szerszego ogółu społeczeństwa wymaga się, aby akcją tą jaknajwięcej się zainteresowało i moralnie i materialnie ją wspierało.

Figure 3- 2 Image 00762016 with show-through artefact

- **Smeared Ink**

This is a typical problem presented in a wide variety of historical documents. Smeared ink occurs when ink that has not been properly absorbed into paper is smeared over the document due to contact, this usually obscures other characters of the image thereby affecting legibility. It can occur in different areas of the document, making it less readable and leading to binarization algorithms failure. Figure 3-3 shows a typical example.

Figure 3- 3 Image 00674680 with smeared ink

- **Faded Ink**

Some documents that are written by hand could over time cause some characters ink to fade. This also applies to machine printed documents that were printed with old ink that were less absorbent and lost intensity over time. Also, due to the machine-driven nature in those times, some characters may appear with more ink intensity than others due to the typing process. Figure 3-4 illustrates an example of faded ink in some characters.


GROOT B

Denburg den 16 Dec
Loudon bedragen o
is commandant e
Schotland geworden. De h
chef het corps der rebell
kanon en eenige veldstuk
meld dat de bergheden m
zyn ; dat zy daegs te voo
dog dat hun waere deffe
taften. Een fransch tra

Figure 3- 4 Image 00673665 with fading characters

- **Low Scan Contrast**

Another common artefact is the low contrast in some document images. The contrast in this application context is defined as the difference between characters with low intensity and high-intensity pixels associated with background. When contrast is low, a grey tone is prevalent in the image and therefore reading may be complex. This issue is attributed to the acquisition of technology and environmental factors like poor illumination. Figure 3-5 shows two examples of a low-contrast document image. It is easy to perceive the prevalence of grey tones in both images.


Jons Gutenbergs
un jo izgudrojums.

Figure 3- 5 Image 00674642 (left) and 00674654 (right) with Low Scan Contrast

- **Broken Characters**

In this research, some document images with broken characters were collected. The ENP provides many images with this keyword. Although like the faded ink artefact, the main difference is that some character's shapes could be cut off completely or disappear. This happens because of the nature of the typing machine used to enter the information. Figure 3-6 shows an example of a document image with broken characters.

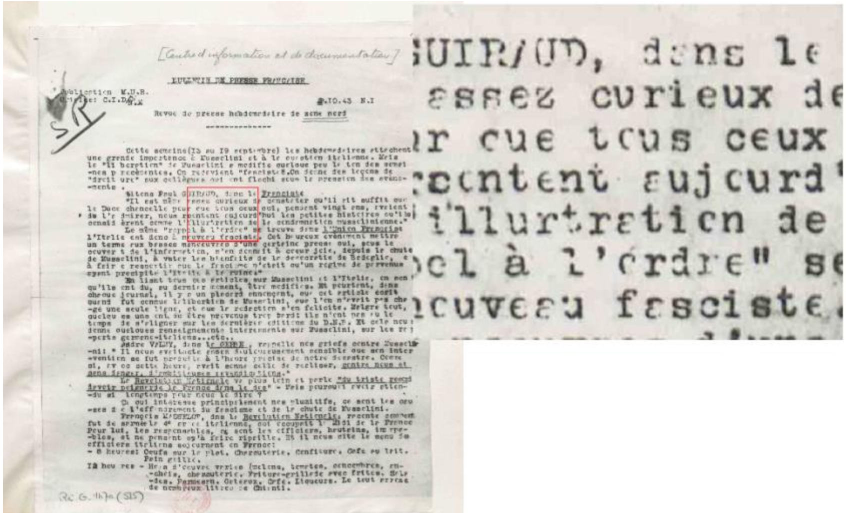


Figure 3- 6 Image 00675211 with broken characters

Summary

This chapter has thus far given a brief background of the ENP dataset, and a review of the challenges and degradation/ artefacts contained in it. The remainder of the chapter will discuss various knowledge sources, procedures and algorithms involved in historical document image enhancement.

3.3 Subset Derivation

The need for performance evaluation, quality assurance and availability of data and resources inspired the creation of a subset of the ENP representing the quality of the dataset for easier experimentation. These artefact subsets are carefully selected from the ENP dataset in the PRImA extensive historical document library to organize a structure on how the algorithms perform on subsets with their specific challenges and attributes. The organisation of these subsets also provides the opportunity to test well-known enhancement and binarization methods and see how they perform with the recent state-of-the-art enhancement and binarization techniques.

The approach taken to build the subset necessary to carry out this research using best practice from previous innovative dataset creation (Antonacopoulos et al., 2009; Clausner et al., 2015; Papadopoulos et al., 2013) is considered using three main characteristics:

- Realistic: The subset must represent a cross-section of real documents likely to be scanned, maintain the representativeness of the individual dataset as much as possible and be representative of artefacts needing enhancement.
- Comprehensive: The subsets must contain detailed information containing metadata and ground-truth to enable detailed evaluation.
- Flexibly structured: It should be structured to allow access to external systems of evaluation and OCR.

With respect to the need for performance evaluation based on artefact type, quality assurance, availability of data and resources. This research classifies a subset of the comprehensive ENP dataset that can be useful for experimentation by researchers in image degradation analysis.

The ENP is a large comprehensive newspaper dataset with ground truth. This published dataset contains 528 images that are either 300dpi or 400dpi and a broad distribution of grayscale, bitonal and colour pages. There is no such dataset that focuses on newspapers and ground truth at this scale as at the time of production. This makes it a good selection for the experiment and research. The closest dataset of similar nature is that of the IMPACT project (Papadopoulos et al., 2013), which focuses mostly on books.

Subset

The subset was created in two stages as detailed in the preceding subsections:

1. Representative image aggregation
2. Refine the selection to a realistic subset.

```
Start

|-> Image Aggregation
|   |-> Image Selection
|   |   |-> Identify artefact types in representative titles
|   |   |-> Consider artefact occurrence frequencies
|   |   |-> Use Prima keyword selection tool for classification
|   |   |-> Select images with highest artefact presence
|   |   |-> Classify images with multiple artefacts as strongest degradation
|   |   |-> Allocate project-unique IDs (ENP)
|   |   |-> Convert images to standard format with lossless compression
|   |   |-> Examine and sort images into individual folders
|   |   |-> Use basic metadata for indexing and grouping
|   |   |-> Group images based on metadata (title, language, script, source, etc.)
|   |   |-> Make image conversions (cropping) for emphasis on artefacts
|   |
|   |-> Subset Selection
|       |-> Restrict image selections by artefact type
|       |-> Maintain dataset representativeness
|       |-> Limit images to 28 per institution
|       |-> Select a total of 300 representative images from ENP dataset
|       |-> Ensure sufficient dataset quantity for research (average text lines/page)
|
End
```

Figure 3- 7 Flow chart of subset creation

I. Image Aggregation

The aggregation process for creating subset as shown is clearly defined as follows. The first step was image selection, where the images were identified by artefact type in representative titles from the larger subset. This action was carried out with consideration of what type of artefacts and their occurrence frequencies in digitised datasets. It was however clear that even though the selection would be mostly from already digitised material, it should be representative of the artefacts to expect in any given library. This, in most cases, meant that selections had to showcase representative prevalence of the degradation in the dataset. The postulated approach is to classify the ENP images using the Prima keyword selection tool to specify the degradation characteristics desired for the research. It was decided that the images collected will be the images with the highest artefact presence, and where they were more than one artefact present which can be expected in several images, the image will be classed as the strongest degradation present.

The collected images were examined and sorted into individual folders. The images have their project-unique IDs already allocated in the ENP which helped to make the classification for the subset easier, all images were also already a conversion in standard image format with lossless compression for smaller file size and easier analysis. Although the ENP dataset is a collection of versions with the best image quality possible or as close as possible from the Europeana libraries, this research would show the need for further enhancement even in cases where the best selections are collated in libraries.

The subset classification in this research did not rely on random image selections but on significant time and effort to manually select which titles, issues, and pages to include for a representative selection. The image selection was also grouped with their individual Metadata present in the ENP dataset. This basic metadata formed the basis for indexing the images when uploaded to an online repository, the metadata also contain details of the images which include title, primary language, primary script, original image source. Some optional information in the metadata include scanner model, image artefact etc. These unique resources were examined and grouped in this research into folders showing artefact type, image conversions were made for some of the images by cropping the images on the pages to emphasise the artefact and show how the enhancement performs to the smallest details.

II. Subset Selection

The sub-selection of the ENP dataset to form the subset was driven by two main factors:

- To restrict the image selections further down to conform with the research requirements of dataset sortation by artefact type.
- To maintain the representativeness of the dataset as far as possible.

Thus, it was determined to limit the number of images to 28 per institution. Thus, a total of 300 images were selected as a representative subset of the ENP dataset. This number gives a sufficient dataset for this research, because of the quantity of information on each page as compared to books, with an average of 383 text lines on each document. The subsets were independently verified by the supervising team of this research.

Table 3. 7: PRIMARY LANGUAGE CLASSIFICATION

Language	No. of Pages	Language	No. of Pages
Dutch	11	Polish	21
English	28	Russia	3
Estonia	28	Serbia	28

Finnish	18		Swedish	11
French	28		Ukrainian	2
German	96		Yiddish	2
Latvian	23			

Table 3. 8: PUBLICATION DATE CLASSIFICATION

Publication Period		**Number of Pages**
17^{th} Century		3
18^{th} Century		7
19^{th} Century		106
20^{th} Century	1900-1925	88
	1926-1950	95

The selection of languages, scripts, title pages, publication-period and artefact type were sustained as close to the original selection as possible in representing the ENP dataset. Table 3.8 depicts the selection with regards to language distribution, and Table 3.7 shows a summary of the dataset per publication date. This subset page images are either 300 dpi or 400 dpi TIFF files with a colour depth distribution of 28% bitonal, 48% Grayscale and 24% Coloured. The most occurring artefacts and other page characteristics are shown in Figure 3-8.

Ground truth Selection

The ground truth utilised in this project was derived from the existing selection in the primaresearch repository(Clausner et al., 2015). Ground truth refers to the ideal outcome in an optimal OCR (Optical Character Recognition) workflow. It is pivotal in assessing document analysis methods against what would be deemed the correct result. Generating ground truth is usually a manual or semi-automated task due to the imperfect state of current OCR engines, particularly with historical documents. To expedite and ensure high-quality and consistent outcomes (targeting 99.95% accuracy), the creation of the ENP ground truth utilised in this research was outsourced to commercial service providers. This involved producing specific ground truth elements: precise region outlines, region type labels, full text (Unicode encoded, encompassing symbols and ligatures), and reading order. These ground truth files adhered to the PAGE (Page Analysis and Ground Truth Elements) format, recommended by the IMPACT Centre of Competence in Digitisation and utilized in numerous EU projects.

To enhance efficiency, service providers received preliminary-processed OCR output files in PAGE format. They were granted the option to either rectify or manually create ground truth based on the provided material's quality. Additionally, a customized version of Aletheia, a widely used semi-automated ground truth production system, was provided to the service providers. Detailed instructions on interpreting and representing specific content elements were also given. For quality control, a three-stage process was established. Initially, the ground truth validator, a system implemented by the authors, conducted automated checks against programmatically verifiable ground truth rules. This included verifying text presence in all regions, ensuring regions were in the correct reading order, and detecting overlapping region outlines.

Following successful automated checks, manual quality assurance was conducted, focusing initially on layout-related elements such as region outlines, type labels, and reading order. Once layout approval was attained, the files were forwarded to the respective content provider (library) for text verification. Minor issues were promptly rectified during quality control, while more substantial deficiencies were referred to the service provider.

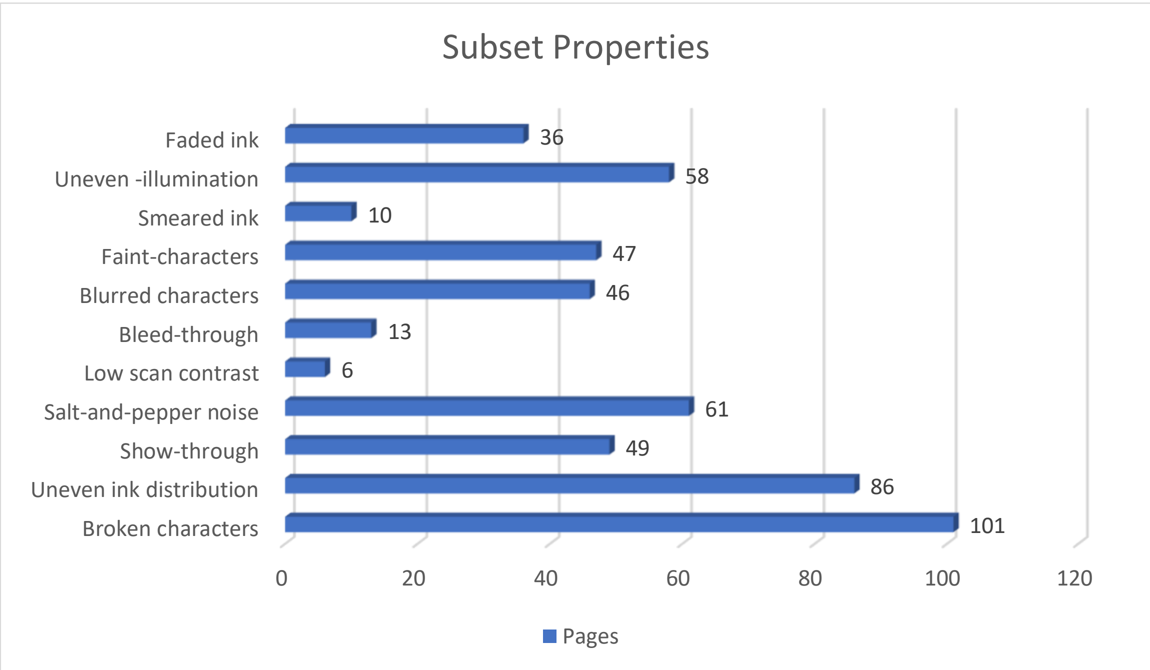


Figure 3-8 Most frequent image artefacts

Each classified image in this subset is also grouped with essential metadata and ground-truth for easy experimentation and evaluation. Researchers will be able to use the ground-truth for testing binarization and the metadata to test OCR performance.

3.4 Methodology

In this section of the report, the procedures, and overall methods to be applied for this research project are described. The research stages are divided into four segments as shown in figure 3-9 namely, *Design, implementation and evaluation*.

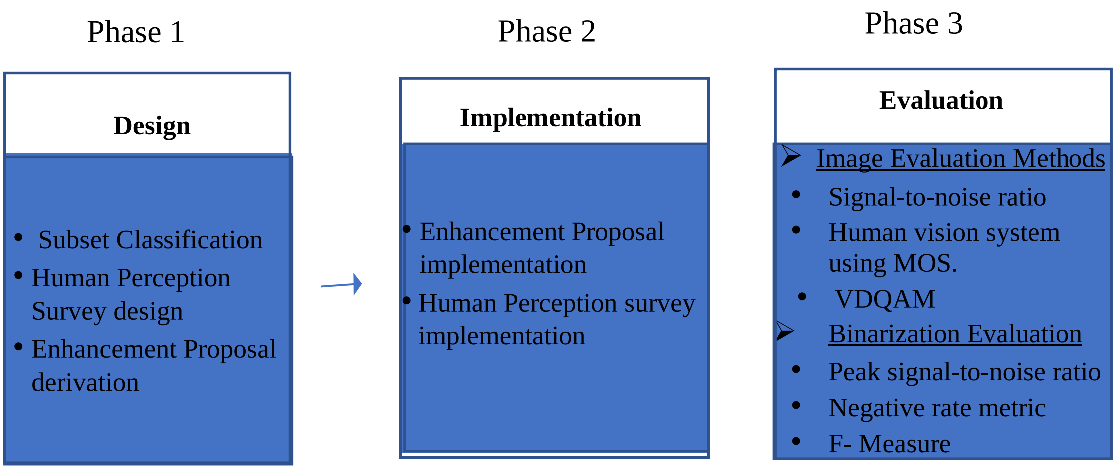


Figure 3-9 Research Plan

At the initial phase of this research, an in-depth review of related literature in the research area is continually carried out. This process gives an insight into the present state of research on the chosen area identifying research questions about the topic that needs additional research and ascertain methodologies utilised in past studies of similar areas of research. So far, from the works of literature reviewed, it has been identified that there are many challenges in enhancing historical document images compared to the enhancement of general document images. One of which is the absence and limitations of a proposal method that enhances visually while also improving binarization results. It is also revealed that the different approaches for image enhancement have their separate limitations, therefore this research intends to confer an adaptative enhancement method that can be applied to most artefacts.

Image enhancement methods are classified into two wide-ranging categories: Spatial domain methods and frequency domain methods. The spatial domain method generally deals with the direct manipulation of pixels in an image on the image plane itself. The frequency domain, however, deals with modifying the Fourier transform of the image as shown in figure 4-3.

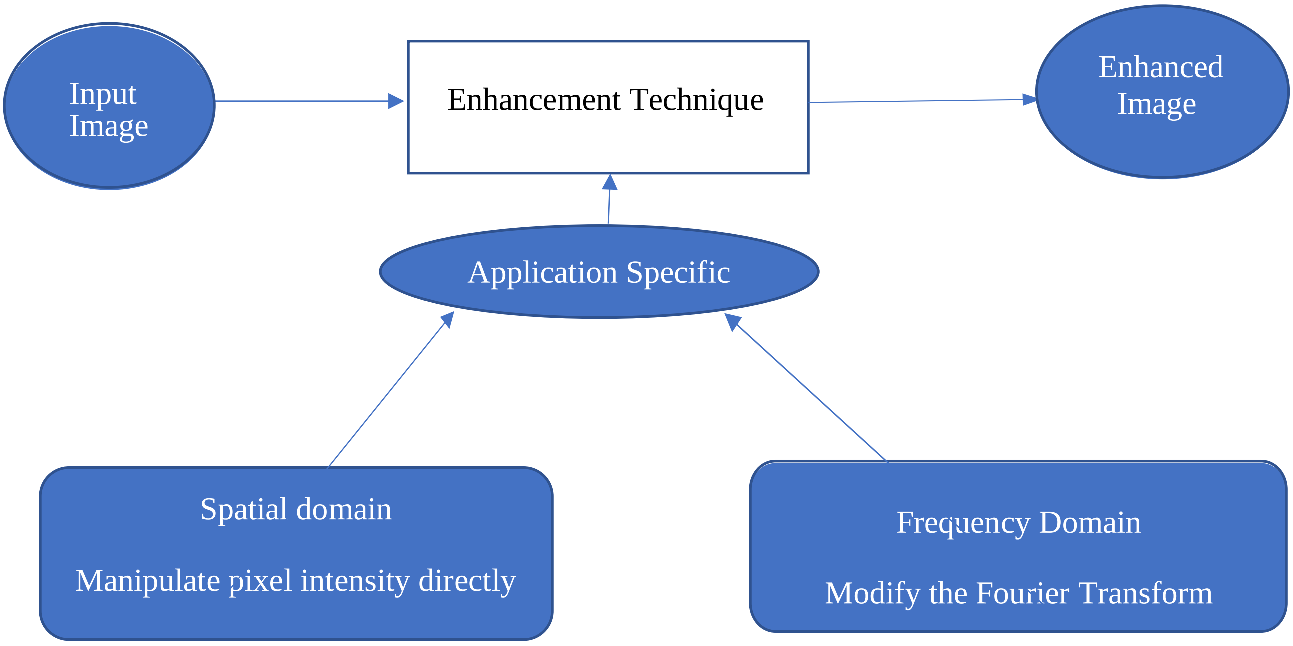


Figure 3- 10: Image enhancement Technique

However, the research focus is on reviewing and finding the gap in methods in the spatial domain, using direct manipulation of the pixel intensities.

Design: The Framework of this research as shown in the methodology in Figure 3-9 includes extraction and manipulation of historical document image datasets (Europeana Newspaper Project Dataset) from PRImA to improve clarity, visibility and overall quality for improved human viewing and OCR recognition of digital document library. An elaborate way of conducting human vision experiment is designed using on google form using the subset images. Human perception surveys allow for a holistic evaluation that considers multiple factors simultaneously, such as image clarity, readability, and overall visual appeal. This comprehensive understanding is crucial for image enhancement methods that aim to improve the overall quality of the visual content. Human perception surveys provide a valuable subjective evaluation of image quality. While objective metrics can quantify certain aspects, human perception captures the nuanced and subjective aspects that automated metrics may miss.

The proposed image enhancement method is a novel system comprising of Contrast stretching, Wiener filtering, Bilateral filtering, Adaptive histogram matching and Median filtering developed and implemented to analyse and manipulate the digital images for quality improvement. The proposed enhancement stands to achieve results that would guarantee higher image quality and reduce OCR challenges in the common artefacts due to some factors. These factors from the literature include the contrast stretching characteristics of maximising dynamic range through the desired range of values, the wiener filter characteristic of removing blurring noise which historical documents are well known to contain. Furthermore, bilateral filtering smoothens the images while preserving the edges before introducing the novel adaptive histogram matching technique to create the ideal histogram for an optimized image. The final part of the method which is the median filtering is a post operation step to eliminate any salt and pepper noise that may be present after the adaptive histogram process. The pipeline is structured in this order after several experiments to determine the most efficient and effective model order for a desired result. When the system is placed in any other order or in the omission of any part of the model, the system underachieves.

To the best of my knowledge, there has been no implementation of the technique in this structure. The true novelty of this method is the adaptative Histogram matching which is a first of its kind in enhancement. It is described in detail in the subsequent chapter.

The entire working procedure is strategically structured into several steps. The first step comprises of classifying a sub-selection representative dataset (subset) of the entire ENP dataset in the PRImA database for simplicity, effective testing and to save computing time and resources while experimenting and evaluating the various algorithm performances. This is achieved by grouping the dataset by their various relevant deterioration types e.g., faded ink, low scan contrast, show through, smeared ink, uneven illumination, etc.

Implementation: The second step of implementation comprises of an up-to-date evaluation of the impacts of a set of classical and state-of-the-art document enhancement techniques and applying them to the created ENP subset to evaluate their performance and make recommendations on how to improve the quality of the images accordingly. Furthermore, the implementation of the enhancement proposal is compared to the classical image enhancement methods CLAHE, Histogram Equalization, BMO, BBTRSHD and ACMHE. These methods were chosen to consider the distribution of the histogram as a key factor in improving the document quality. The aim is to turn the original histogram into a bimodal histogram and therefore adaptively improve the discrimination between the foreground (letters) and background (paper). This enhancement technique is configured to enhance historical documents without the problem of over-enhancement and loss of edges. As the research progresses, further evaluation of the proposed method is to be carried out using classical binarization methods Otsu, Niblack, Sauvola, Bernsen, Howe and the more recent Dilu method to evaluate how the proposed method affects binarization. In summary, qualitative, and quantitative analysis is to be conducted among the different approaches in terms of image quality assessment, word recognition (OCR) rate of the document samples and the image quality of the resulting text samples after experimentation using human perception survey experiment.

Evaluation: It is imperative to analytically examine and evaluate results from the developed technique. This substantiates whether the proposed method offers effective improvement as the goal of image enhancement is to improve an image so that the resultant image is better, more readable, and/or improve postprocessing procedure. This research proposes to evaluate the image performance through Signal to noise ratio, Human vision system using Mean Opinion Score, Visual Document Image Quality Assessment Metric (VDQAM), binarization and OCR improvement (post-processing).

At the end of this research journey, a comprehensive PhD thesis would be formed. The thesis would capture all the relative contributions, information and recommendations achieved. This includes the performance of the method in comparison to other state of the art and recent methods.

CHAPTER 4

THE EVOLUTION OF ADAPTIVE HISTOGRAM MATCHING

The current state-of-the-art in historical document enhancement and development of the research framework shows considerable achievements as disclosed in Chapter 2. However, most effective historical document enhancement techniques still bring out image results that are either under enhanced, over enhanced, or not edge preserved in addition to being unsuitable for multiple degradations. This chapter introduces the background of the proposed pipeline of the method and demonstrates its working theory.

4.1 Process Background Description

This section introduces the background of the basic concept of the components of the proposed pipeline method including the state-of-the-art model. It also shows the equations of the working model and introduces the theories relevant to the techniques.

4.1.1 Contrast Stretching

Contrast stretching is an image enhancement technique that operates by stretching the range of intensity values throughout the desired range of values (usually the full range of pixel values allowed on the image). It produces a less harsh enhancement on images due to its linear scaling function(Winiarti et al., 2017). The low contrast of an image occurs due to several factors such as low lightning condition, low illumination of camera, background light or atmospheric conditions. Contrast stretching is one of the simplest and resourceful pre-processing techniques that increases the grey level of an image dynamically to normalize the pixel in its maximum range. For example, in an image with lower and upper limits of 0 and 255, if the image histogram does not cover the entire range, then contrast stretching can extend the pixels to the limits filling the gaps of the histogram thereby improving contrast and detail of the image. The stretching is represented as:

$$P_{out} = (P_{in} - c)\left(\frac{b-a}{d-c}\right) + a \qquad (4.\ 1)$$

Where,

P_{out} and P_{in} are the output and input pixels

a and b are the lower and upper limits respectively; a= 0 and b=255 for standard 8-bit grayscale

c and d are the lowest and highest pixel or intensity values currently present in the image.

Contrast Stretch is the simplest algorithm that stretches the pixel values of a low-contrast image or high-contrast image by extending the dynamic range across the whole image spectrum. Contrast stretching is only conceivable if the minimum intensity value and maximum intensity value are not equivalent to the possible minimum and maximum intensity values or else, the image generated after contrast stretching will remain the same as the input image. It is a valued pre-processing technique that applies the highest possible contrast variation which is valuable for handling historical document images and other image types (Ooi & Isa, 2010)

4.1.2 Wiener Filtering

The Wiener filter is a Mean square error ideal linear filter for improving images with additive noise and blurring. The computation of this filter is on the hypothesis that both signal and noise are second-order stationery. Wiener filters are incapable of reconstructing frequency components that have been degraded by noise. They can only suppress them. Also, Wiener filters are unable to restore components for which $H(u,v)=0$. This means they are unable to undo blurring caused by band-limiting of $H(u,v)$. Such band-limiting occurs in any real-world imaging system(Gatos et al., 2006). The Wiener filter is a combination of inverse filtering and noise smoothing which helps invert blurring while removing noise simultaneously. Wiener filter is very effective in mean square error improvement from the process of inverse filtering and noise smoothing. For this effectiveness, the Wiener filter is implemented as part of the image enhancement goal in this research. The Wiener filter is represented mathematically.

$$I(x,y) = \mu + (\sigma^2 - v^2)(I_s(x,y) - \mu)/\sigma^2 \qquad (4.\ 2)$$

Where μ is the local mean, σ^2 is the variance of an m×n neighbourhood around each pixel and v^2 is the estimated average of all estimated variances for each neighbourhood pixel.

4.1.3 Bilateral Filtering

Bilateral filtering operates on a principle that involves smoothing images while preserving the edges. The concept is to replace the pixel intensity of each pixel with a weighted average of intensity values from nearby pixels. Each neighbour is weighted by a spatial component that penalises distant pixels and range component that penalises pixels with different intensities (Gavaskar & Chaudhury, 2018). The combination of both components ensures that only nearby similar pixels contribute to the result. It is usually based on Gaussian distribution and is very effective in historic documents.

The idea starts with a low-pass domain filter applied to the image f(x)(Tomasi & Manduchi, 1998)

$$h(x) = k_d^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathcal{E}) c(\mathcal{E}, x) d\mathcal{E} \qquad (4.\ 3)$$

When low pass filtering is to preserve the dc component of low pass signals, the formula below is obtained

$$k_d(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e(\mathcal{E}, x) d\mathcal{E} \qquad (4.\ 4)$$

Range filtering is similarly defined as

$$h(x) = k_d^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathcal{E}) s(f(\mathcal{E}), f(x)) d\mathcal{E} \qquad (4.\ 5)$$

However, s(**f(ε)**, **f** (**x**) measures the photometric similarity between the pixel at the neighbourhood centre x. The normalization constant is replaced by

$$k_r(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(f(\mathcal{E}), f(x)) d\mathcal{E} \qquad (4.\ 6)$$

The next step is to combine domain and range filtering thereby enforcing geometric and photometric locality. Combined filtering is denoted as.

$$h(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathcal{E}) c(\mathcal{E}, x) s(f(\mathcal{E}), f(x)) d\mathcal{E} \qquad (4.\ 7)$$

After normalization

$$h(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\mathcal{E}, x) s(f(\mathcal{E}), f(x)) d\mathcal{E} \qquad (4.\ 8)$$

The combined and range filtering denotes **bilateral filtering.**

4.1.4 Histogram matching

This section explains the background of conventional histogram matching. The proposed adaptive histogram is described in detail in the implementation section 4.2.5.

Histogram matching process involves the modification of an image such that its histogram matches that of another reference dataset. Histogram matching is a smart and easy way to "calibrate" one image to match another. In mathematical terms, it is the procedure of transforming an image so that the cumulative distribution function (CDF) of values in each band matches the CDF of bands in another image. The general application of histogram matching is computed by considering a grayscale input image X, with probability density function $p_r(r)$, where r is a value in grayscale, and $p_r(r)$ is the probability of that value. This probability is calculated from the image histogram by.

$$p_r(r_j) = \frac{n_j}{n} \qquad (4.\ 9)$$

Where, n_j is the frequency of the grayscale value r_j , and n is the pixel number total.

For a desired probability density function output $p_z(z)$. A $p_r(r)$ transformation is required to convert it to $p_z(z)$.

The probability density function (pdf) is easily mapped to its cumulative distribution function by

$S(r_k) = \sum_{j=0}^{k} p_r(r_j),$ $\qquad k = 0,1,2,3, \dots, L-1$

$G(z_k) = \sum_{j-0}^{k} p_z(z_j),$ $\qquad k = 0,1,2,3, \dots, L-1$

Where L denotes the total gray level number (256).

The goal is to map the r value in X to the z value with the same probability as the desired pdf. i.e., $S(r_j) = G(z_j)$ or $z = G^{-1}(s(r))$ as shown in Figure 5-1

(Correal et al., 2014).

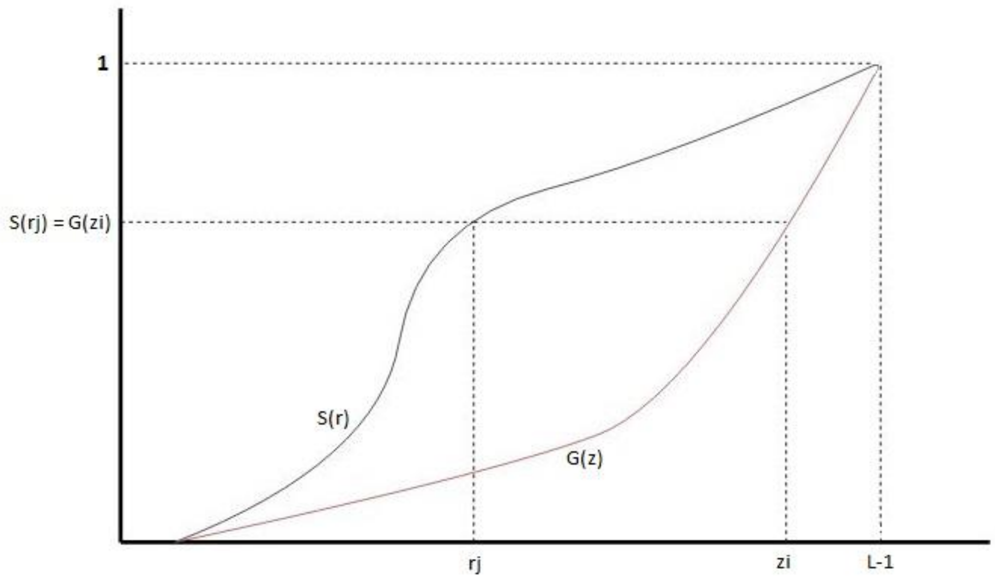


Figure 4- 1: Input image matched to desired output CDF

4.1.5 Median Filtering

Median filtering is a non-linear digital filtering technique well suited for the removal of noise for historical documents. Its edge-preserving nature also helps with OCR computing time and performance. As reviewed in the preceding chapter, most smoothing techniques are effective in removing noise but usually blur the edges. It is however very important to preserve edges as they affect image appearance and detection. It is effective in removing residue salt and paper noise (Hambal et al., 2017). The working process of median filtering takes an output sample computed as the median value of the input samples under the window. For relatively uniform areas, the median filter estimates grey level and when an edge is crossed, the other side dominates the window thereby switching up the values. This ensures the edges are not blurred (Erkan et al., 2018).

4.2 Process Evolution and Operation

This section comprises of the description and step by step process on the working operation of the proposed novel algorithm that enhances degraded document images adaptively.

The proposal considers the distribution of the histogram as a key factor to improve the document quality. The aim is to turn the original histogram into a bimodal histogram and therefore adaptively improve the discrimination between the foreground (letters) and background (paper). This enhancement technique is configured to enhance historical documents without the problem of over enhancement and loss of edges. The stages in the proposed pipeline are demonstrated in Figure 4-2 and source code shown in appendix D and appendix E:

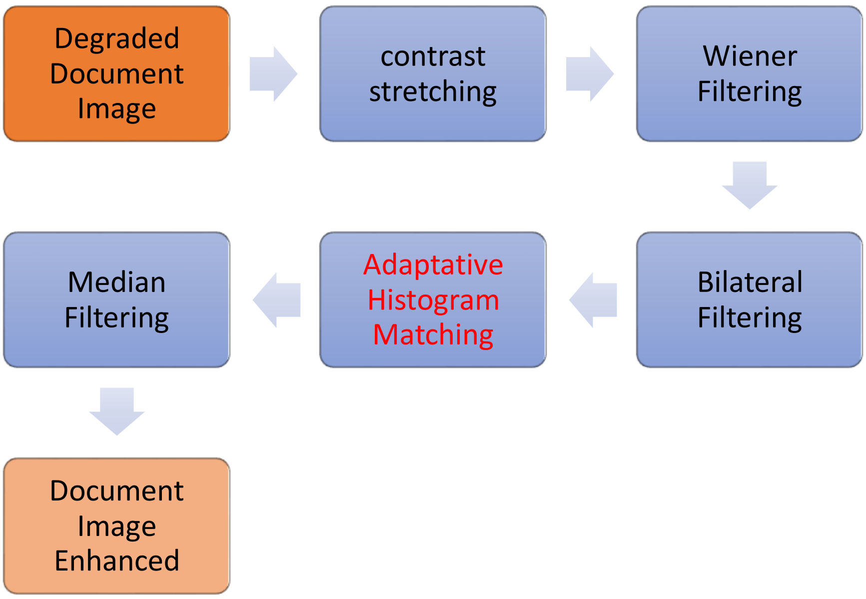


Figure 4- 2: Diagram showing the flowchart of the proposed processing stages

The algorithm of the processing stages built is stated algorithmically as:

Algorithm: EnhanceDocumentImage

Input: Degraded Document Image

Procedure:
1. Apply Contrast Stretching:
   - Input: Degraded Document Image
   - Output: Image after Contrast Stretching

2. Apply Wiener Filtering:
   - Input: Image after Contrast Stretching
   - Output: Image after Wiener Filtering

3. Apply Bilateral Filtering:
   - Input: Image after Wiener Filtering
   - Output: Image after Bilateral Filtering

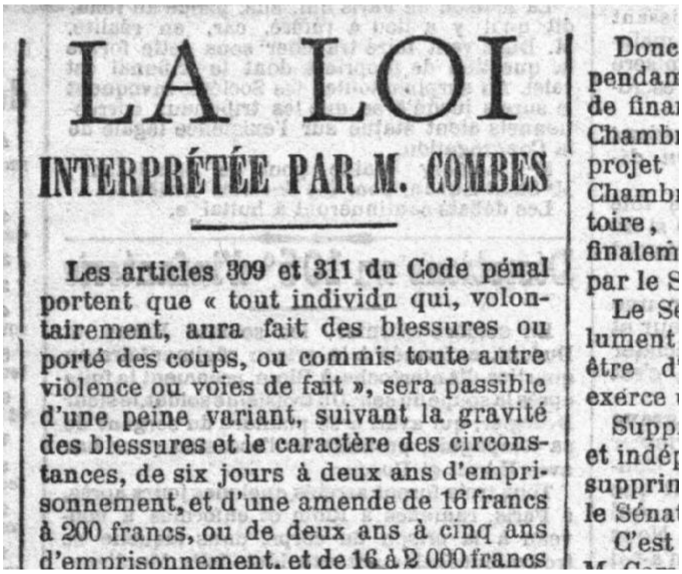
4. Apply Adaptive Histogram Matching:
   - Input: Image after Bilateral Filtering
   - Output: Image after Adaptive Histogram Matching

5. Apply Median Filtering:
   - Input: Image after Adaptive Histogram Matching
   - Output: Image after Median Filtering

6. Output: Enhanced Document Image
   - Final result after applying all enhancement algorithms

4.2.1 Degraded Document Image

The experiment is initiated with the transformation of the degraded document image to grayscale values to reduce computational time, reduce complexity and improve edge detection. The signal to noise ratio (SNR) and entropy are the initial metrics used to record and observe the gain and performance of the method in each stage of the enhancement pipeline process to indicate improvement(Krbcova & Kukal, 2017).



LA LOI

INTERPRÉTÉE PAR M. COMBES

Les articles 309 et 311 du Code pénal portent que « tout individu qui, volontairement, aura fait des blessures ou porté des coups, ou commis toute autre violence ou voies de fait », sera passible d'une peine variant, suivant la gravité des blessures et le caractère des circonstances, de six jours à deux ans d'emprisonnement, et d'une amende de 16 francs à 200 francs, ou de deux ans à cinq ans d'emprisonnement, et de 16 à 2 000 francs

SNR [dB]: 5.0208 Entropy: 5.8909

Figure 4- 3: Diagram showing sample ENP image.

In this patch image example in figure 4-3, from the ENP, the common undesired artefacts and degradations like uneven illumination, show-through and faded ink are present. The following histogram in figure 4-4 corresponds to this historical image.

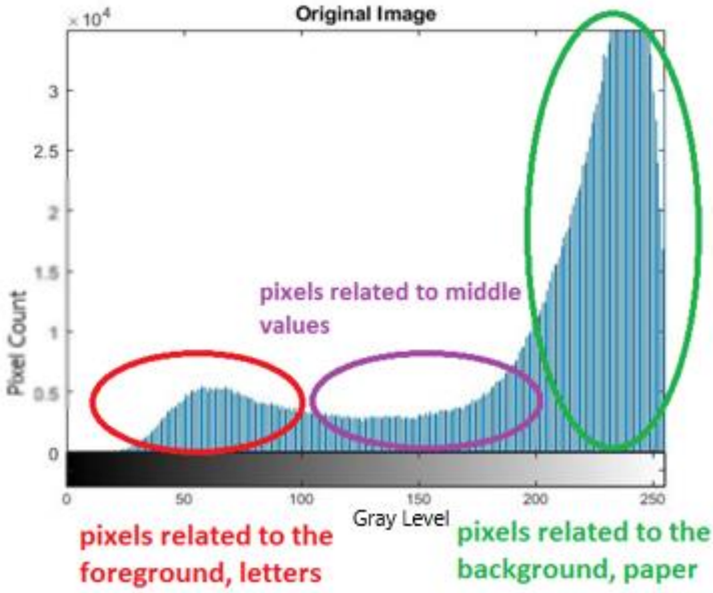


Figure 4- 4: Histogram showing original image pixels

The experiment aims to isolate both peak bins further, thereby increasing the bimodality and diminishing the middle pixel values. The middle pixel values are associated with undesired show-through artefacts and signal interference in the background, thus reducing this amount will reduce the undesired noise. In that direction, the final histogram will be more bimodal and the letters better perceivable from the background.

4.2.2 Contrast Stretching

The first pre-operation carried out in the document image is contrast stretching. This is a normalisation algorithm as described. It is deployed to improve the image contrast by stretching the range of intensity values it contains to span the full grayscale range of pixel values. This linear operation is considered as a pre-operation technique to normalise the pixels giving enough range to modify the pixels.

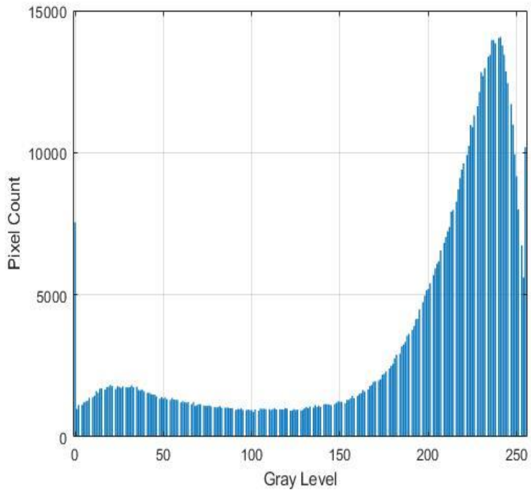

contrast stretching

LA LOI

INTERPRÉTÉE PAR M. COMBES

Les articles 309 et 311 du Code pénal portent que « tout individu qui, volontairement, aura fait des blessures ou porté des coups, ou commis toute autre violence ou voies de fait », sera passible d'une peine variant, suivant la gravité des blessures et le caractère des circonstances, de six jours à deux ans d'emprisonnement, et d'une amende de 16 francs à 200 francs, ou de deux ans à cinq ans d'emprisonnement, et de 16 à 2 000 francs

Donc pendan de fina Chamb projet Chamb toire, finalem par le S Le S lument être d exerce Supp et indé supprin le Séna C'est



SNR[dB]: 6.3021, Entropy: 5.9926

Figure 4- 5: Image and Histogram after contrast stretching

The above image in Figure 4-5 shows the sample image histogram extended to cover the full grayscale range and also an improved contrast distribution on the original image. This operation is important to set the basis for further operation.

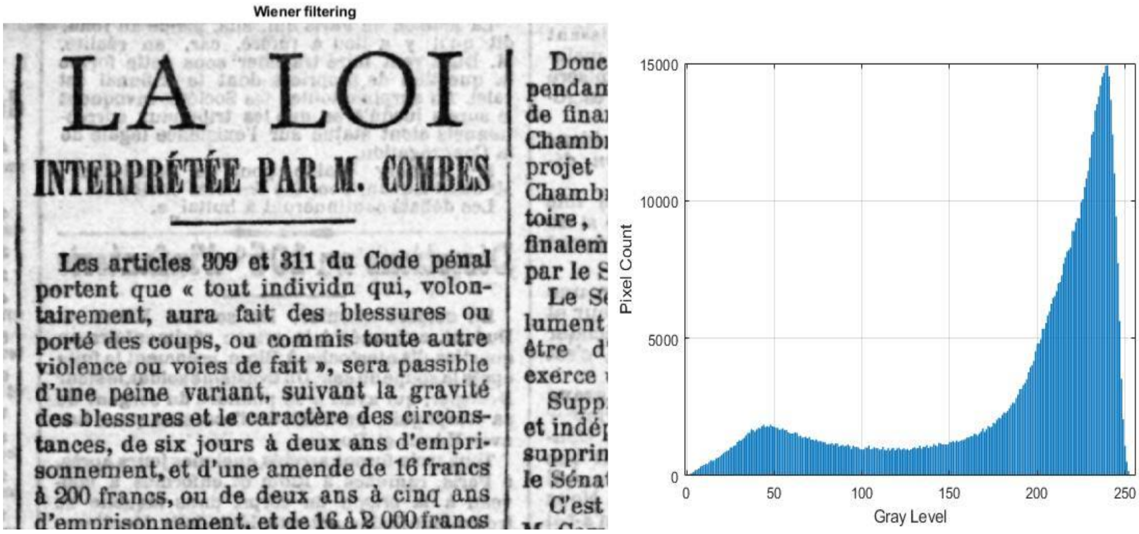
4.2.3 Wiener Filtering

Following contrast stretching, a 2-D adaptive low-pass noise-removal wiener filter is applied to the image, as described in the previous chapter, this filter is applied also as a pre-operation to further prepare the image by removing the noise and smoothening the image.

It is applied using the in-built Matlab function

`J = wiener2(I,[m n])`, where `[m,n]` is set to `[5,5]` specifying the `size (m x n)` of the neighbourhood used to estimate the local image mean and standard deviation.

The result obtained and the histogram of the image is shown in Figure 5-6 below:



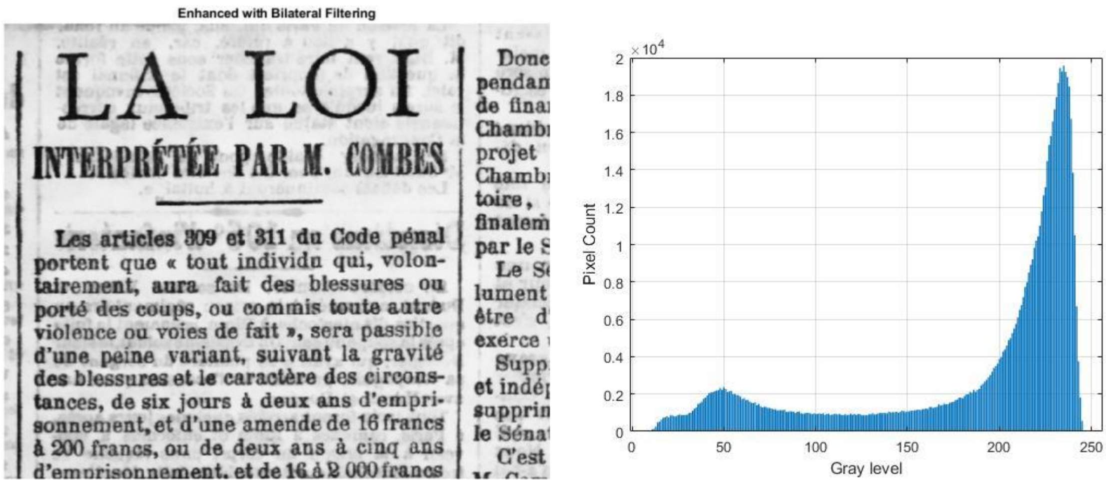
SNR [dB]: 8.9022, Entropy: 6.0239

Figure 4- 6: Image and Histogram after Wiener filter

The smoothing effect of the filter on the image can be noticed. The histogram of the image maintains its shape.

4.2.4 Bilateral Filtering

Subsequently, bilateral Gaussian filtering is applied to the pre-operation process after wiener filtering. As described in the previous chapter, a bilateral filter is a nonlinear filtering approach that will smooth the background of the image without affecting the edges of the letters. It is a low pass filter that can maintain the edges of the image. The technique has two parameters to modify from, he first one is the degree of smoothing that controls the degree of edge preservation, and the second is the spatial Sigma which is the sigma of the Gaussian function that controls the degree of filtering background.



SNR [dB]: 10.2432, Entropy: 6.2362

Figure 4- 7: Image and Histogram after Bilateral filter

The image in Figure 4-7 shows some improvements made to further prepare the image. The show-through artefact is shown to be diminished considerably after this operation.

4.2.5 Adaptive Histogram Matching

This stage depicts the novel part of the algorithm proposed in this research. Studies carried out by Han et al. (2019) classified image histogram into three categories and claim that global thresholding methods perform well on document images when the histogram approximates a balanced bimodal type. They also describe how a balanced bimodal type of histogram depicts a homogenous and uniform illumination which is ideal for the human visual system.

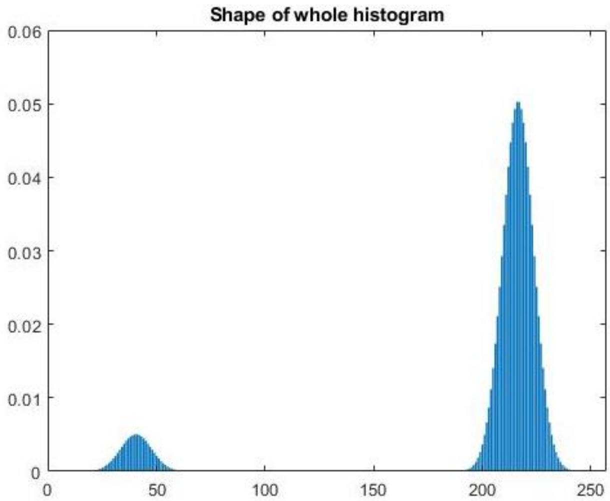


Figure 4- 8: Generic Bimodal Histogram

This concept inspired the research idea position on how a histogram of a perfect enhanced document should look like. The goal is to turn the final histogram into a histogram closer to the ideal shape. Kavallieratou (2005), Gatos et al. (2004), Bannigidad and Gudada (2016) describe that an ideal bimodal document histogram has two main distributions, the first one related to the foreground and the second one with the pixels of the background. Studies have shown that using bimodal image enhancement can greatly improve the accuracy of recognition and human evaluation(Zhu et al., 2011). Different bimodal image enhancements will impact the accuracy of image recognition.

From research, the following points were deduced to influence the enhancement guidelines for creating a novel adaptive histogram technique that creates its ideal reference histogram based on individual attributes and artefacts of each dataset:

- The two distributions must be separated as far as possible, this feature provides a better contrast.
- The pixels of middle values must be attenuated because high histogram middle pixel values associated with grayscales relates to show-through and other undesired artefacts.
- The second distribution will always be higher than the first one. This is because there must be more pixels on the background than in the foreground (letters).
- Moreover, a histogram of this shape is advantageous due it is the shape where Otsu binarization performs its best in case binarization is needed in future works.

The technique of histogram matching consists of adjusting the histogram of an input image into the desired histogram reference. Since the goal of this experiment is to transform the histogram obtained from the bilateral Gaussian filtering stage and to improve it, therefore the histogram matching technique is utilised.

This research contributes a new type of histogram matching technique (adaptive) that generates its reference histogram taking the histogram of the filtered image obtained after bilateral Gaussian filtering as a basis. The following pipeline in figure 4-9 depicts this technique:

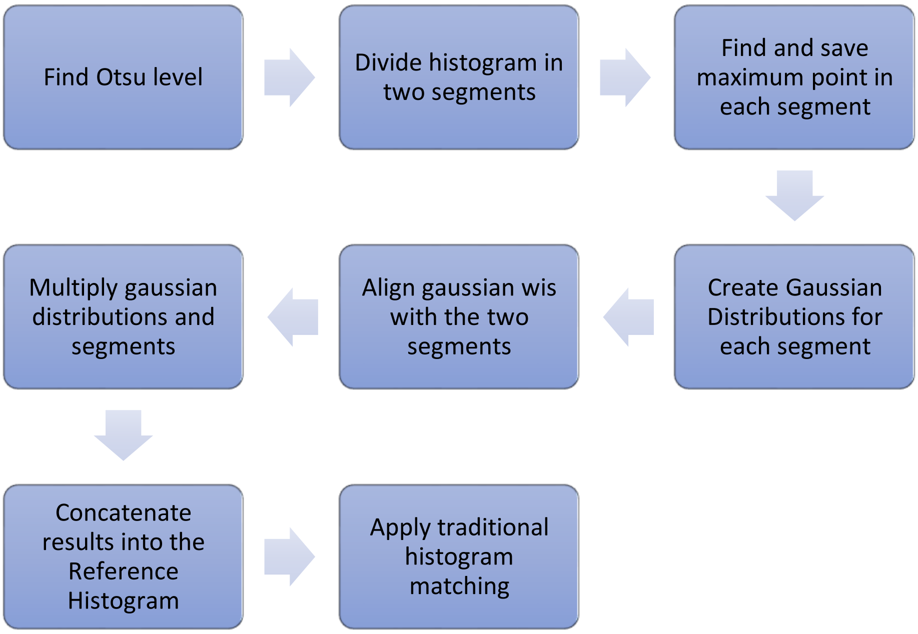
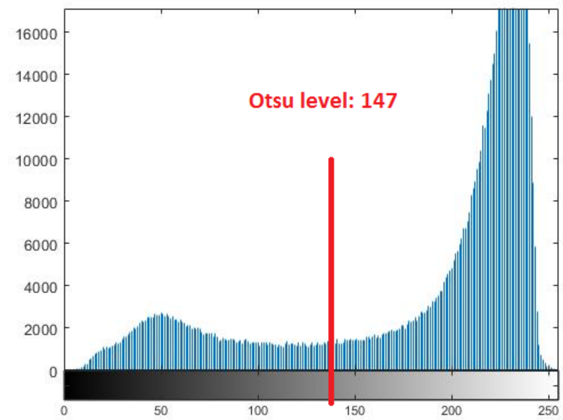


Figure 4- 9: Adaptive Histogram method pipeline

Find Otsu level and divide the histogram

The first two steps consist of finding the Otsu grey-level and dividing the histogram into two segments using the Otsu level as the boundary. This level is very crucial because it enables the separation of the two main distributions. The following histogram images in figure 4-10 below show the Otsu level and the separation process from after the bilateral filtering process which is the input image.



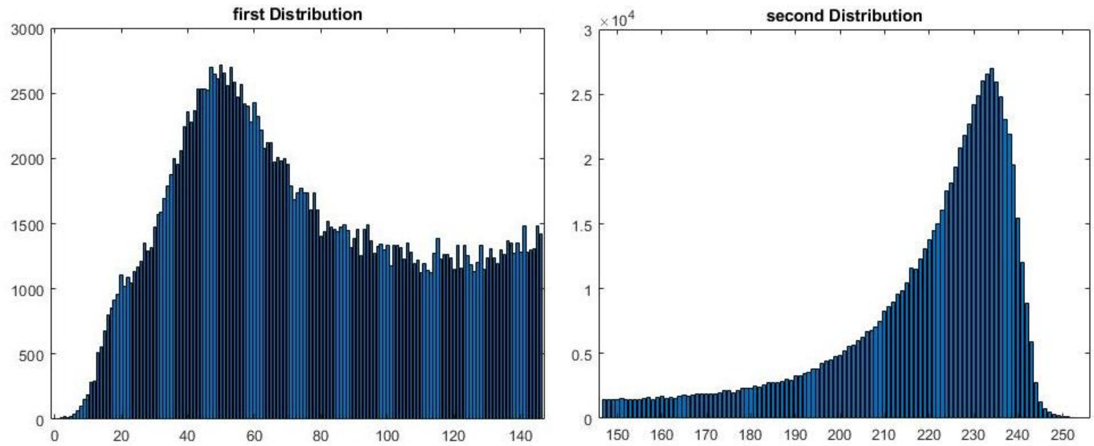


Figure 4- 10: Sample Histogram showing first and second distribution

Find and save the maximum point of each segment

The next step is to compute **the maximum of each distribution** (red line).

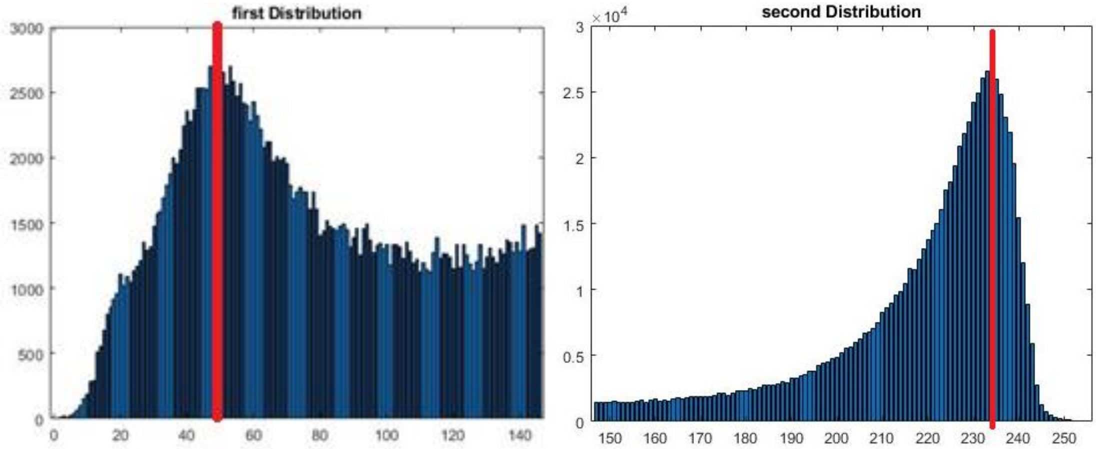


Figure 4- 11: Split Histogram showing the maximum of each distribution

Create Gaussian distributions for each segment

Since the log of a Gaussian produces a parabola, Gaussian distributions can be used for nearly exact quadratic interpolation in frequency estimation in all signal types(Hansen, 2014). This interpolation and its ability to control and manipulate signals is the baseline for the design algorithm to give weights of importance to the histogram distribution. This technique is used to ensure a better bimodal shape of the histogram using the multiplication of the Gaussian bell in each segment by giving importance to the pixels following the parabolic curve.

Two Gaussian distributions are created for each segment. Each Gaussian distribution is structured as three times the length of its corresponding segment and both have the same standard deviation. The following figure shows both Gaussian distributions:

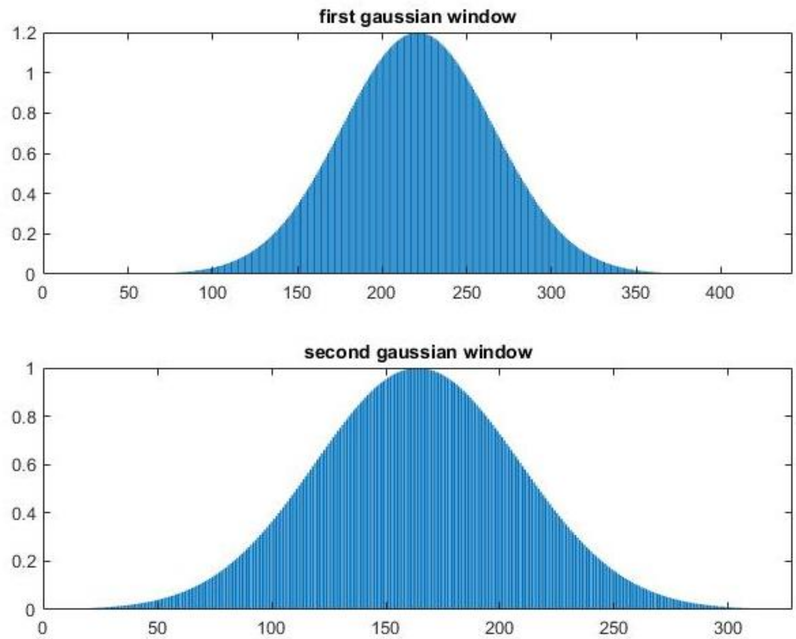


Figure 4- 12: Image showing generated Gaussian distributions

Gaussian distribution coefficients are computed from the following equation (Hansen, 2014):

$$w(n) = e^{-\frac{1}{2}(\alpha\frac{n}{(L-1)/2})^2} = e^{-n^2/2\sigma^2} \quad (4.1)$$

where $-(L-1)/2 \leq n \leq (L-1)/2$ (4.2)

and α is inversely proportional to the standard deviation, σ, of a Gaussian random variable. The standard deviation of a Gaussian probability density function is $\sigma = (L-1)/(2\,\alpha)$

The input arguments are given by;

L is the specified Window length,

α is the width factor which is inversely proportional to the width of the window,

w is the Gaussian distribution, returned as a column vector

From experimentation, the proposed Gaussian distribution is designed with the following variables to achieve the desired pixel weights and distribution:

$\sigma 1$ = 45; standard deviation for first segment

$\sigma 2$ = 45; standard deviation for second segment

$w1$= 1.2.

$w2$= 1;

$L1$= 3*length (segment1);

$\alpha 1 = (L1-1)/(2\,\sigma 1)$; (4.3)

$L2$= 3*length (segment2)

$\alpha 2 = (L2-1)/(2\,\sigma 2)$; (4.4)

The first Gaussian distribution is an estimated shape to achieve the desired distribution related to the foreground pixels. The maximum value of 1.2 and all other parameters are chosen after experimentation with different set values as the parameters that produces the desired emphasis level of font intensity of the letters and consequently the best readability. The values of both Gaussian functions will perform as weights to give importance to different values of each segment of the histogram. The length of the Gaussian distribution was experimentally chosen after experimentation of other values to be three times the length of each histogram segment as the parameters that ensure a better expansion of the gaussian bell to accommodate and truncate the required pixels. These values are critical to ensuring the desired performance of the method.

Align Gaussian distributions with the two segments

After the creation of the Gaussian distributions, it is important to align the two Gaussian distributions with the two segments of the histogram. This is achieved by aligning the two maximum points of each Gaussian distribution to the maximum point position of the two histogram segments. This operation is performed before multiplication. The following figure 4-13 shows this step.

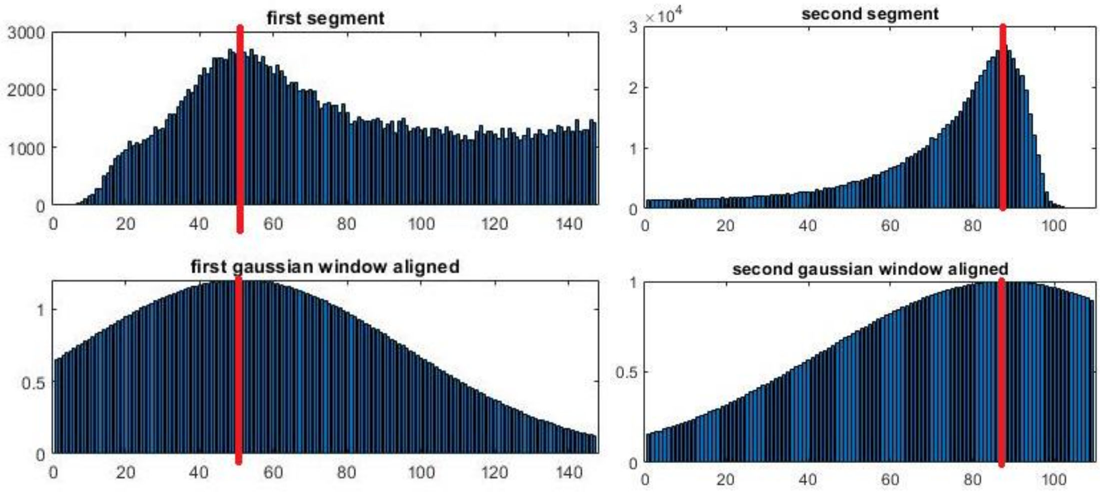


Figure 4- 13: Image showing generated Gaussian distributions in alignment with histogram distributions

The length of the Gaussian distribution shifted must be equal to the length of the two segments. This must be accomplished before multiplication.

Multiplication of the Gaussian distributions with the two segments

When the segments and the windows are aligned, the multiplication of both variables is carried out. This scalar multiplication aims to give 'importance' to the most relevant pixels and attenuate those which are not. The maximum points in both segments will keep their shape because they represent the amount of prevalent visual pixels (letters and background). Nevertheless, the pixels that are far from those maximum points will be smoothly attenuated, the middle pixel values will be attenuated. The following figure shows this multiplication result:

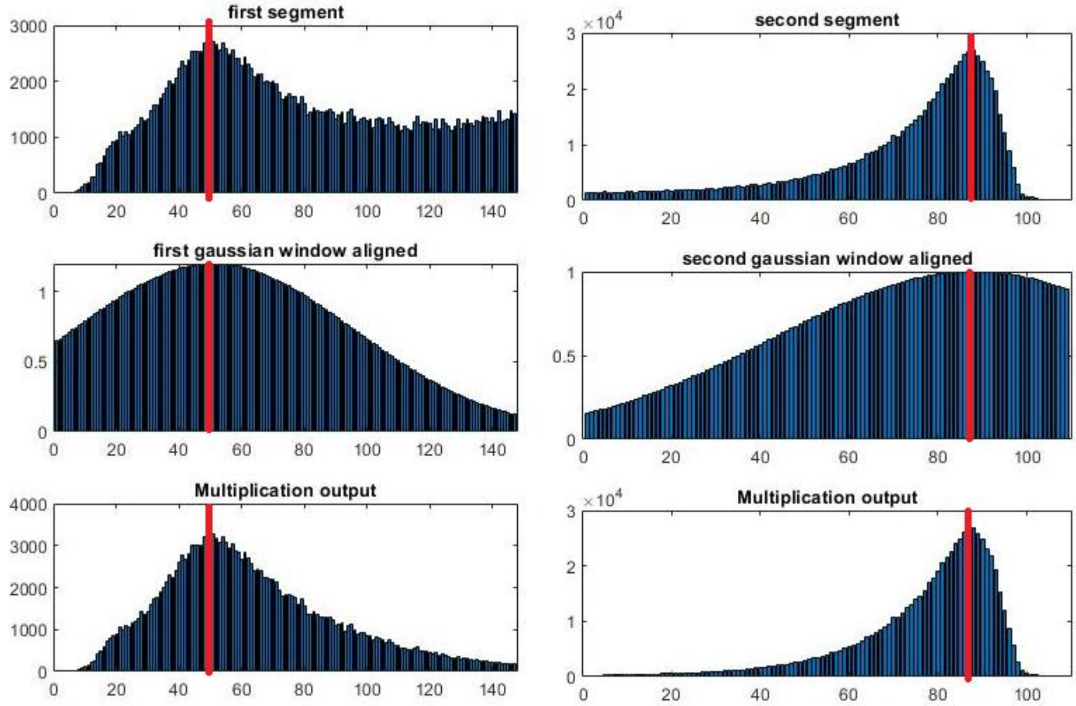


Figure 4- 14 Image showing generated Gaussian distributions, histogram, and output

From the figures above it can be observed how middle pixel values are attenuated in both results. The first segment shows substantially how much the multiplication output has attenuated its middle values (in this case values forward right).

Concatenate results into the reference histogram

This step consists of combining the two-multiplication output to achieve the final reference histogram that will be used for the histogram matching technique. This histogram is the approximate shape for the desired reference histogram for matching the enhanced document. From the resultant histogram image shown below, the attenuation of the middle values was attenuated and how the two distributions were enhanced and smoothed.

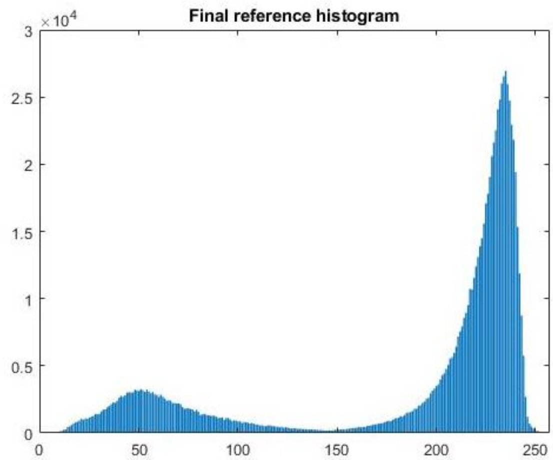
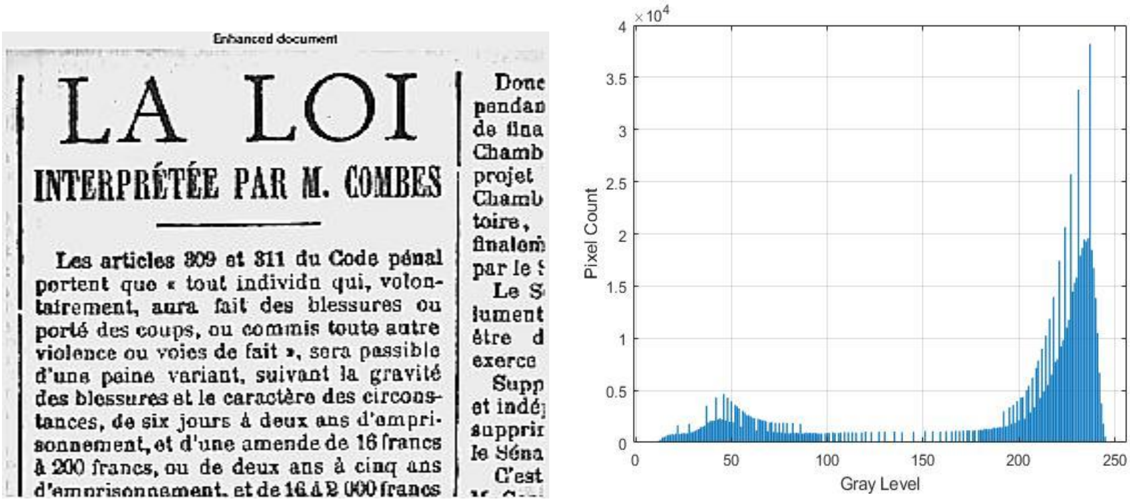


Figure 4- 15 Image showing the final generated histogram

Apply traditional histogram matching

The document image obtained at the output of the bilateral filtering process above is transformed using the traditional histogram matching function with the reference histogram generated. This is achieved by mapping each pixel of the original image histogram with the index of the reference histogram generated following the range of 0 to 255 pixels. The following figure depicts the resultant image and its histogram.



SNR [dB] Original: 15.0300, Entropy: 6.8967

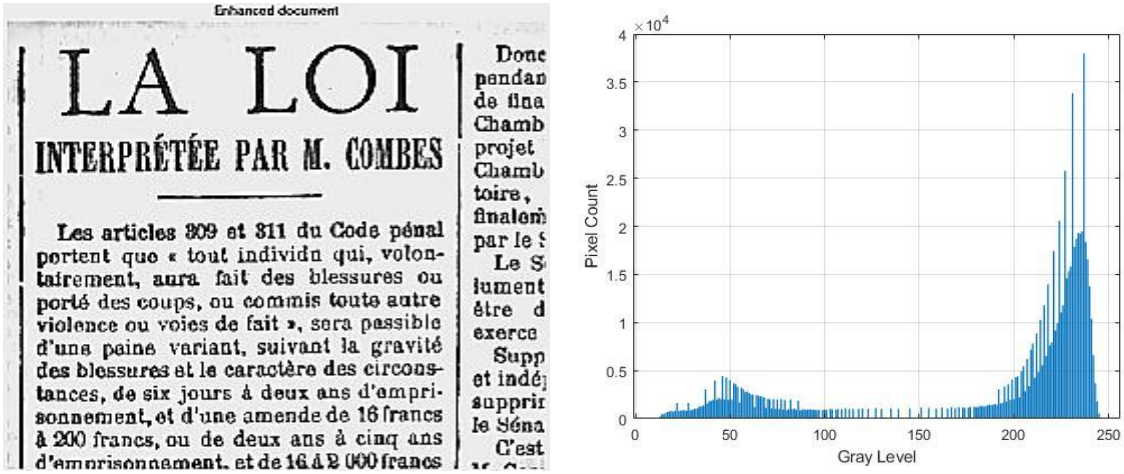
Figure 4- 16 Image showing the final image and histogram after Histogram Matching

The final histogram tends to adopt the shape of the reference histogram and shows fewer middle grey values than the first histogram. Enhancement in the amplitude of each distribution is also clearly depicted. The final histogram shows some disparity to the reference because both histograms are discrete values, digital image signals are typically represented as two-dimensional (2D) arrays of discrete signal samples and theory states that they can only be equal when the histogram is continuous (Wang et al., 2005). However, as shown, the resultant enhanced document achieves comprehensive visual quality.

Median Filtering:

This is a post-operation step to make a more robust pipeline in the proposal. The median filter is incorporated in the system as a contingency in case some document images present salt and pepper noise after the adaptive histogram process. In that event, a median filter operation application suppresses the noise and produces a better result.

The mask neighborhood applied within the median filtering is of 3x3 size.



SNR [dB] Original: 17.2300, Entropy: 7.6967

Figure 4- 17 Image showing histogram after Median Filter

The shape of the histogram shows no changes and the same distribution as the adaptive histogram stage as shown in the figure above. This is due to the lack of salt and pepper noise in the tested document. As discussed, this is a contingency operation in the pipeline in the eventuality of salt and pepper noise which the median filter is proven to suppress(Liu, 2013).

Discussion: The whole pipeline process controls the histogram shape by emphasising the bimodality and isolating the peak distributions further from one another. Both peak bin values increase while the middle values attenuate to present an ideal bimodal histogram. The increasing SNR and Entropy at each stage of the enhancement indicates improvement of the image. A substantial increase in SNR and entropy is recorded at the novel segment (adaptive histogram matching) of the enhancement. There was an increase of approximately 5dB in SNR and an entropy increment of 0.6605 after the adaptive histogram matching application.

4.3 Performance Evaluation

The performance evaluation metrics considered in this research are Signal to Noise Ratio, Mean opinion score from surveys and a new blind state-of-the-art VDIQA metric. These evaluations are carried out in the next chapter.

CHAPTER 5

Performance Analysis

In this chapter, the experimental task of the research is carried out. This involves selecting the ENP dataset from the PRImA database as discussed in chapter 1. As discussed, a representative subset of the dataset is selected by keywords using the PRImA tool to represent the prevalent problems in the database. This subset was created to help identify the challenges more precisely as subjective experimentation with huge amount of dataset could prove cumbersome to the test subjects. Also, the objective and the subjective within the same scope, the same amount of images are used in both cases.

5.1 Experiment

The degradation faults for classification of the subsets are the prevalent artefacts broken characters, faded ink, low scan contrast, show through, smeared ink, uneven illumination, and uneven ink distribution. Some of the images were observed to possess a mixture of a combination of random prevalent problems. Due to the size of the images, a cut-out zoomed patch of each image is employed to give a better depiction and detail of the image enhancement process and result. An initial random experiment consisting of 18 selected images consisting of a combination of the degradation artefacts is experimented upon for a subjective analysis. The selected images were chosen to be 18 in number to enable the test subjects to be able to carry out visual evaluation on the images. Creating a survey with the entire subset will prove cumbersome for the test subjects to effectively evaluate the performance due to its large number of images (300), therefore a selection of the dataset with a mix of the prevalent problems is chosen with a limited number for simplicity of evaluation.

To also make a correlation between the human visual experimentation and the objective metric-based evaluation in the enhancement experiment, the same number of images are adopted to avoid disparity when comparing performance in the metrics.

Figure 5- 1 Cross-section of selected images from the subset

5.2 Pre-possessing Evaluation (Enhancement)

The experiment is structured to test the selected datasets shown in figure 5-1 using the MATLAB ® software by implementing the empirical methods in comparison to the proposed method and testing them with performance metrics. Then a survey using Google docs on human subjects for an opinion score of the performances on the same test images to form a correlation of the performance of the method for both subjective and objective baseline evaluation. Ethical approval is not required in the evaluation survey as the images are freely available in the public domain. The enhancement of the selected datasets is evaluated after implementing the proposed enhancement technique on the document images. The performance metrics used in this research for evaluating the enhancement technique proposed are SNR, MOS from the visual evaluation(subjective), VDIQA (objective) Shahkolaei et al. (2018), and binarization performance. The proposed technique is compared side by side with state-of-the-art methods Histogram Equalization, CLAHE, BMO, and ACMHE for the subjective evaluation. At the time of the commencement of the subjective evaluation, the implementation code for Sun et al. (2016) method Blind bleed-through removal for scanned historical document image with conditional random fields (BBTRSHD) had not been acquired and thus could not participate in the subjective evaluation. The results from the experiment are displayed in the figures below.

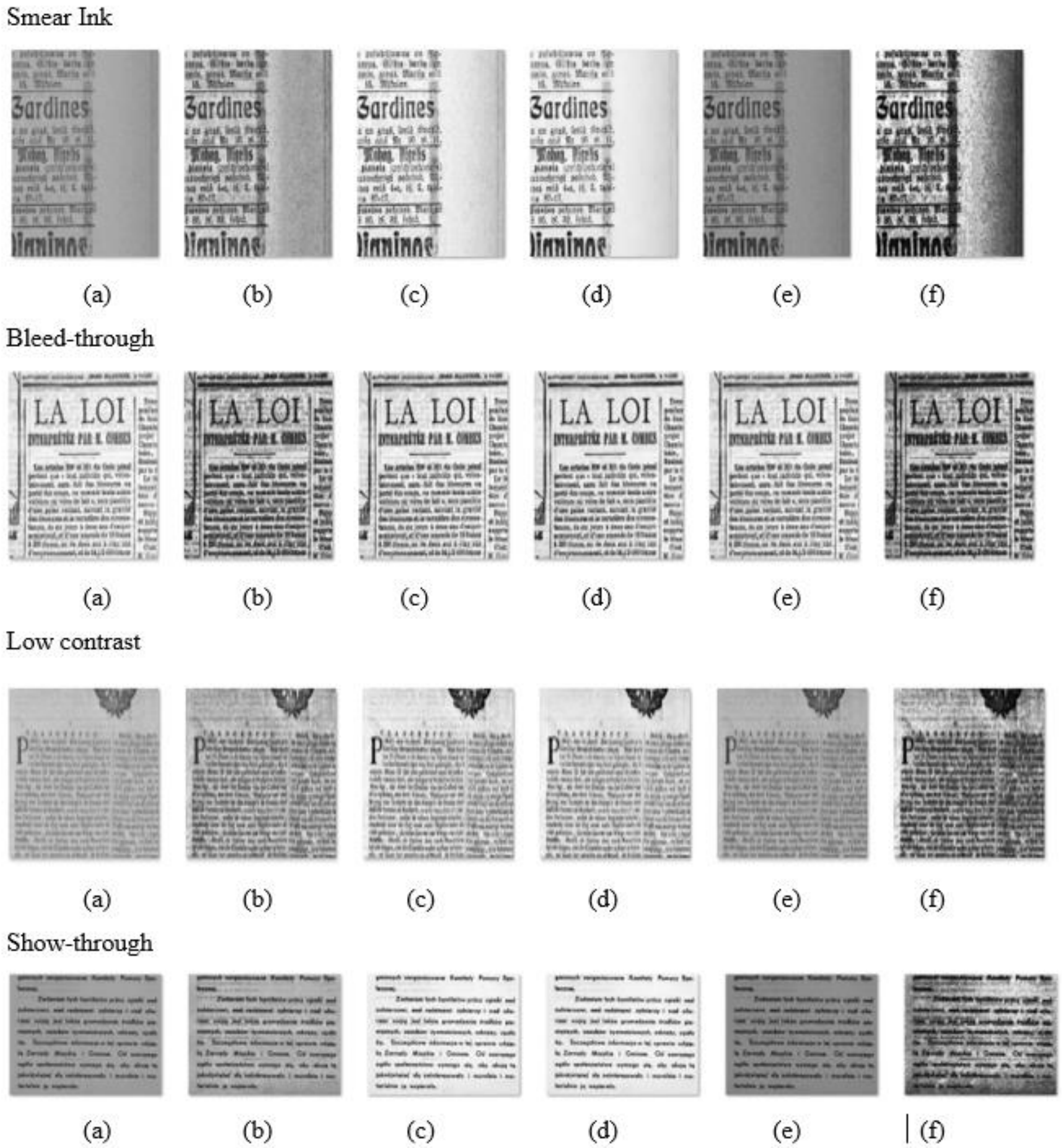
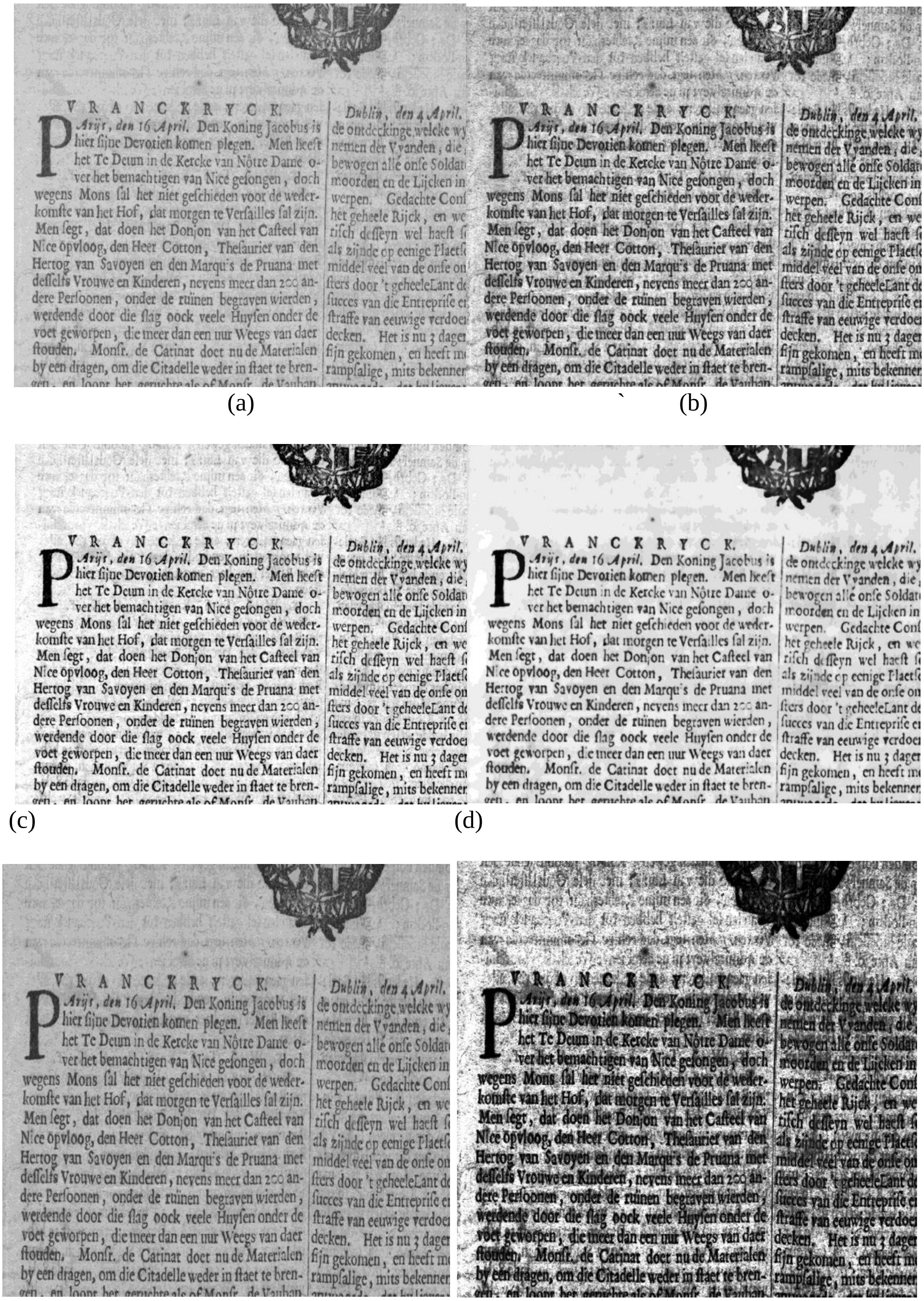


Figure 5- 2: Summary of Visual results of the experiment on four examples of degraded documents in the dataset; (a) original image (b) CLAHE result (c) BMO (d) Enhancement proposal (e) ACMHE (f) Histogram Equalization



VRANCKRYCK.

PArijs, den 16 April. Den Koning Jacobus is hier ſijne Devotien komen plegen. Men heeft het Te Deum in de Kercke van Nôtre Dame over het bemachtigen van Nice geſongen, doch wegens Mons ſal het niet geſchieden voor de wederkomſte van het Hof, dat morgen te Verſailles ſal zijn. Men ſegt, dat doen het Donjon van het Caſteel van Nice opvloog, den Heer Cotton, Theſaurier van den Hertog van Savoyen en den Marquis de Pruana met deſſelfs Vrouwe en Kinderen, nevens meer dan 200 andere Perſoonen, onder de ruïnen begraven wierden, werdende door die ſlag oock veele Huyſen onder de voet geworpen, die meer dan een uur Weegs van daer ſtouden. Monſr. de Catinat doet nu de Materialen by een dragen, om die Citadelle weder in ſtaet te brengen, en loopt het gerughte als of Monſr. de Vauban

Dublin, den 4 April.
de ontdeckinge welcke wy
nemen der Vyanden, die,
bewogen alle onſe Soldat
moorden en de Lijcken in
werpen. Gedachte Conſ
het geheele Rijck, en we
riſch deſſeyn wel haeſt ſ
als zijnde op eenige Plaetſ
middel veel van de onſe on
ſters door 't geheeleLant de
ſucces van die Entrepriſe e
ſtraffe van eeuwige verdoe
decken. Het is nu 3 dage
ſijn gekomen, en heeft m
rampſalige, mits bekenne

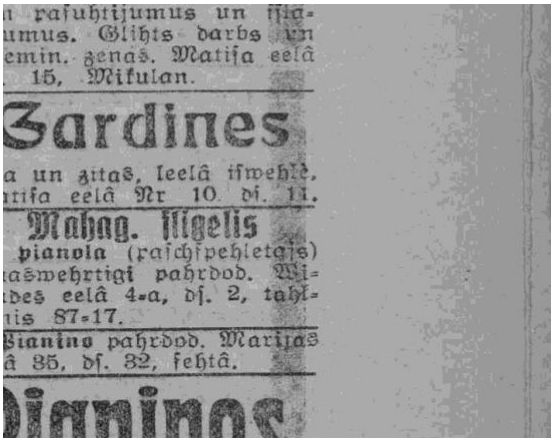
(a) (b)

(c) (d)

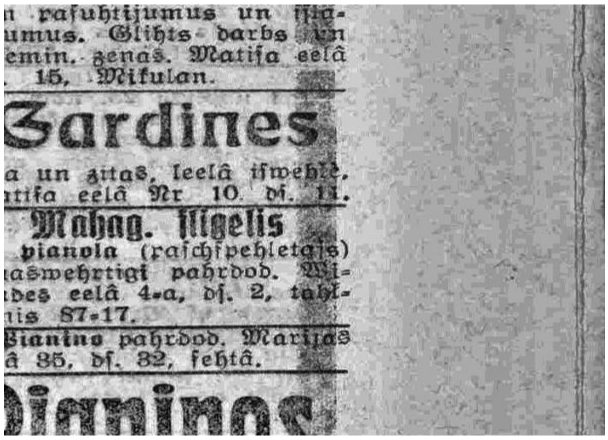
(e) (f)

Figure 5- 3: Visual results of the experiment on image 0067358 with Low scan contrast, Faint character, noise and show-through present; (a) original image (b) CLAHE result (c) BMO (d) Enhancement proposal (e) ACMHE (f) Histogram Equalization

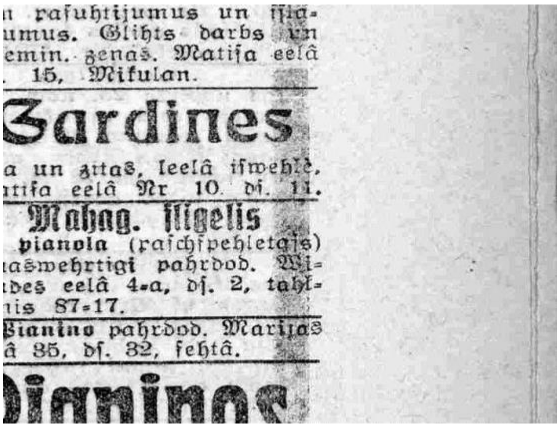
Figure 5- 4: Visual results of the experiment on image 00673977 with Low scan contrast, Faint character, noise and bleed-through present; (a) original image (b) CLAHE result (c) BMO (d) Enhancement proposal (e) ACMHE (f) Histogram Equalization



(a)

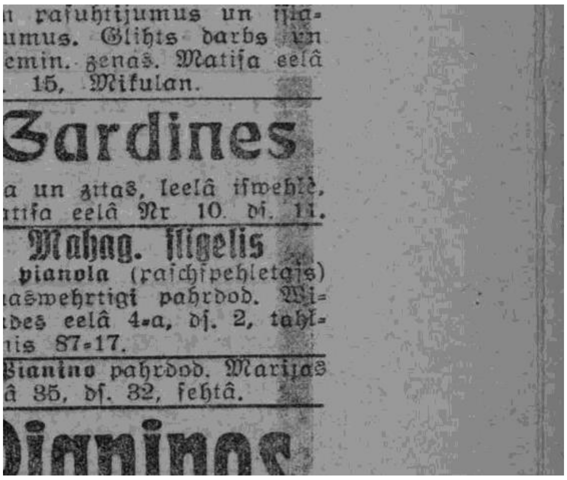


(b)

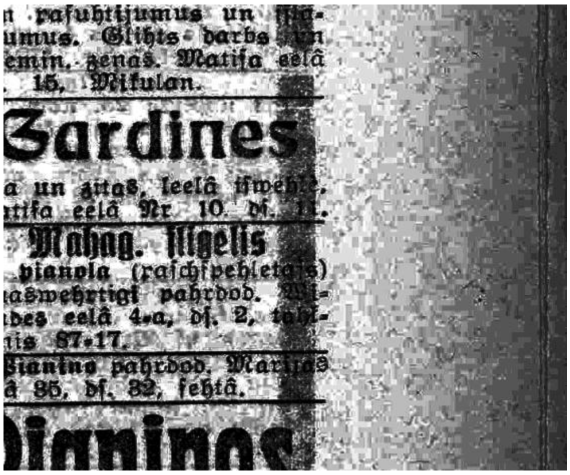


(c)

(d)



(e)



(f)

Figure 5- 5: Visual results of the experiment on image 00674680 with Low scan contrast and Smear-ink present; (a) original image (b) CLAHE result (c) BMO (d) Enhancement proposal (e) ACMHE (f) Histogram Equalization

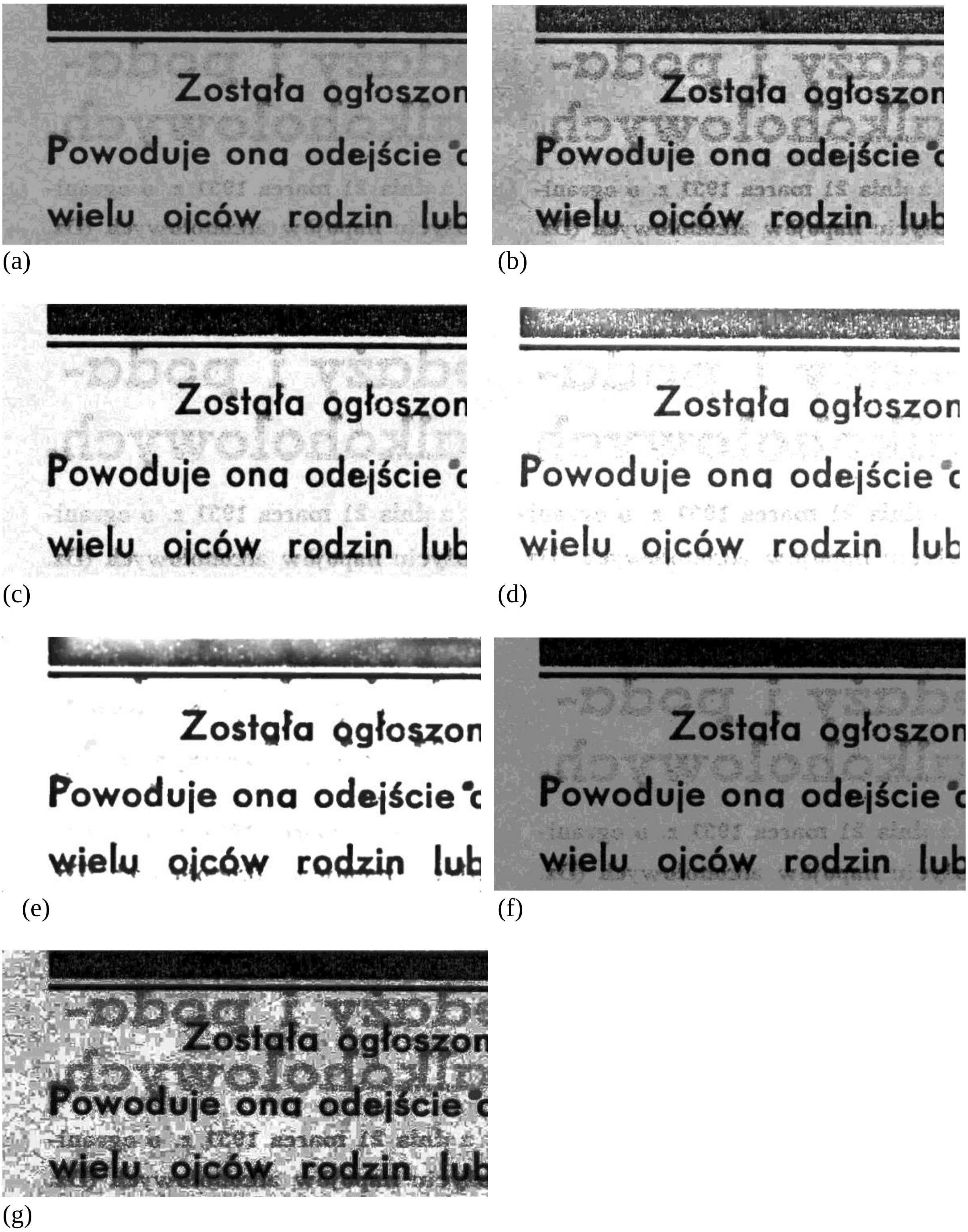


Figure 5- 6: Visual results of the experiment on image 00762015 with Low scan contrast and Bleed-through present; (a) original image (b) CLAHE result (c) BMO (d) BBTRSHD (e) Enhancement proposal (f) ACMHE (g) Histogram Equalization

After the implementation of the proposed algorithm, the image in figure 5-2 shows a side-by-side visual evaluation of a summary of some images for the prevalent degradations. Figure 5-3 shows the visual results of the experiment on image 0067358 containing low scan contrast, faint character, noise, and show-through artefacts, the proposed enhancement method performs best among all the other methods in visual resultant image quality from the apparent reduction of background noise, improvement of contrast, improvement of font intensity and removal of the minor bleed-through. The document images in figure 5-4 displays the result of the experiment on image 00673977 with low scan contrast, faint character, noise and show-through present, the proposed enhancement method also performs best for this set of degradations by the apparent reduction in the degradations. Figure 5- 5 shows the results of the experiment on image 00674680 with Low scan contrast and Smear-ink present and Figure 5-6 shows the results of the experiment on image 00762015 with low scan contrast and bleed-through present, the bleed-through degradation is shown to be eliminated completely in the enhancement proposal as compared to the other enhancement methods. The side-by-side illustration clearly shows a better performance and higher quality image in the proposed method as compared to the other methods when visualised with the HVS.

5.2.1 Subjective evaluation

To validate the subjective evaluation results, an extensive survey on the images is implemented on 22 test subjects consisting of a set of graduate student candidates whose areas of research range from computer science, electrical engineering, and image processing. They were commissioned to score the quality of the images based on the enhancement performance of each method. The survey[1] was carried out on a total of 126 images which were grouped into 21 questions with 6 enhancement options namely, original image, BMO, CLAHE, the proposed enhancement, ACMHE and histogram equalization as shown in figure 5-8. The test was carried out using google forms where each candidate was asked to give a score for the overall quality of the images by a ranking method. A quality score of 1 means excellent or highest perceived quality, 2 means good, 3 means fair, 4 means poor and 5 means bad or worst perceived quality. The sequence of the images was randomized in each presentation for a proper blind experiment and to stall pattern voting. The subjects could also rate the images as equal in instances where it was not easy to judge. The methods selected for the experiment were the methods which had been implemented at the time. The method BBTRSHD was not a part of this experiment as the implementation code for the method hadn't been attained at the time of collecting the survey.

[1] https://docs.google.com/forms/d/1O-8B6Y1mtcAvzfWy8cxfDcQxZlMnW0-4AfLgIThKlMk

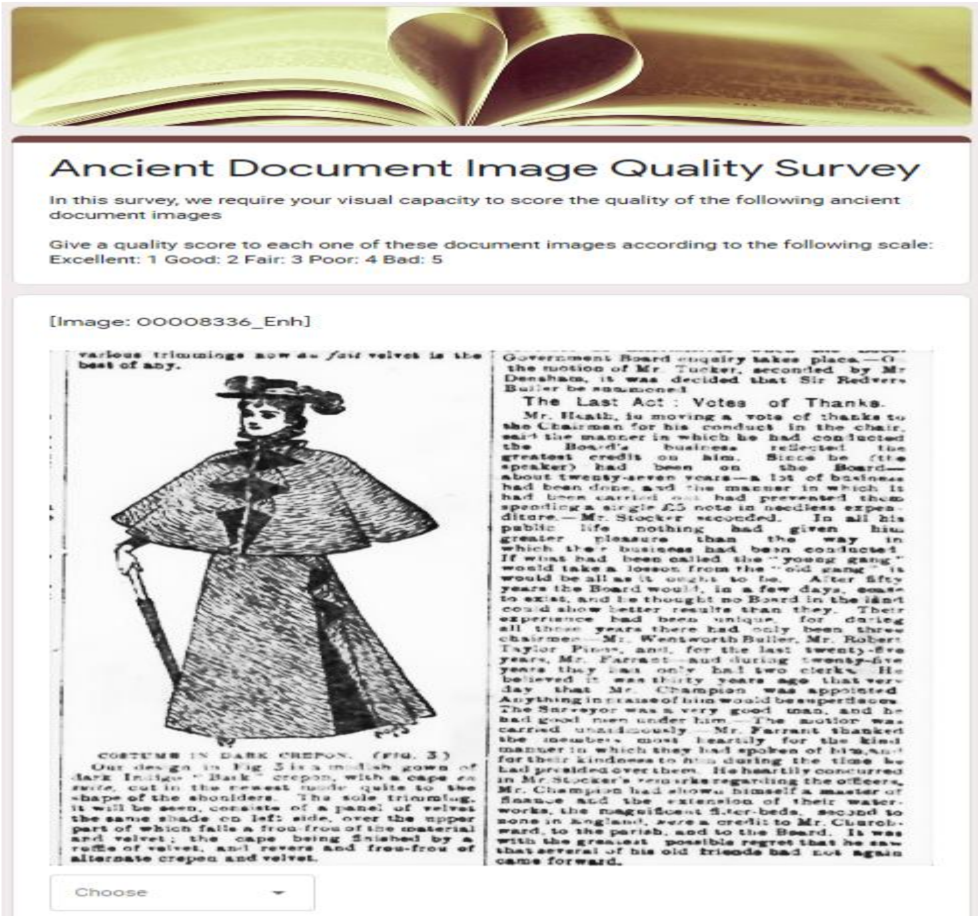


Figure 5- 7: Screenshot of the subjective evaluation interface

The HVS survey result on the test image 0067358 in figure 5-3 is shown in table 5.1 and figure 5-8. The HVS survey on test image 00673977 in figure 5-4 is shown in table 5.2, and figure 5-9. The survey shows how the percentage of the 22 subjects voted according to image quality for the different methods. In the case of the proposed enhanced image, 77.3% of subjects voted the enhancement proposal as excellent, 18.2% voted as good and no subject voted it fair or poor. However, 4.5% voted the proposed enhancement as bad. From the charts, the average scores by the test subjects recorded the proposed enhancement method as the best performing therefore validating the method. Some HVS results for the experiment on the other selected images are shown in the appendix.

Table 5. 1: Subject evaluation experiment result for test image 0067358

Method	Excellent (%)	Good (%)	Fair (%)	Poor (%)	Bad (%)
Original Image	15	35	35	15	0
Proposed Method	77.3	18.2	0	0	4.5
BMO	41	40.9	9.1	4.5	4.5
ACMHE	4.5	40.9	27.3	18.2	9.1

Histogram Equalization	13.6	27.3	13.6	18.2	27.3
CLAHE	13.6	22.7	36.4	18.2	9.1

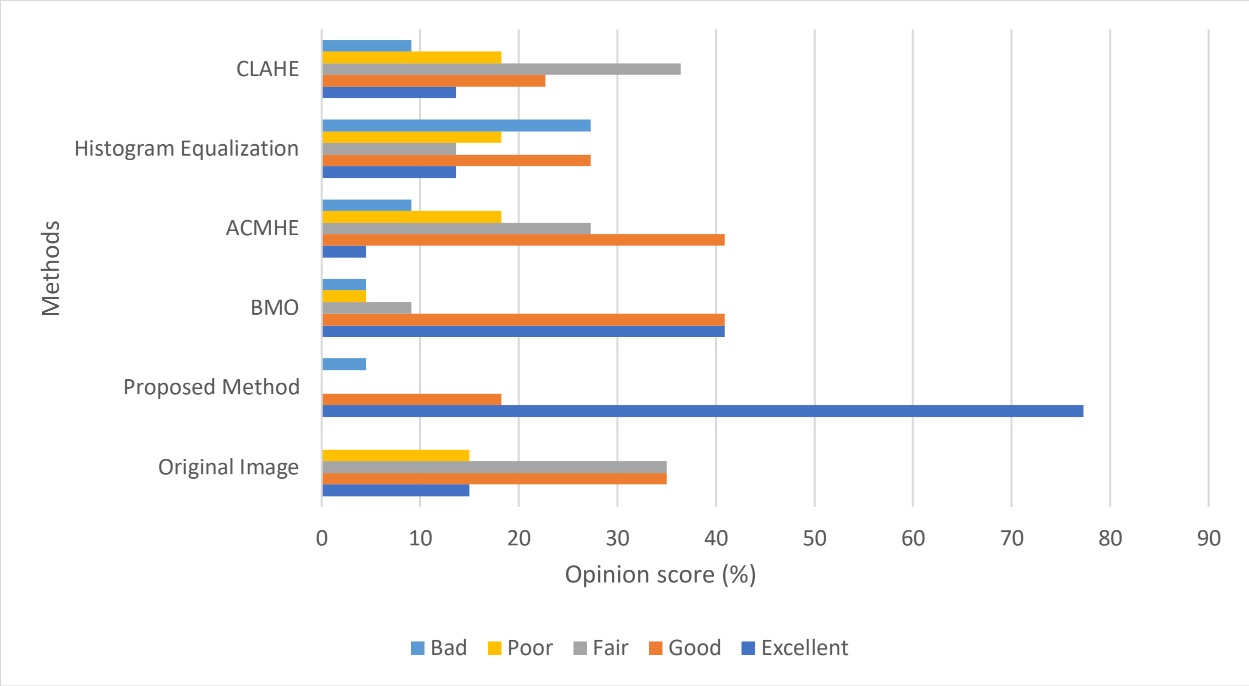


Figure 5- 8: Bar graph of subject evaluation experiment graph result for test image 0067358

Table 5. 2: Subject evaluation experiment result for test image 00673977

Method	Excellent (%)	Good (%)	Fair (%)	Poor (%)	Bad (%)
Original Image	0	13.6	54.5	31.8	0
Proposed Method	27.3	18.2	36.4	18.1	0
BMO	0	18.2	50	27.3	4.5
ACMHE	0	22.7	50	18.2	9.1
Histogram Equalization	0	9.1	13.6	31.8	45.5
CLAHE	0	18.2	54.5	18.2	9.1

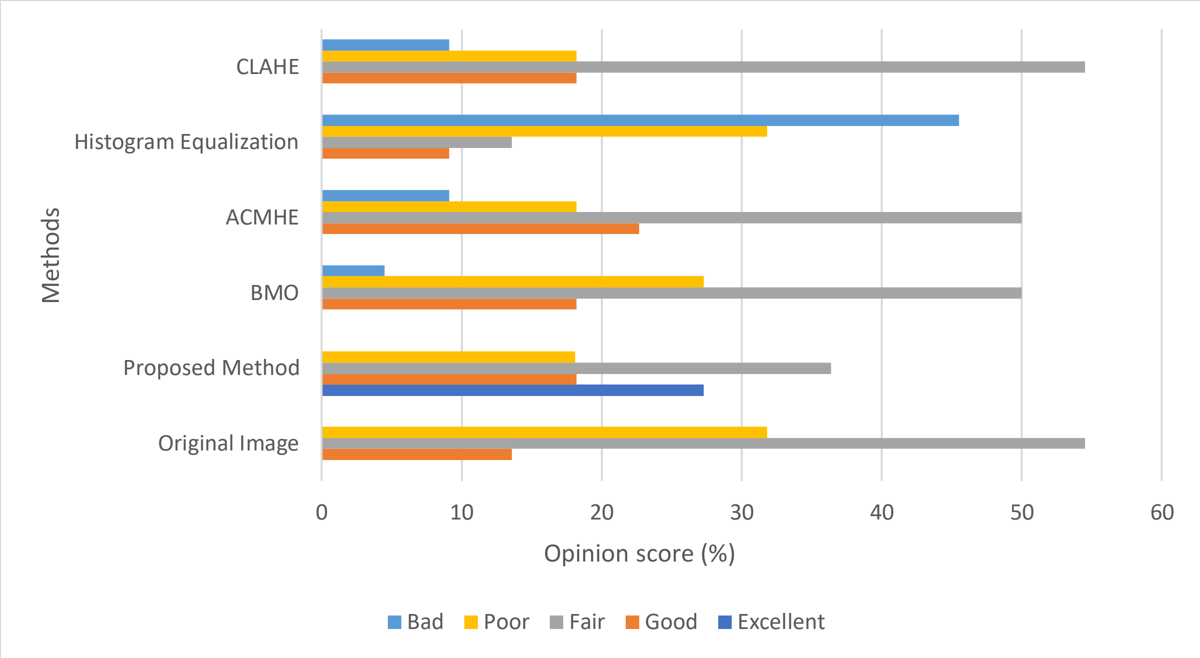


Figure 5- 9 Bar graph of subject evaluation experiment graph result for test image 00673977

Mean Opinion Score:

The MOS is calculated as the arithmetic mean over individual ratings performed by the human subjects.

The MOS is defined as (Katsigiannis et al., 2018):

$$MOS = \frac{\sum_{n=1}^{N} R_n}{N} \tag{5.1}$$

Where R is the individual ratings for a given stimulus by N number of subjects.

From the chart in figure 5-8 showing a sample analysis of the image 0067358, the majority (77.8%) of the candidates gave the enhancement proposal an excellent ranking. The MOS for the same image is calculated using equation (5.1) and has a maximum possible outcome of 5. In this evaluation, the proposal performed best with a score of 4.636, a 26% increment on the average score of the other methods, 10.92% increment compared to the BMO (the second-best performing method) and 22.72% better than the original image.

For image 00673977 with original image MOS of 2.818 and proposed enhancement MOS of 3.545, 27.3% of the subjects scored the enhancement proposal as excellent and none of the subjects scored an excellent performance for any other method, an 18.2% increment on the average score of the other methods, 13.64% increment compared to the ACMHE (the second-best performing method) and 14.4% better than the original image.

Table 5. 3: MOS results for images

Image	Original Image	Proposed method	BMO	ACMHE	Histogram equalization	CLAHE
1	4.000	4.015	4.000	3.997	2.378	3.191
2	3.500	4.636	4.090	3.136	2.818	3.136
3	3.5	3.650	3.850	3.350	1.600	3.000
4	3.728	3.455	3.815	3.774	2.904	3.820
5	2.818	3.545	2.818	2.863	1.863	2.818
6	3.314	3.549	3.409	3.411	2.366	3.087
7	3.684	4.364	4.543	3.592	2.047	3.453
8	3.350	4.678	4.358	3.964	2.817	3.395
9	3.269	4.578	3.060	3.560	1.964	3.120
10	3.450	3.920	3.650	3.720	2.950	3.350
11	3.672	4.650	3.870	3.930	2.700	3.180
12	3.325	4.256	3.920	3.864	2.870	3.345
13	3.560	4.328	4.126	3.814	2.720	3.350
14	3.250	4.538	4.258	3.724	2.970	3.235
15	3.240	3.862	3.658	3.864	2.870	3.045
16	3.350	3.850	3.780	3.720	2.950	3.087
17	3.805	4.560	4.000	3.950	2.700	3.905
18	3.950	4.650	4.150	3.980	2.950	3.560

Table 5. 4: Average MOS results for images

	Original Image	Proposed method	BMO	ACMHE	Histogram equalization	CLAHE
Mean	3.487	4.188	3.853	3.678	2.579	3.282
Standard Deviation	0.281	0.428	0.418	0.304	0.423	0.269

Table 5.3 shows the MOS results on 18 images from the experiment. From the results, the proposed method clearly exceeds the other methods in enhancement from the survey carried out by having an average score of 4.188 across the 18 images, the original image scored an average MOS of 3.487, the second-best performing method BMO averaged 3.853, ACMHE scored an average MOS of 3.678, HE which was the least performing method averaged a score of 2.580, and CLAHE recorded an average of 3.282 respectively as shown in Table 5.4. The t-test of the proposed method vs original method is approximately 5.805 which is an indication that the data readings are strong and not due to chance.

Discussion: Analysis of the voting pattern in the MOS results indicated that the enhancement proposal scored a lower enhancement margin in images with degradations of very uneven ink distribution and very faint text as shown in its lower significant performance on the MOS of Images 1, 4, 6, and 16 in Table 5.3. The proposed enhancement showed an improvement in the image contrast; however, it diminishes some very faded ink in the text close to the background. The method performed remarkably in images with the degradations; salt and pepper noise, low contrast, uneven illumination, smeared ink, faded ink, uneven-ink, show-through and bleed-through.

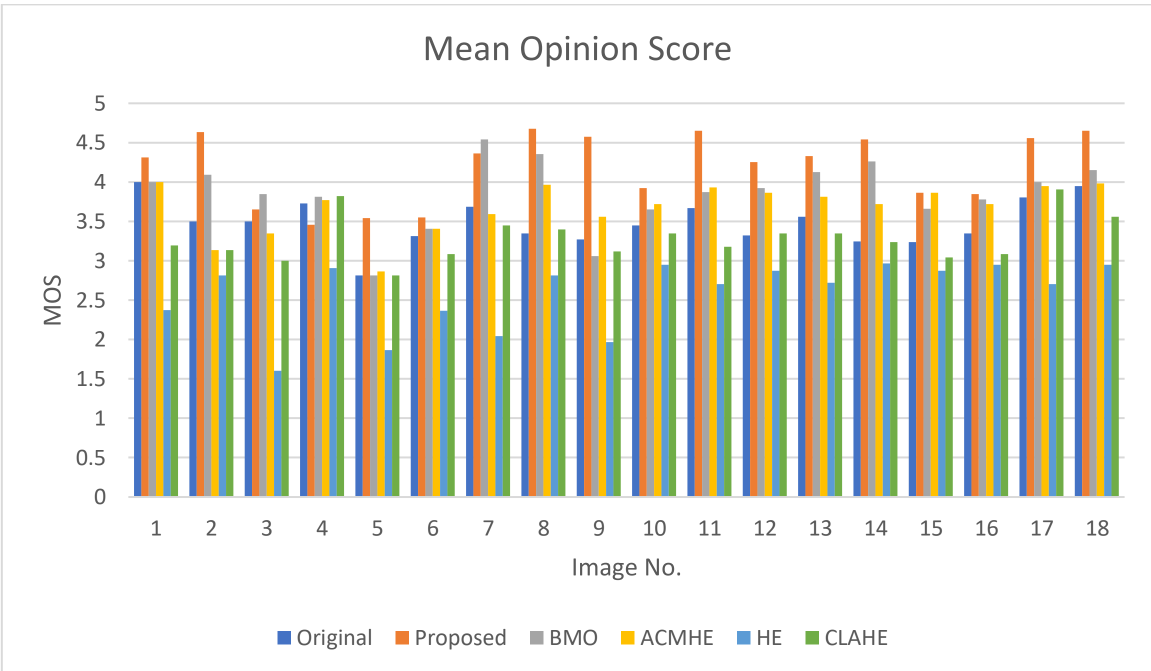


Figure 5- 10: Graph of MOS experiment result for test images

5.2.2 Objective Evaluation

Signal to Noise Ratio

Signal to Noise ratio is a metric used to quantify image quality, it essentially compares the desired signal level to the undesirable background noise within an image. SNR is thus the outcome of dividing the average (mean) signal by the variation (standard deviation) of the image signal and the higher the SNR value, the better the image quality.

$$\mathrm{SNR} = \frac{\text{Mean Signal Value} = \text{Useful Image Information}}{\text{Standard Deviation} = \text{Noise/Random Information}}$$

Images captured in low light conditions have low dynamic range, a low SNR and consequently degraded by noise since low light images contain much noise (Gaussian, Salt and pepper, etc.) in flat and dark regions which in turn impacts negatively the output of OCR and evaluation by HVS. Most traditional image enhancement do not consider noise characteristics, thus leading to amplification of noise while improving contrast. (Su & Jung, 2017).

The signal to noise ratio was tested as part of the objective evaluation of the experiment. A high SNR indicates a better image quality compared to noise in the image. From Table 5.5, the SNR for the 18 images (objective evaluation) of each method is presented. To proper evaluate each method, the mean SNR across the dataset by the individual method was calculated and presented in Table 5.6. The proposed method scored the highest mean of approximately 12.60. The second-best performing method was the BBTRSHD with an SNR mean of 10.70. In comparison, the least performing method is the histogram equalisation method with a mean of 4.41 due to its quality of improving contrast in some cases but may introduce noise and other undesired effects. HE tends to produce intensity saturation problem due to the shifting of the original mean brightness. In the experiment, HE leads to over-enhancement, which produces unnatural phenomena and amplifies the noise to the enhanced images. The other methods averaged between 6.36 to 10.70. The proposed method performs best due to its denoising property.

Table 5. 5: SNR results of the proposed method compared to state-of-the-art methods

Image ID	Original	Enhancement Proposal	HE	CLAHE	BMO	ACMHE	BBTRSHD
1	4.85	8.64	3.19	4.67	4.84	6.47	7.35
2	14.71	18.89	6.19	8.95	14.44	14.56	15.67

3	6.99	8.39	2.48	4.55	6.78	7.09	7.29
4	8.09	10.66	2.76	5.14	7.37	7.2	9.54
5	12.83	15.03	5.17	7.82	13.6	12.58	13.89
6	13.84	15.79	7.69	9.31	15.89	14.73	14.10
7	4.33	7.15	1.41	3.01	5.65	4.79	5.26
8	6.85	8.45	2.57	4.48	5.05	6.03	7.10
9	3.43	6.27	1.64	2.30	4.85	3.84	4.67
10	6.58	8.71	3.86	4.46	5.6	7.93	7.21
11	4.89	6.21	2.82	2.54	4.58	4.06	5.20
12	13.10	19.50	6.05	6.78	15.22	11.97	14.34
13	14.83	19.10	6.88	8.88	15.80	13.70	17.67
14	6.39	11.45	3.47	4.23	5.50	5.69	8.32
15	5.64	9.3	3.15	4.19	4.67	5.09	6.54
16	11.24	13.01	7.83	9.02	10.94	10.51	11.96
17	17.07	20.07	5.37	12.35	18.47	15.92	18.35
18	17.21	20.11	6.80	11.97	17.69	15.98	18.17

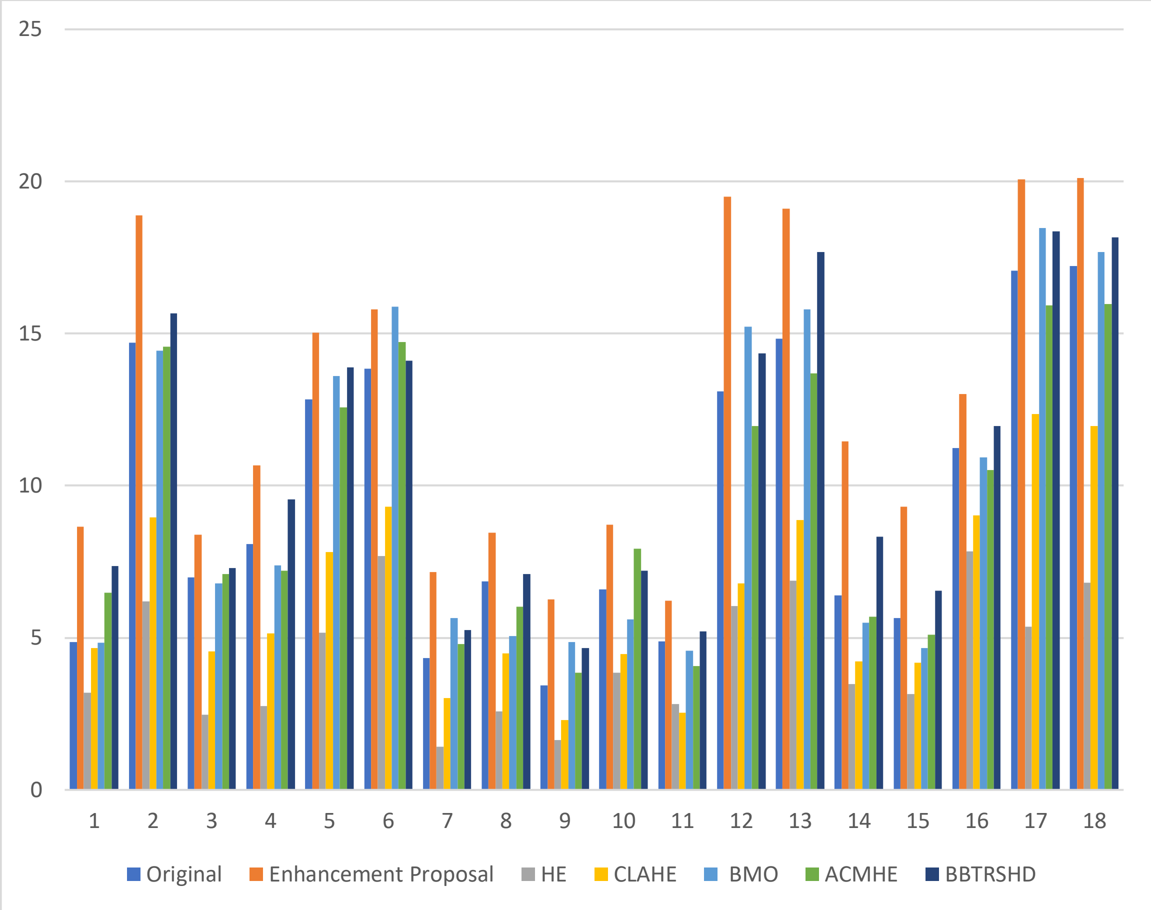


Figure 5- 11: Graph of SNR results of the proposed method compared to state-of-the-art methods

Table 5. 6: Mean and Standard Deviation of SNR results

	Original	Proposed	HE	CLAHE	BMO	ACMHE	BBTRSHD
Mean	9.604	12.596	4.407	6.369	9.830	9.341	10.701
Standard Deviation	4.551	5.001	2.040	3.026	5.113	4.232	4.670

5.2.3 VDIQA

In this section, the results based on the proposed metric VDQAM in Shahkolaei et al. (2018) will be discussed. This metric is a no-reference metric based on the statistics of the mean subtracted contrast normalized coefficients computed from segmented layers of each document image as previously discussed in Chapter 3. The segmentation is based on the hypothesis that the sensitivity of the human vision system (HVS) is different at the locations of text and non-text. The proposed metric is ideal for document images that contain different types of severity of physical distortions. As it mimics the HVS, this is an ideal metric to further verify the fidelity of the proposed enhancement. In this evaluation, the pre-trained model for 80% training and 20% test on dataset is adopted from the list of available models published in their paper and available online. This was adopted due to the correlation of the study by Pham et al. (2020) where an increase in training set size could also enhance the testing performance. Since the system is designed to mimic the HVS based off the Mean opinion score, the default score range for the metric in this experiment is set between 1 to 8, with 1 being the lowest score and 8 the highest ranging from bad to excellent quality.

Table 5.7 shows the VDIQA performance from the experiment. An increase in value denotes better enhancement and vice versa. Furthermore, the metric correlates with the MOS experiment values in Table 5.3 according to performance. The mean of the performance shown in table 5.8 shows that the proposed method performs better than the other implemented techniques with a value of approximately 5.41. The graphical representation of the experiment is also denoted in Figure 5-12. Furthermore, the bar chart in Figure 5-13 shows the number of images where each method performed best in each condition. The proposed method performed best in the instance of 12 images (66.7% of the total sample images). Compared to other methods, BMO and BBTRSHD performed best in approximately 11.1% and 22.2% of the sample respectively. HE, ACMHE and CLAHE on the other hand both did not perform best in every image tested in this experiment. It was observed that BBTRSHD performed best in instances where there was presence of strong bleed-through/show-through artefacts in the images.

Table 5. 7: VDIQA metric results of the proposed method compared to state-of-the-art methods.

Image ID	Original	Proposed	BMO	ACMHE	HE	CLAHE	BBTRSHD
1	3.4456	**5.5172**	3.6257	3.7253	2.8739	3.8150	5.1166
2	3.8260	**5.7880**	3.7253	3.8854	2.4535	3.0536	4.8574
3	3.3646	5.4177	3.6146	3.5644	3.0835	3.2438	5.1475

4	3.9361	5.5656	5.9876	3.8574	2.9137	3.1445	4.0362
5	3.8171	5.2174	4.6471	3.8863	3.2648	3.3651	4.2572
6	2.9443	2.8742	3.4144	2.8542	3.1439	2.4338	3.2143
7	3.8350	5.6782	3.4154	3.6147	3.1335	3.2136	4.5367
8	4.7782	7.2599	6.5198	5.5487	6.6193	3.3198	5.3286
9	3.6476	7.0492	3.7578	3.5374	3.2435	3.3435	4.6182
10	2.5148	3.2257	2.9547	2.8146	2.2139	2.1241	2.4147
11	3.6449	5.0766	3.5548	3.4746	3.2541	3.0142	5.3567
12	3.4159	4.7664	3.3556	3.2753	2.4339	2.2338	4.8179
13	2.9351	4.8565	2.9945	2.8143	2.2335	2.1237	5.0672
14	3.1141	3.9852	3.3441	3.4741	3.0640	3.1241	3.4143
15	3.4258	4.5468	3.3753	3.2250	3.2135	3.3735	3.8560
16	3.8478	5.8684	3.9581	4.2474	3.5463	2.9859	4.4683
17	3.3158	7.2191	3.9163	3.8060	2.8448	3.6462	4.9670
18	3.8573	7.4195	4.0578	3.6371	3.4670	3.7871	7.9297

Table 5. 8: Mean of VDIQA

Original	Proposed	BMO	ACMHE	HE	CLAHE	BBTRSHD
3.537	5.4073	3.7899	3.6245	3.1667	3.074	4.6113

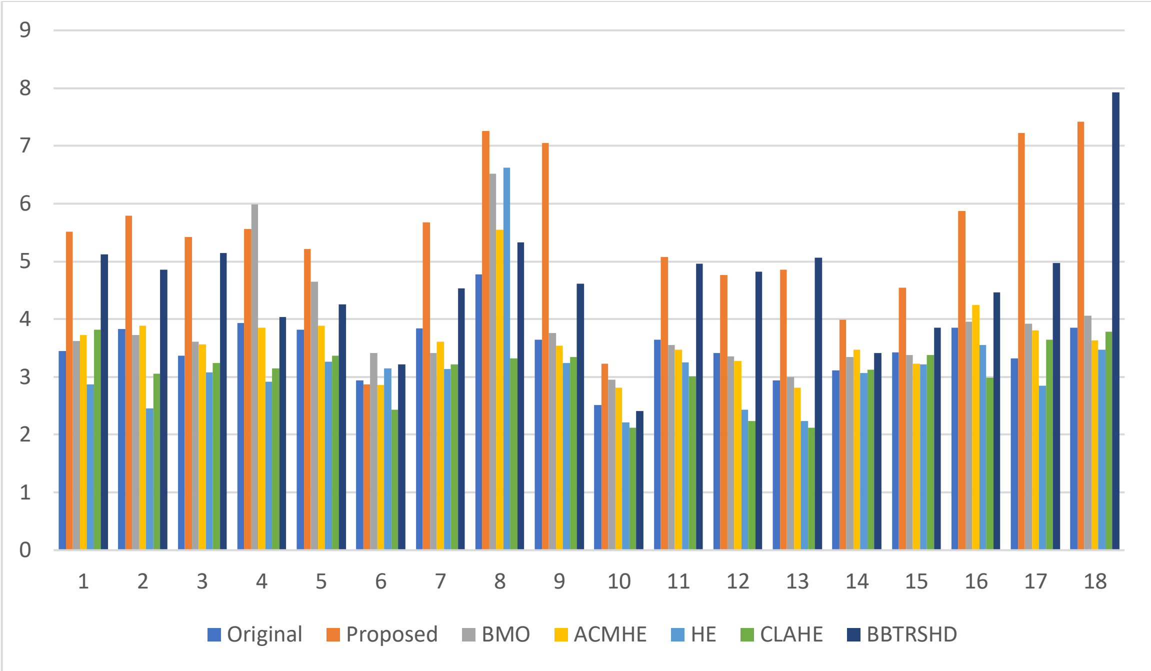


Figure 5- 12: Graph of VDIQA performance per method

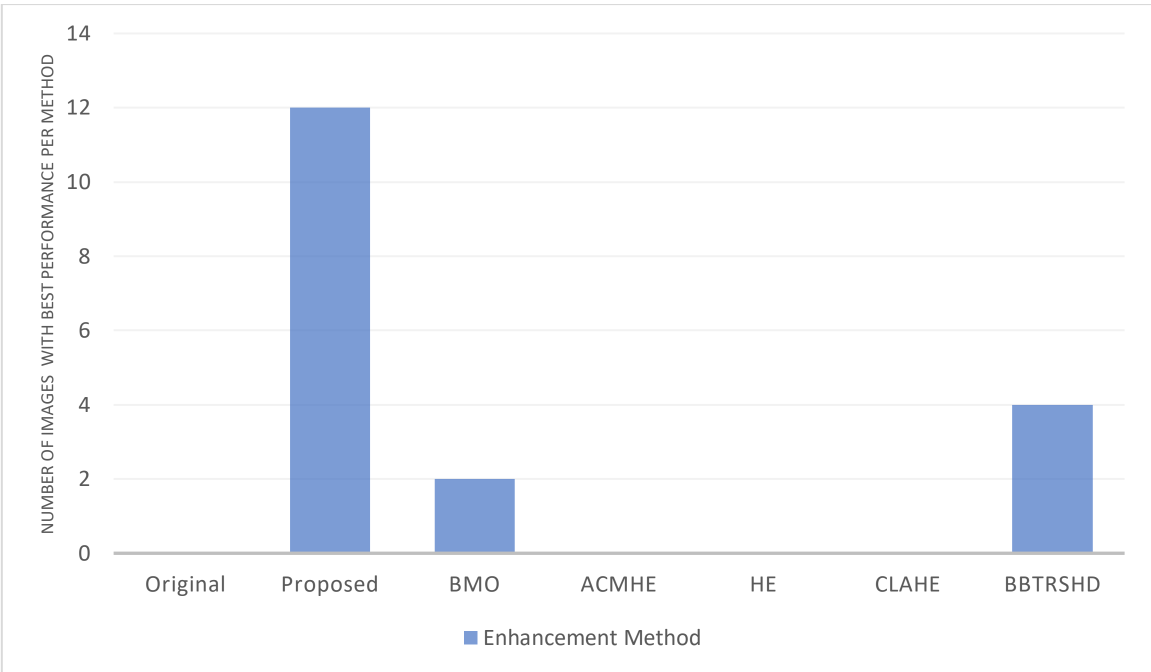


Figure 5- 13: Bar chart of distribution of VDIQA best performance per method

Summary

The proposed method has proven its efficiency across all performance evaluation metrics used (Mean opinion score MOS, signal to noise ratio SNR and visual display image quality assessment VDIQA) against state-of-the-art methods HE, CLAHE, BMO, ACMHE, and BBTRSHD.

5.3 Post-processing Evaluation (Binarization Metrics)

To further validate the outcomes of the proposed image enhancement method, the enhanced subset test images are binarized using Otsu (1979), Niblack (2003), Sauvola and Pietikäinen (2000), Bernsen (1986), Bradley and Roth (2007), Lu et al. (2018) and Howe (2013) method.

5.3.1 Performance Metrics to Evaluate Binarization Methods

There exists a wide spectrum of performance metrics for evaluating binarization methods. In the state-of-the-art reviews presented, the main performance benchmark metrics frequently used by different authors are F-measures, Pseudo F-Measure (FMp), Peak Signal to Noise Ratio (PSNR), Misclassification Penalty Metric (MPM), Negative Rate Metric (NRM), Average Quality Score and Distance Reciprocal Distortion (DRD).

In this research, three very prominent metrics used in literature (Pratikakis et al., 2017) are selected: Peak Signal-to-Noise Ratio (PSNR), F-Measure and Negative Rate Metric (NRM).

Peak Signal-to-Noise Ratio (PSNR)

The PSNR measures the amount of signal in comparison with the amount of noise(Pratikakis et al., 2017). The higher the values of PSNR the better the performance and quality of the signal over the noise. Concerning document binarization, PSNR provides a measure of the quality of binarization against ground truth image, and it can be measured as given in equation (5).

$$PSNR = 10 * log10\left(\frac{MAX^2}{MSE}\right) \quad (5)$$

Where MSE is the Mean Square Error between the binary image B and the Ground Truth GT and can be computed as:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}(B(i,j) - GT(i,j))^2 \quad (6)$$

F-measure

This metric depends on the recall and precision metric (Pratikakis et al., 2017). It is a unified form of representing both metrics. The higher the F-measure value the better the performance; its maximum value is 1. It is mathematically denoted in equation (7), (8) and (9) as:

$$F - measure = \frac{2*recall*precision}{recall+precision} \tag{7}$$

Where recall and precision are:

$$recall = \frac{TP}{TP+FN} \tag{8}$$

$$precision = \frac{TP}{TP+FP} \tag{9}$$

where TP, TN, FP, and FN denote the number of true positives, true negatives, false positives, and false negatives, respectively.

Negative Rate Metric (NRM)

The negative rate metric (NRM) calculates actual mismatches of pixels between the results output and ground truth images based on a pixel-by-pixel estimation. Also, it is a ratio of total false-negative and total false-positive pixels in combination, and it can be expressed as in the following equation. In contrast with previous metrics, the lower the NRM, the better the binarization result.

$$NRM = \frac{\frac{FN}{FN+TP}+\frac{FP}{FP+TN}}{2} \tag{10}$$

Metric to Evaluate OCR performance

One of the most crucial goals in document binarization is improving the Optical Character Recognition (OCR) performance. Therefore, a metric to measure how OCR is improving its performance after binarization of document images is desired. As is the case of this research, the OCR is analysed before enhancement and compared with OCR after enhancement to measure how much the enhancement can improve OCR results. There are several metrics (Karpinski et al., 2018) to measure OCR performance in terms of textual content recognition. The metric proposed is based on the edit distance also known as the Levenshtein distance (Haldar & Mukhopadhyay, 2011). This is a way of quantifying how dissimilar two texts are to one another by counting the minimum number of operations required to transform one text into the other. In other words, the model checks for how many edit operations are required to make both texts equal. Those operations are removal, insertion, or substitution of a character in the text.

Edit distance is a widespread computational algorithm (Navarro, 2001) across the literature. The lower the edit distance the better similarity of both texts. One text is the output of the OCR and the other one is the ground truth. However, we propose a metric based on edit distance, Edit Distance Rate (EDR):

$$EDR \ = 100 * \frac{maxED - ED(ocrText, GroundTruthText)}{maxED} \qquad (11)$$

Where *maxED* is the maximum edit distance possible, ocrText is the output text after OCR recognition engine, *GroundTruthText* is the *Ground Truth Text* expected in the output. Finally, this metric will show a percentage of similarity for both texts based on edition operations.

5.4 Implementation of the Binarization Methods

This section explains the procedures followed for implementing the list of chosen methods. Some algorithms are already in-built MATLAB functions while others were available in the MATLAB community or implemented by research. The Otsu's and Bradley's methods are classical image segmentation techniques available as MATLAB function. In contrast, Niblack's, Sauvola's and Bernsen's are based on the works (Motl, 2013; Niblack, 2003), (Najafi & Salehi, 2015; Sauvola & Pietikäinen, 2000), and (Bernsen, 1986; Eyupoglu, 2017) respectively. Also, Howe's method is derived from the author's website (Howe, 2013) while Dilu's method was implemented from the instructions shown in the author's published paper.

In Bradley's MATLAB implementation, the parameter sensitivity that indicates sensitivity towards thresholding more pixels are foreground. According to some local experiments, 0.4 is a good performance value in the context of this application.

This method has two parameters k_1 & k_2. The original paper does not provide values for both parameters, therefore an experiment to figure out if it is existing as a particular optimized combination is carried out. In doing so, a variation from $k_1, k_2 \in [0,1.5]$ was carried out, and the pair combination with the maximum F-measure selected. The optimized values are $k_1 = 0.2$ $k_2 = 1.5$. Particularly, this experiment shows the same results for various images and consequently, these are the values of the parameters used for all further experiments.

5.5 Experiment results and discussion

In this section, the binarization methods are tested on the subset created in this research to showcase research findings on the binarization process on different artefacts. The collection of the ENP subset selected in this research comprises of folders with 6 artefacts: smeared ink, low scan contrast, faded ink, uneven illumination, broken characters and show through, with each folder comprising of the set of images, and their binarization ground truth. The subset also contains the text ground truth, for measuring OCR performance. Table 5.9 provides information on the number of images per artefact selected. The binarization methods deployed over each set of images are Otsu, Niblack, Sauvola, Bernsen, Bradley and Howe. Also, the binarization results for each method and the measures PSNR, F-measures and NRM are shown. The results for smeared ink, low scan contrast, faded ink, uneven illumination, broken characters, show through can be depicted in Table 5.10, Table 5.11, Table 5.12, Table 5.13 Table and Table 5.15 respectively. On the other hand, Table shows the average of all these experiments.

Table 5. 9: Artefact's selections and quantity

Artefacts	Quantity	Artefacts	Quantity
smeared ink	17	uneven illumination	28
low scan contrast	7	broken characters	151
faded ink	35	show through	131

Table 5. 10: Smeared Ink

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	64.01	64.09	62.43	57.42	57.42	63.32	65.74
F-measure	0.88	0.93	0.89	0.74	0.74	0.89	0.97
NRM	0.09	0.06	0.09	0.20	0.20	0.08	0.05

Table 5. 11: Low scan contrast results

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	67.74	66.25	62.09	56.00	66.03	64.80	66.73
F-measure	0.93	0.94	0.88	0.65	0.94	0.91	0.89
NRM	0.05	0.04	0.11	0.26	0.06	0.05	0.02

Table 5. 12: Faded ink

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	58.84	63.84	60.02	56.40	59.13	58.41	67.04
F-measure	0.76	0.92	0.81	0.68	0.78	0.73	0.95
NRM	0.18	0.07	0.15	0.24	0.17	0.19	0.02

Table 5. 13: Uneven Illumination

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	61.60	63.14	61.53	56.54	63.27	60.98	71.55
F-measure	0.86	0.91	0.88	0.70	0.90	0.84	0.96
NRM	0.11	0.08	0.11	0.23	0.08	0.11	0.02

Table 5. 14**:** Broken Characters

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	66.12	64.26	63.56	57.80	65.91	63.98	66.73
F-measure	0.93	0.92	0.91	0.73	0.94	0.90	0.89
NRM	0.05	0.07	0.08	0.21	0.04	0.06	0.03

Table 5. 15: Show through

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	65.28	64.18	63.00	57.49	64.86	63.01	66.60
F-measure	0.92	0.92	0.90	0.73	0.93	0.89	0.96
NRM	0.06	0.06	0.09	0.21	0.05	0.07	0.22

Table 5. 16: Average of all artefacts

	Otsu	Niblack	Sauvola	Bernsen	Bradley	Dilu	Howe
PSNR	63.93	63.75	62.10	56.94	64.29	62.16	67.40
F-measure	0.88	0.90	0.88	0.70	0.92	0.85	0.94
NRM	0.09	0.06	0.10	0.22	0.06	0.10	0.06

In the above tables, the results corresponding to the three measures presented in each subset of the document images for all the methods is shown. The values highlighted in red, and green are the first- and second-best results respectively. These tables provide an experimental approach to discern which methods are best in several artefacts in the subset.

The average results showed that the best two methods are Howe and Bradley as seen in Table . In overall performance, these methods seem to stand out in this dataset. It has been shown by other authors how well Howe's method performs and how its results perform remarkably in

DIBCO dataset. However, Bradley's method stood out in second place unprecedentedly and it was discovered that this method has not been tested enough on ancient document binarization due to its lack of literature, but it has shown considerable performance in this experiment. An argument to consider is that the Bernsen's method performs very poorly in most of the experiments and therefore it is suggested limiting the use in more state-of-the-art experiments.

5.6 Enhancement Validation

In this section, the proposed method is compared with the state-of-the-art methods used in this research and the ground-truth in terms of how they perform with binarization. The assessment of the enhanced image, and other state-of-the-art methods with the ground truth is measured with 4 metrics namely; F-measure(Pratikakis et al., 2017), Negative Rate Metric (NRM), and Peak Signal to Noise Ratio (PSNR)(Hore & Ziou, 2010)

5.6.1 Binarization Evaluation (Subjective)

A visual example is illustrated in figure 5-14, 5-15 and 5-16 of image 00674388 with noticeable presence of uneven illumination, faded ink, faint character and noise. The images show a side-by-side evaluation of the original image, proposed enhancement image and the subsequent binarized image for both as applied to binarization methods Otsu, Niblack, Sauvola, Bernsen, Bradley, Dilu and Howe. From the example, the proposal shows a more improved binarization result for all binarization methods tested. Image examples with other degradations like bleed-through are shown in the index section.

Original Image


Gef. sof. tücht. Frau für Säcke flicken u. stopfen. J. Boldt, Marthastr. 21, 1. Etage.

Arbeiterinnen gesucht. Dampfwäscherei Te Korb, Steilshooperstraße 136 a–140.

Ges. eine geübte Arbeiterin. Plathe, Konf.-Fabrik, Müggenkampstraße Nr. 58.

Gesucht sofort tüchtige Frauen als Zeitungsträgerinnen für Touren in der Inneren Stadt

Zuverläs Zeitungsträge für den Be St. Geo gesucht. Zu meld

Damen-Ka gesucht. K. Freihe

Geübte Tütenkleber ebenfalls eine ti

Enhanced Image Proposal

Binarization Ground-truth

No previous enhancement: Otsu

With the enhancement: Otsu

Figure 5- 14: Visual results of the Binarization on image 00674388; (a) original image (b) Enhanced Proposal (c) Ground-truth (d) Otsu (Before enhancement) (e) Otsu (After Enhancement

No previous enhancement: Niblack

With the enhancement: Niblack

No previous enhancement: Sauvola

With the enhancement: Sauvola

No previous enhancement: Bernsen

With the enhancement: Bernsen

Figure 5- 15: Visual results of the Binarization on image 00674388; (a) Niblack (Before enhancement) (b) Niblack (After enhancement) (c) Sauvola (Before enhancement) (d) Sauvola (After enhancement) (e) Bernsen (Before Enhancement) (f) Bernsen (After enhancement)

No previous enhancement: Bradley

With the enhancement: Bradley

No previous enhancement: Dilu

With the enhancement: Dilu

No previous enhancement: Howe

With the enhancement: Howe

Figure 5- 16: Visual results of the Binarization on image 00674388; (a) Bradley (Before enhancement) (b) Bradley (After enhancement) (c) Dilu (Before enhancement) (d) Dilu (After enhancement) (e) Howe (Before Enhancement) (f) Howe (After enhancement)

The subsequent result from this experiment shows that the improved font intensity and elimination of noise after enhancement subsequently improves the binarization result visually as shown. Artefacts like noise, low font intensity and show-through that impact binarization were significantly diminished with the enhancement method. The appendix section of this report shows some of the images.

To objectively show the improvement of binarization, the next section validates the method through binarization performance metrics.

5.6.2 Binarization Evaluation (Objective)

To further validate the enhancement method, the enhanced images from the subset are binarized using Otsu's binarization and compared with the original image and ground-truth using 3 metrics, namely; Peak Signal Noise Ratio (PSNR) (Pratikakis et al., 2017), F-Measure (Pratikakis et al., 2017) and Negative Rate metric.

The table 5.8 shows the objective results for image 00674388 (Figure 5-14) with regards to PSNR, F-Measure and NRM for all binarization methods implemented. From the results, the enhancement method performed better than the original in all Binarization methods, regarding the characterized image containing degradation artefacts uneven illumination, faded ink, faint character and noise for PSNR, F-Measure and NRM. Although with improved performance from original to the enhanced image, Otsu and Howe recorded the lowest gain when applied to the enhancement method for this degradation as compared to the original image.

Table 5. 17: Evaluation of Image 00674388 (Original and Enhanced) using multiple Binarization methods

Method	PSNR		F-Measure		NRM	
	Original	Enhanced	Original	Enhanced	Original	Enhanced
Otsu	64.948	66.483	0.95994	0.96815	0.030607	0.026882
Niblack	65.242	67.065	0.96271	0.97725	0.029951	0.021208
Sauvola	61.217	66.753	0.9126	0.97379	0.080292	0.029032
Bernsen	60.312	**66.876**	0.89452	**0.97685**	0.095384	**0.032165**
Bradley	64.879	66.636	0.95985	0.97083	0.034619	0.02593
Dilu	63.758	66.483	0.94536	0.96815	0.028466	0.026882
Howe	61.758	64.882	0.91952	0.94135	0.067347	0.058955

The result in Table 5.17 denotes the result on an image 00675802 with show-through/ bleed-through, and low-contrast, the enhancement also performed better in all Binarization methods tested. Bernsen method which performs least also performed marginally better when comparing the original image PSNR, F-Measure and NRM to the enhanced result output.

Table 5. 18: Evaluation of Image 00675802 (Original and Enhanced) using multiple Binarization methods.

Method	PSNR		F-Measure		NRM	
	Original	Proposed	Original	Proposed	Original	Proposed
Otsu	76.92	78.899	0.9896	0.99289	0.0030966	0.002562
Niblack	67.164	79.164	0.91087	0.99774	0.081742	0.029378
Sauvola	68.851	79.178	0.93779	0.99787	0.058564	0.028813
Bernsen	54.767	70.493	0.37055	0.74225	0.38629	0.23863
Bradley	70.487	79.097	0.95646	0.99643	0.04165	0.027616
Dilu	72.694	78.888	0.97185	0.99264	0.001888	0.001519
Howe	65.62	75.063	0.87229	0.90873	0.099638	0.098376

5.6.3 Subset Binarization results

This section describes the evaluation of the enhancement technique with respect to the artefacts in the ENP subset. These include 6 artefacts namely, smeared ink, uneven illumination, low scan contrast, faded ink, show-through, and broken-characters. The original image, enhancement method, the second and third best performing methods (BMO and BBTRSHD) in Table 5.6 and 5.8 in terms of SNR and VDIQA performance were selected for experimentation. In turn, the most popular global binarization method (Otsu), and the most recent (Dilu) among the methods is used to evaluate the methods based on PSNR, F-Measure and NRM. The mean result of each artefact selection is presented as result. This section describes the evaluation of the enhancement technique in respect to how it responds to different degradations.

Table 5. 19: Smeared Ink

Binarization Method	**Otsu**				**Dilu**			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	64.01	72.52	65.82	68.32	63.32	68.45	66.32	67.80
F-Measure	0.88	0.93	0.90	0.91	0.89	0.96	0.91	0.92
NRM	0.09	0.06	0.08	0.07	0.08	0.06	0.07	0.07

Table 5. 20: Low scan contrast

Binarization Method	**Otsu**				**Dilu**			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	67.74	79.66	65.82	68.32	64.80	70.45	67.22	68.50
F-Measure	0.93	0.97	0.92	0.94	0.91	0.97	0.93	0.94
NRM	0.05	0.04	0.05	0.05	0.05	0.03	0.05	0.04

Table 5. 21: Faded ink

Binarization Method	Otsu				Dilu			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	58.84	72.66	69.82	69.32	58.41	69.32	62.32	59.80
F-Measure	0.76	0.96	0.92	0.94	0.91	0.96	0.91	0.92
NRM	0.18	0.10	0.12	0.17	0.19	0.12	0.14	0.15

Table 5. 22: Uneven Illumination

Binarization Method	Otsu				Dilu			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	61.60	73.14	69.82	69.32	60.98	70.32	64.42	65.80
F-Measure	0.86	0.97	0.88	0.90	0.84	0.97	0.91	0.92
NRM	0.11	0.06	0.08	0.08	0.11	0.05	0.07	0.06

Table 5. 23: Broken characters

Binarization Method	Otsu				Dilu			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	66.12	67.66	66.82	64.78	63.98	69.32	67.32	66.80
F-Measure	0.93	0.93	0.92	0.94	0.90	0.96	0.91	0.92
NRM	0.05	0.08	0.09	0.11	0.06	0.05	0.07	0.07

Table 5. 24: Show-through

Binarization Method	Otsu				Dilu			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	65.28	72.66	69.82	70.85	63.01	76.32	62.32	73.80
F-Measure	0.92	0.96	0.94	0.94	0.89	0.96	0.93	0.93
NRM	0.06	0.04	0.05	0.04	0.07	0.06	0.07	0.07

Table 5. 25: Average of all artefacts

Binarization Method	Otsu				Dilu			
Enhancement method	Original	Proposed	BMO	BBTRSHD	Original	Proposed	BMO	BBTRSHD
PSNR	63.93	73.05	67.99	68.48	62.16	70.70	64.98	67.08
F-Measure	0.88	0.95	0.91	0.93	0.85	0.96	0.92	0.93
NRM	0.09	0.06	0.08	0.08	0.10	0.06	0.08	0.07

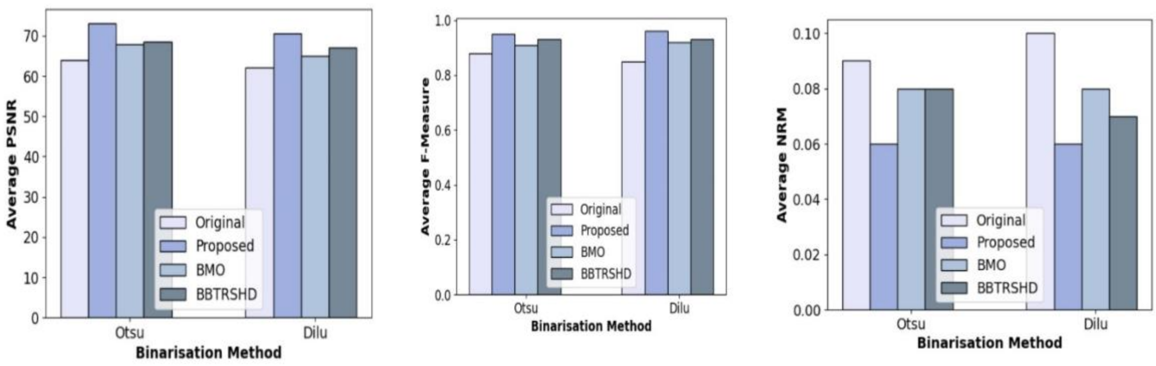


Figure 5- 17: Graph of average PSNR, F-Measure and NRM of all artefacts

The average result among the tested images shown in table 5.25 and figure 5-17 demonstrates that the proposed method improves the quality of the images better than other enhancement methods tested on the artefacts. From the results, the most significant improvements were recorded in the artefacts smeared ink, low scan contrast, faded ink, uneven illumination and show-through as shown in tables 5.19, 5.20, 5.21, 5.22, 5.23 and 5.25. The improvements in all 3 metrics are significant in terms of PSNR, F-Measure and Negative Rate metric. The lower NRM indicates a better binarization result by indicating lower mismatch of pixels between enhanced image and ground-truth. The graphs of the rest of the results are attached in the appendix.

Limitations of the Proposed Method on Different Degradations

The ENP dataset mainly contains degradations like slight broken characters, bleed-through, uneven illumination, faint characters, and random noise as shown in the tables 5.9. The enhancement method may not be ideal in conditions of broken character, very faded ink (some faded ink is approximated to background), and severe show-through/ bleed-through as shown in their relatively lower performance gain margin in those artefacts.

5.7 Summary

This chapter describes the subset by artefact type, the evaluation of the original images and the enhanced in terms of PSNR, NRM and F-Measure. This evaluation section addresses the quantitative and qualitative evaluation objective of this research and the limitations. From the results, the enhancement method performed best in all artefacts tested. Multiple experiments were carried out to validate the proposed technique with the experiment outcomes presented, described and discussed in this chapter.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In summary, this research introduces an adaptive method for historical document library enhancement that is functional for most degradation artefacts. In the introductory chapter, the background and motivation of the study and insight to the challenges in PRImA historical document image libraries (Europeana newspaper) is presented. It also depicts the project's research problems, aims and objectives. The second chapter describes the background of the ENP dataset, challenges, typical degradations, and artefacts to identify the problems. The third chapter shows that despite the success of enhancement methods in the literature, several challenges are still prevalent which have motivated this research study. To identify these challenges and propose solutions, an in-depth review of the literature was conducted. Challenges of historical document image libraries, image enhancement evaluation methods, noise filtering methods, enhancement approaches based on contrast, enhancement based on hybrid and CNN are reviewed. This demonstrates insight into the ideas, techniques, and contributions provided by researchers from the proposed research area. These problems include restriction of enhancement to single degradation or artefact type and not multiple degradations, enhancement methods do not improve human vision and machine vision simultaneously in most cases.

To address these challenges, a hybrid novel adaptive enhancement method called adaptive histogram equalization is proposed in Chapter 4 of this research study to accommodate other degradations namely uneven ink distribution, bleed-through ink, faded ink and low scan contrast. Some other existent noise filtering pre-operation tools were employed as part of the process for a more robust result. A subset of the ENP is also created for effective testing and evaluation of the method. An implementation model and a subset of images with a cross-section of degradations are implemented to show the proposal performance.

The experimental results in Chapters 4 and 5 show promising performance and efficiency of the proposed method when evaluated with SNR, no reference VDIQA and MOS against state-of-the-art methods, histogram equalization, and Contrast Limited Adaptive Histogram Equalization (CLAHE). The other more recent methods tested against the proposed method are adaptive contrast enhancement using modified histogram equalization (ACMHE), Blind bleed-through removal for scanned historical documents (BBTRSHD) and Barnacles Mating Optimizer (BMO) methods.

6.2 Research limitations and Future Work

The proposed method exhibits commendable accuracy and adaptability; however, certain limitations warrant attention for future improvements. One prominent constraint lies in its dependence on approximating foreground and background components through identified pixel intensities. This approach encounters challenges when confronted with pages exhibiting intense show-through or extremely faded ink. In cases of intense show-through, the system might erroneously approximate artefacts as text, while it could tend to approximate towards the background in instances of severely faded ink.

Future endeavors in this research area could prioritize evaluating and measuring the efficacy of the image enhancement technique against more contemporary, well-cited methodologies outlined in the existing literature. This comparative analysis would facilitate a comprehensive understanding of the proposed technique's strengths and weaknesses in contrast to the latest advancements in image enhancement.

Further avenues for future work might include:

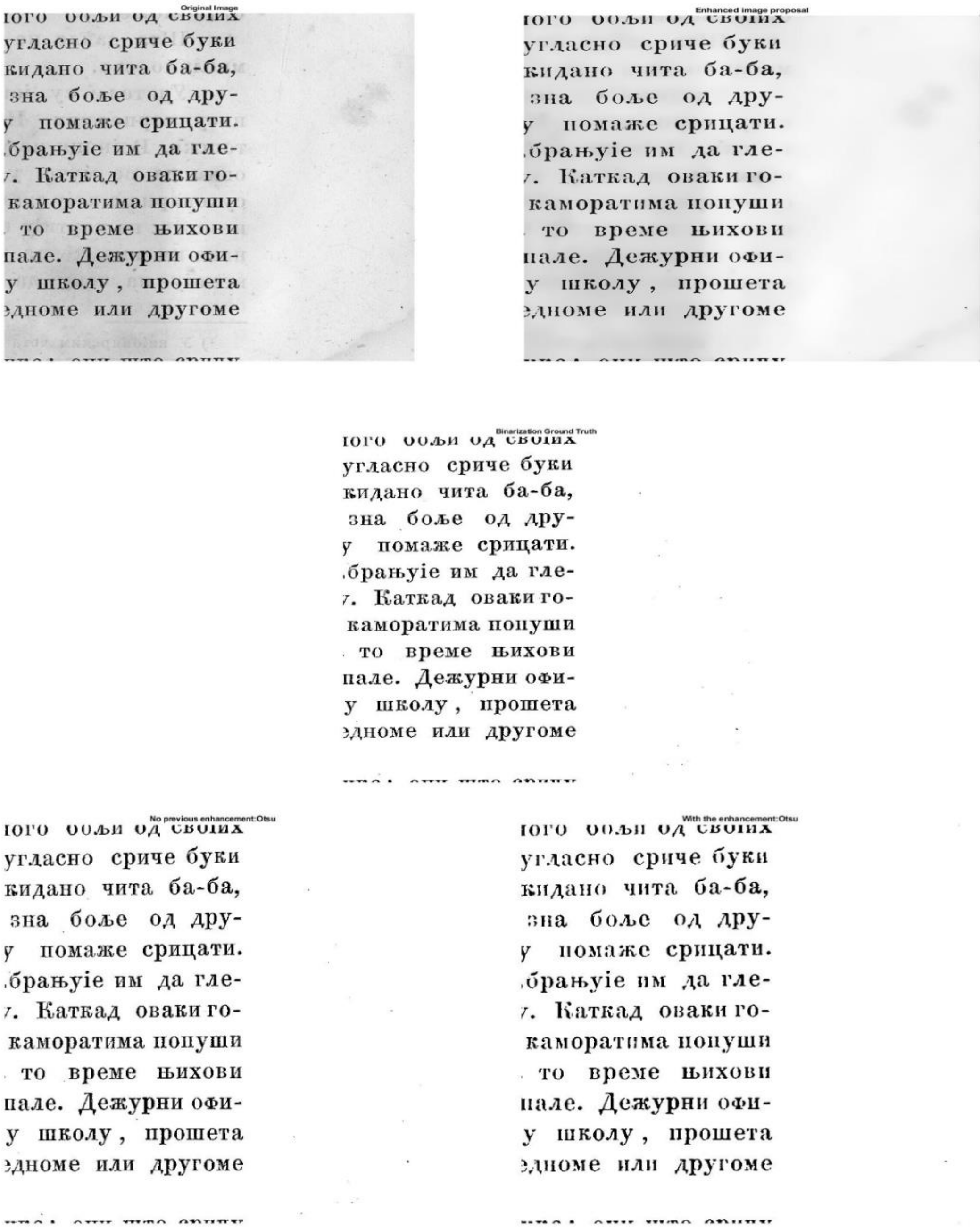
1. Refinement of Foreground-Background Separation: Developing more sophisticated algorithms or models for robustly distinguishing between foreground and background elements, especially in scenarios involving intense show-through or faded ink, to improve accuracy in image enhancement.
2. Adaptive Enhancement Strategies: Designing adaptive enhancement strategies that dynamically adjust parameters based on the specific characteristics of the input image, thus enabling the system to address varying degrees of show-through or faded ink more effectively.
3. Integration of Advanced Image Processing Techniques: Exploring and integrating cutting-edge image processing methodologies, such as deep learning-based approaches or advanced feature extraction techniques, to enhance the system's capability to handle diverse image characteristics.
4. Real-World Testing and Validation: Conducting extensive real-world testing using a diverse dataset comprising a wide range of document types and conditions to ensure the method's applicability and robustness across different scenarios.

5. Generalization and Scalability: Focusing on refining the technique for broader generalization across various historical document types and scalability to handle larger datasets efficiently.

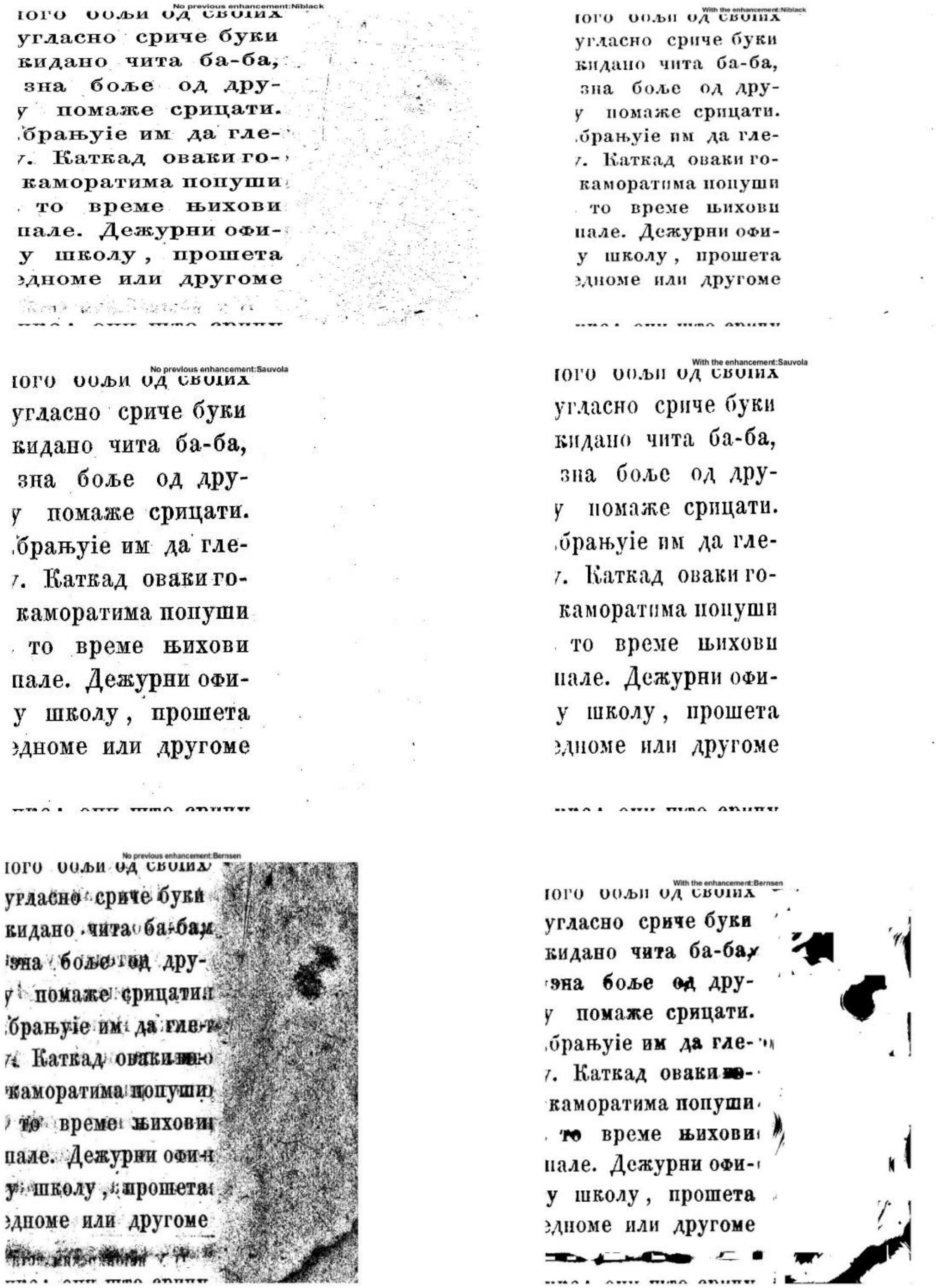
Addressing these future research directions could significantly enhance the proposed image enhancement method, making it more reliable, versatile, and capable of handling diverse document characteristics encountered in practical applications.

Appendices

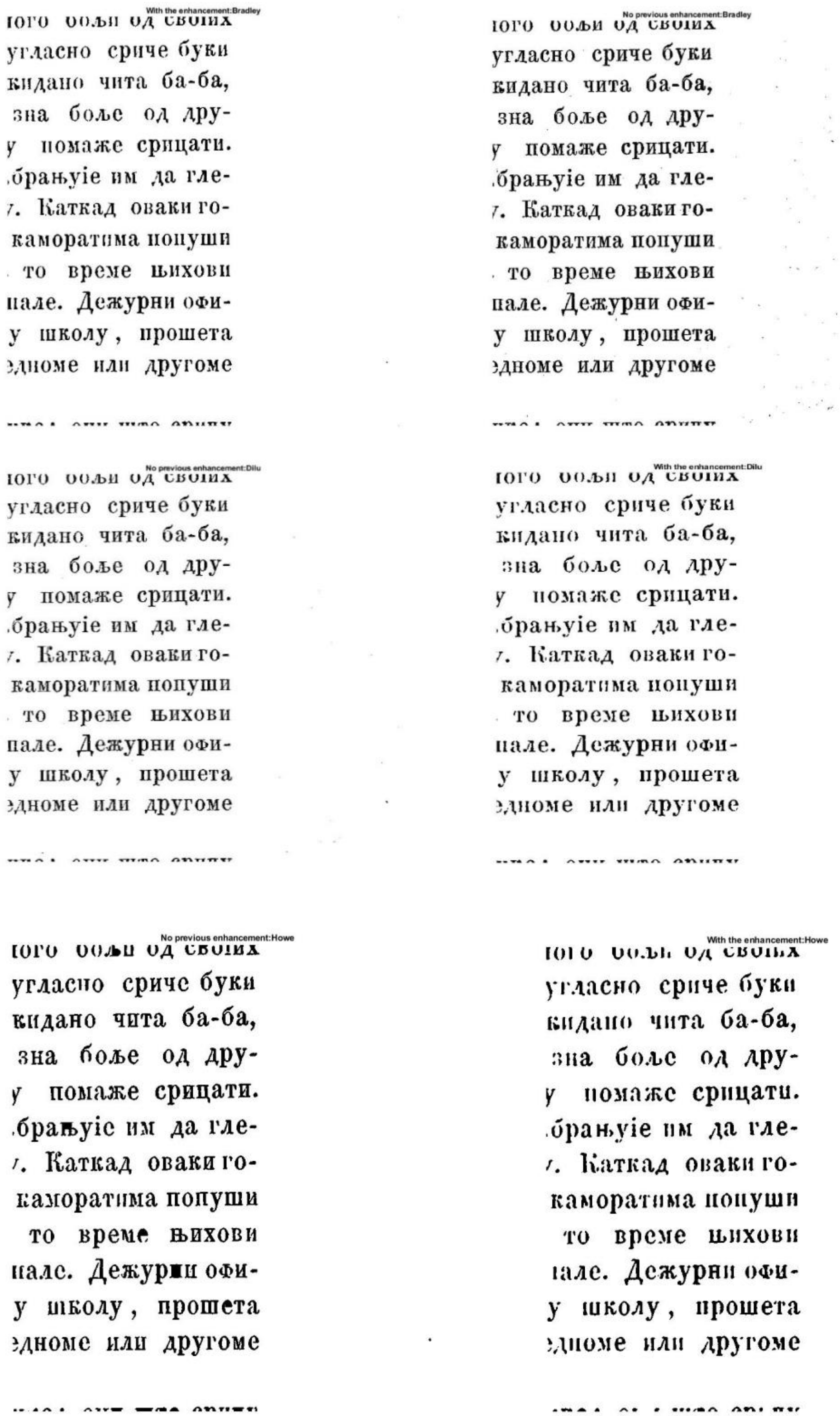
Appendix A



Visual results of the Binarization on image with Bleed-through/ Show-through; (a) original image (b) Enhanced Proposal (c) Ground-truth (d) Otsu (Before enhancement) (e) Otsu (After Enhancement)

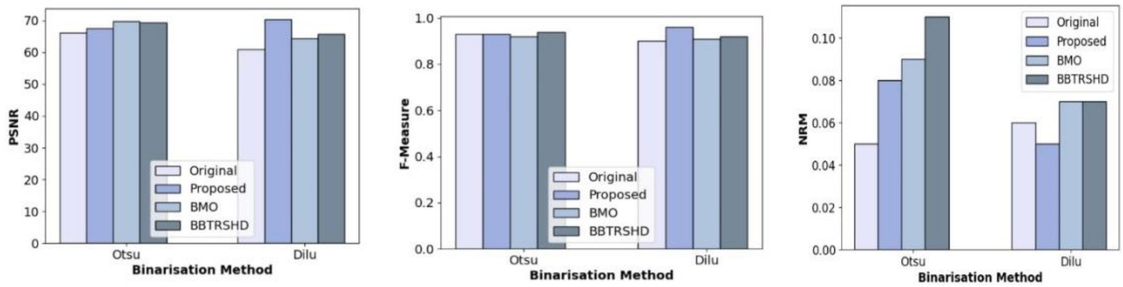


Visual results of the Binarization on Bleed-through/ Show-through; (a) Niblack (Before enhancement) (b) Niblack (After enhancement) (c) Sauvola (Before enhancement) (d) Sauvola (After enhancement) (e) Bernsen (Before Enhancement) (f) Bernsen (After enhancement)

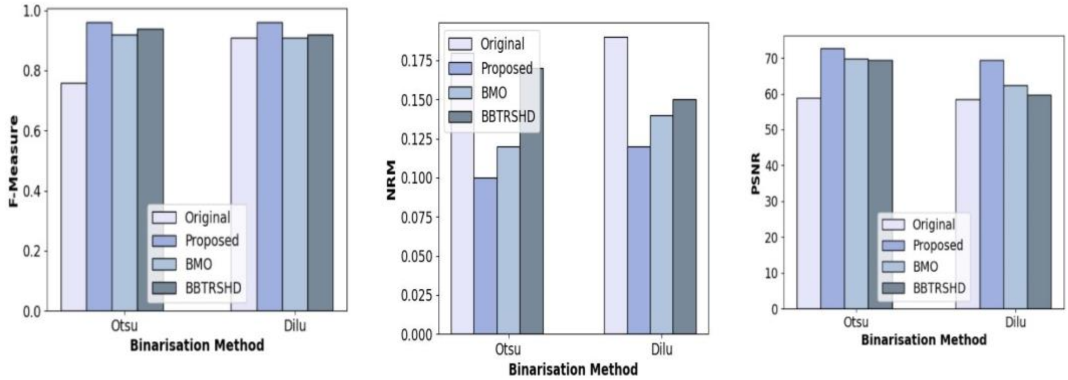


Visual results of the Binarization on Bleed-through/ Show-through; (a) Bradley (Before enhancement) (b) Bradley (After enhancement) (c) Dilu (Before enhancement) (d) Dilu (After enhancement) (e) Howe (Before Enhancement) (f) Howe (After enhancement)

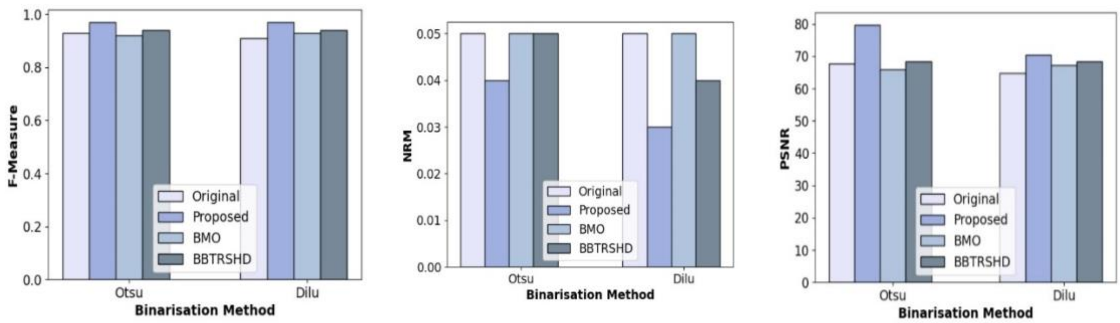
Appendix B



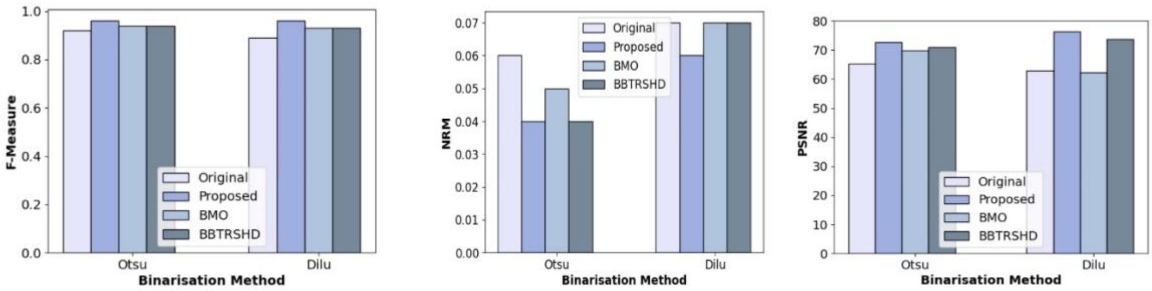
Graph of average PSNR, F-Measure and NRM on Broken Characters artefacts



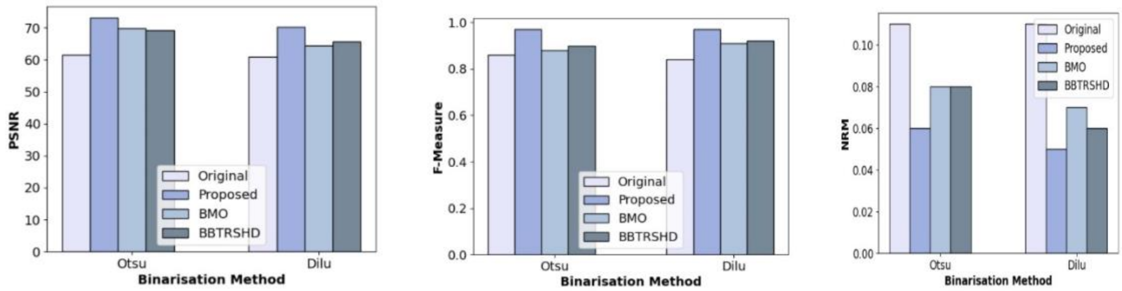
Graph of average PSNR, F-Measure and NRM on Faded ink artefacts



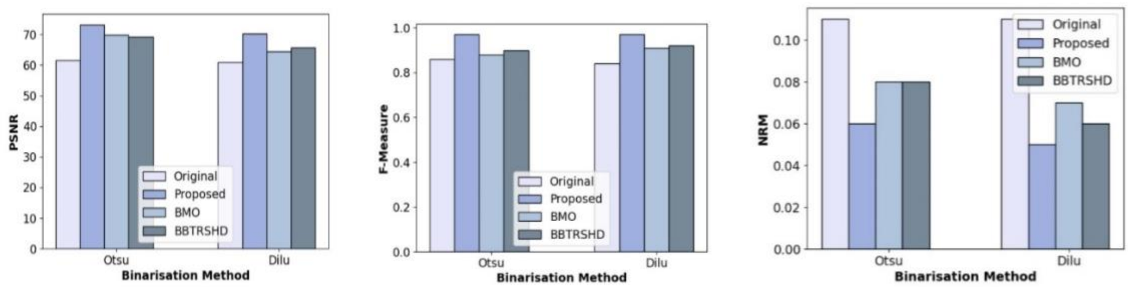
Graph of average PSNR, F-Measure and NRM on low scan contrast artefacts



Graph of average PSNR, F-Measure and NRM for show-through artefacts



Graph of average PSNR, F-Measure and NRM for uneven illumination artefact



Graph of average PSNR, F-Measure and NRM for Smear ink artefact

Appendix C

Source Code

```
%% Visual comparisons

% this script runs several enhancement methods in order to make a visual
% comparision with traditional enhancement techniques against the proposal

clc; clear; close all;

% Adding required functions and paths
addpath('../Standard-Algorithms');
addpath('./VDQAM code');
addpath('./VDQAM code/SVM');

% dataset selected for enhancement
imageDir = '..\ImageDataset\Europeana Newspapers Project
Dataset\EnhancementExperiments';

% folder to save images for the human evaluation
savePath = './HumanVisualEvaluation/';

% creating Datastorage obj
imds = imageDatastore(imageDir);


EnhancementMethods = {'HgramEq','CLAHE','ACMHE','BMO'};

% snr storage
SNRstorage = zeros(length(imds.Files),length(EnhancementMethods)+2); % plus
2 for original snr and enhanced one

% VDQAM metric scores storage
VDQAMscores = zeros(length(imds.Files),length(EnhancementMethods)+2);

% running over each image
for i=1:length(imds.Files)

    name = imds.Files{i};
    newStr = split(name,'\');
    newName = newStr{end};
    newName = split(newName,'.');

    im = imread(name);

    % ensure grayscale
    if(size(im,3)>1)
        im = rgb2gray(im);
    end

    % saving grayscale original version
    imwrite(im,strcat(savePath,newName{1},'.',newName{2}));

    % computing the propousal
    im_enhanced = EnhanceDocument(im);
```

```
    % saving the enhanced version
    imwrite(im_enhanced,strcat(savePath,newName{1},'_Enh','.',newName{2}));

    % original and enhanced SNR
    SNRstorage(i,1) = round(imsnr(im),2);
    SNRstorage(i,2) = round(imsnr(im_enhanced),2);

    % original and enhanced SNR
    VDQAMscores(i,1)  = VDQAMscore(im);
    VDQAMscores(i,2)  = VDQAMscore(im_enhanced);

    % showing original image and enhanced and snr
    %figure; imshowpair(im,im_enhanced,'montage');
    %title([ 'SNR[dB] Original: ' mat2str(SNRstorage(i,1)) '           '
'SNR[dB] enhanced: ' mat2str(SNRstorage(i,2))]);

    % running classical methods
    EnhancedStorage =
zeros(size(im,1),size(im,2),length(EnhancementMethods));

    for j=1:length(EnhancementMethods)

        % computing classical enhancements methods
        classicalEnhanced =
ClassicalEnhanceImage(im,string(EnhancementMethods{j}));

        % saving SNR for each case
        SNRstorage(i,j+2) = round(imsnr(classicalEnhanced),2);

        % saving VDQAMscores
        VDQAMscores(i,j+2) = VDQAMscore(classicalEnhanced);

        % saving output images to the folder

%imwrite(classicalEnhanced,strcat(savePath,newName{1},'_',EnhancementMethod
s{j},'.',newName{2}));

        EnhancedStorage(:,:,j) =
ClassicalEnhanceImage(im,string(EnhancementMethods{j}));

    end

    %figure; montage(EnhancedStorage,'DisplayRange', []);
end

variablesNames = {'original', 'enhancedP','HgramEq','CLAHE','BMO','ACMHE'};

disp('SNR_Table');
SNRTable = cell2table(num2cell(SNRstorage),'VariableNames', variablesNames
);
save SNRTable SNRTable;
disp(SNRTable);

disp('VDQAM_Table');
VDQAMtable = cell2table(num2cell(VDQAMscores),'VariableNames',
variablesNames );
save VDQAMtable VDQAMtable;
disp(VDQAMtable);
```

```
%% Anaylising results

% This section must anaylise results looking for the best methods according
to
% the metrics and check statistic metrics

% finding indexes maximum
[~,ind] = max(VDQAMscores,[],2);

% showing histogram
C = categorical(ind,1:length(variablesNames),variablesNames);
h = histogram(C,'BarWidth',0.5,'Normalization','probability');
h1 = histogram(ind);
m = max(h1.Values);
title('Distribution of better VDQAM per method');

% looking for images indexes to check independent
ImagesNamesFail = cell(length(variablesNames),m);
for i = 1:length(variablesNames)
    idxes = find(C == variablesNames(i));

    for j=1:length(idxes)
        name = imds.Files{idxes(j)};
        newStr = split(name,'\');
        newName = newStr{end};
        ImagesNamesFail{i,j} = newName;
    end

end

ImagesNamesFailTable =
cell2table(num2cell(ImagesNamesFail'),'VariableNames',variablesNames);
disp(ImagesNamesFailTable);
```

```
%% EvaluateEnhance

% This script evaluates the proposal enhancement using as a basis the
% binarization method

clc;clear; close all;

% Adding required functions
addpath('../Standard-Algorithms');

% loading test image and binarized GT
I = imread('00674388.jpg');
GT = imread('00674388-00229940-Gatos.tif');

rect =  [947.5 1143.5 991 783]; % selected area
I = imcrop(I,rect);
GT = imcrop(GT,rect);

% I = imread('H08.png');
```

```
% GT = imread('H08_GT.png');

figure; imshow(I); title('Original Image');
figure; imshow(GT); title('Binarization Ground Truth');

% computing propousal enhancement technique
im_enhanced = EnhanceDocument(I);
figure; imshow(im_enhanced); title('Enhanced image proposal');

% computing binarization using previous methods
Methods = {'Otsu', 'Niblack', 'Sauvola', 'Bernsen', 'Bradley',
'Dilu','Howe'};

% preparing metric variables
PSNR = zeros(length(Methods),2);
Fmeasures = zeros(length(Methods),2);
nrms = zeros(length(Methods),2);

% running
for j = 1:length(Methods)

        docBinarized = binarizeDoc(I,string(Methods{j}));
        docBinarized2 = binarizeDoc(im_enhanced,string(Methods{j}));

        figure; imshow(docBinarized); title(strcat('No previous
enhancement:',string(Methods{j})));
        figure; imshow(docBinarized2);title(strcat('With the
enhancement:',string(Methods{j})));

        PSNR(j,1)  = psnr(uint8(docBinarized),uint8(GT));
        PSNR(j,2)  = psnr(uint8(docBinarized2),uint8(GT));

        Fmeasures(j,1)  = Fmeasure(uint8(docBinarized),uint8(GT));
        Fmeasures(j,2)  = Fmeasure(uint8(docBinarized2),uint8(GT));

        nrms(j,1)  = NRM(uint8(docBinarized),uint8(GT));
        nrms(j,2)  = NRM(uint8(docBinarized2),uint8(GT));
end

disp('PNSR');
cell2table(num2cell(PSNR),'RowNames',Methods,'VariableNames',{'noEnhance',
'enhance'})

disp('Fmeasures');
cell2table(num2cell(Fmeasures),'RowNames',Methods,'VariableNames',{'noEnhan
ce', 'enhance'})

disp('NRMS');
cell2table(num2cell(nrms),'RowNames',Methods,'VariableNames',{'noEnhance',
'enhance'})
```

Appendix D

```
function [im_enhanced] = EnhanceDocument(I)

% This fuction performs the document enhancement method
```

```
% contrast stretching
I = imadjust(I,stretchlim(I),[]);

% Wiener filtering
I = wiener2(I,[5 5]);

% Bilateral filtering
DoS = 218;
Ssigma = 10;
B = imbilatfilt(I,DoS,Ssigma);

% Adaptative Histogram Matching
im_enhanced = AdaptHistMatching(B);

% Median Filtering
m = 3;
im_enhanced = medfilt2(im_enhanced,[m m]);
end
```

Appendix E

```
function [im_enhanced] = AdaptHistMatching(I)

% This function is a novel technique that applies histogram matching using
% two adaptative gaussian windows. The idea is trying to reduce the entropy
% by transforming the histogram of the image in a bimodal histogram

[counts,~] = imhist(I);
threshold = graythresh(I);
graylevel = threshold * 255;

segment1 = counts(1:graylevel);
segment2 = counts(graylevel+1:end);

% figure; bar(binLocations(1:graylevel),segment1); title('first
Distribution');
% figure; bar(binLocations(graylevel+1:end),segment2); title('second
Distribution');

[~,idx1] = max(segment1);
[~,idx2] = max(segment2);

W1 = 1.2;
W2 = 1;
sigma1 = 45;
sigma2 = 45;

N1 = 3*length(segment1);
alpha1 = (N1-1)/(2*sigma1);
gaussWin1 = W1*gausswin(N1,alpha1);

N2 = 3*length(segment2);
alpha2 = (N2-1)/(2*sigma2);
gaussWin2 = W2*gausswin(N2,alpha2);

[~,idxG1] = max(gaussWin1);
[~,idxG2] = max(gaussWin2);

if((idxG1 - idx1)>0)
    dif = idx1;
    idxCut1 = idxG1 - dif;
```

```
    MgaussWin1 = gaussWin1(idxCut1+1:idxCut1 + length(segment1));
else
    dif = length(segment1) - idx1;
    idxCut1 = idxG1 + dif;
    MgaussWin1 = gaussWin1(1:idxCut1);
end

if((idxG2 - idx2)>0)
    dif = idx2;
    idxCut2 = idxG2 - dif;
    MgaussWin2 = gaussWin2(idxCut2+1:idxCut2 + length(segment2));
else
    dif = length(segment2) - idx2;
    idxCut2 = idxG2 + dif;
    MgaussWin2 = gaussWin2(1:idxCut2);
end

outseg1 = segment1 .* MgaussWin1;
outseg2 = segment2 .* MgaussWin2;

% figure; subplot(311); bar(segment1); subplot(312); bar(MgaussWin1);
subplot(313); bar(outseg1);
% figure; subplot(311); bar(segment2); subplot(312); bar(MgaussWin2);
subplot(313); bar(outseg2);

HgramRef = [outseg1;outseg2];
% figure; bar(HgramRef);

im_enhanced = HistogramMatching(I,HgramRef);
% figure; imshow(im_enhanced);

end
```

REFERENCES

Antonacopoulos, A., Clausner, C. and Pletschacher, S. (2011). *Historical Document Layout Analysis Competition*. [online] Primaresearch.org. Available at: https://www.primaresearch.org/www/assets/papers/ICDAR2011_Antonacopoulos_HDLAC.pdf [Accessed 21 Jun. 2019].

Antonacopoulos, A. (2010). *Large - Scale Digitisation and Recognition of Historical Documents:*
Challenges and Opportunities for Image Processing and Analysis. [online] Primaresearch.org.
Available at: https://www.primaresearch.org/www/assets/papers/SSBA2010_Antonacopoulos_KeynoteD igitisationAnalysis.pdf [Accessed 21 May 2019].

Ahmed, S., Ghosh, K. K., Bera, S. K., Schwenker, F., & Sarkar, R. (2020). Gray level image contrast enhancement using barnacles mating optimizer. *IEEE Access, 8*, 169196-169214.

Antonacopoulos, A., Bridson, D., Papadopoulos, C., & Pletschacher, S. (2009). A realistic dataset for performance evaluation of document layout analysis. 2009 10th International Conference on Document Analysis and Recognition,

[Record #76 is using a reference type undefined in this output style.]

Antonacopoulos, A., & Karatzas, D. (2005). Semantics-based content extraction in typewritten historical documents. Eighth International Conference on Document Analysis and Recognition (ICDAR'05),

Aurich, V., & Weule, J. (1995). Non-linear gaussian filters performing edge preserving diffusion. In *Mustererkennung 1995* (pp. 538-545). Springer.

Bankman, I. (2008). *Handbook of medical image processing and analysis*. Elsevier.

Bannigidad, P., & Gudada, C. (2016). Restoration of degraded historical Kannada handwritten document images using image enhancement techniques. International Conference on Soft Computing and Pattern Recognition,

Bernsen, J. (1986). Dynamic thresholding of gray-level images. Proc. Eighth Int'l conf. Pattern Recognition, Paris, 1986,

Bonny, M. Z., & Uddin, M. S. (2019). Degraded Document Enhancement through Binarization Techniques. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI),

Boonserm, N. (2015). *The development of noise reduction model for computed radiography system* School of Information Technology Institute of Social Technology. Suranaree …].

Boudraa, O., Hidouci, W. K., & Michelucci, D. (2019). Degraded Historical Documents Images Binarization Using a Combination of Enhanced Techniques. *arXiv preprint arXiv:1901.09425*.

Bradley, D., & Roth, G. (2007). Adaptive thresholding using the integral image. *Journal of graphics tools*, *12*(2), 13-21.

Calvo-Zaragoza, J., & Gallego, A.-J. (2019). A selectional auto-encoder approach for document image binarization. *Pattern Recognition, 86*, 37-47.

Cheng, H.-D., & Zhang, Y. (2012). Detecting of contrast over-enhancement. 2012 19th IEEE international conference on image processing,

Clausner, C., Antonacopoulos, A., & Pletschacher, S. (2020). Efficient and effective OCR engine training. *International Journal on Document Analysis and Recognition (IJDAR), 23*(1), 73-88.

Clausner, C., Hayes, J., Antonacopoulos, A., & Pletschacher, S. (2017). Creating a complete workflow for digitising historical census documents: considerations and evaluation. Proceedings of the 4th International Workshop on Historical Document Imaging and Processing,

Clausner, C., Papadopoulos, C., Pletschacher, S., & Antonacopoulos, A. (2015). The ENP image and ground truth dataset of historical newspapers. 2015 13th International Conference on Document Analysis and Recognition (ICDAR),

Clausner, C., Pletschacher, S., & Antonacopoulos, A. (2016). Quality prediction system for large-scale digitisation workflows. 2016 12th IAPR Workshop on Document Analysis Systems (DAS),

Correal, R., Pajares, G., & Ruz, J. J. (2014). Automatic expert system for 3D terrain reconstruction based on stereo vision and histogram matching. *Expert Systems with Applications, 41*(4), 2043-2051.

Erkan, U., Gökrem, L., & Enginoğlu, S. (2018). Different applied median filter in salt and pepper noise. *Computers & Electrical Engineering, 70*, 789-798.

Eyupoglu, C. (2017). Implementation of Bernsen's locally adaptive binarization method for gray scale images. *The Online Journal of Science and Technology, 7*(2), 68-72.

Gatos, B., Pratikakis, I., & Perantonis, S. J. (2004). An adaptive binarization technique for low quality historical documents. International Workshop on Document Analysis Systems,

Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. *Pattern Recognition, 39*(3), 317-327.

Gavaskar, R. G., & Chaudhury, K. N. (2018). Fast adaptive bilateral filtering. *IEEE Transactions on Image Processing, 28*(2), 779-790.

Ghosh, M., Bera, S. K., Guha, R., & Sarkar, R. (2019). Contrast enhancement of degraded document image using partitioning based genetic algorithm. International conference on emerging technologies for sustainable development (ICETSD'19),

Guha, R., Alam, I., Bera, S. K., Kumar, N., & Sarkar, R. (2022). Enhancement of image contrast using Selfish Herd Optimizer. *Multimedia Tools and Applications, 81*(1), 637-657.

Gupta, A., Kumar, S., Gupta, R., Chaudhury, S., & Joshi, S. (2007). Enhancement of old manuscript images. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007),

Haldar, R., & Mukhopadhyay, D. (2011). Levenshtein distance technique in dictionary lookup methods: An improved approach. *arXiv preprint arXiv:1101.1232*.

Hambal, A. M., Pei, Z., & Ishabailu, F. L. (2017). Image noise reduction and filtering techniques. *International Journal of Science and Research (IJSR), 6*(3), 2033-2038.

Han, Z., Su, B., Li, Y.-g., Ma, Y.-f., Wang, W.-d., & Chen, G.-q. (2019). An enhanced image binarization method incorporating with Monte-Carlo simulation. *Journal of Central South University, 26*(6), 1661-1671.

Hansen, E. W. (2014). *Fourier transforms: principles and applications*. John Wiley & Sons.

Hao, S., Han, X., Guo, Y., Xu, X., & Wang, M. (2020). Low-light image enhancement with semi-decoupled decomposition. *IEEE Transactions on Multimedia*.

He, S., & Schomaker, L. (2019). DeepOtsu: Document enhancement and binarization using iterative deep learning. *Pattern Recognition, 91*, 379-390.

Hong, L., Wan, Y., & Jain, A. (1998). Fingerprint image enhancement: algorithm and performance evaluation. *IEEE transactions on pattern analysis and machine intelligence, 20*(8), 777-789.

Hore, A., & Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. 2010 20th international conference on pattern recognition,

Hossain, M. F., Alsharif, M. R., & Yamashita, K. (2010). Medical image enhancement based on nonlinear technique and logarithmic transform coefficient histogram matching. IEEE/ICME International Conference on Complex Medical Engineering,

Howe, N. R. (2011). A laplacian energy for document binarization. 2011 International Conference on Document Analysis and Recognition,

Howe, N. R. (2013). Document binarization with automatic parameter tuning. *International Journal on Document Analysis and Recognition (IJDAR), 16*(3), 247-258.

Jajware, R. R., & Agnihotri, R. B. (2020). Image Enhancement of Historical Image Using Image Enhancement Technique. In *Innovations in Computer Science and Engineering* (pp. 233-239). Springer.

Jemni, S. K., Souibgui, M. A., Kessentini, Y., & Fornés, A. (2022). Enhance to read better: a multi-task adversarial network for handwritten document image enhancement. *Pattern Recognition, 123*, 108370.

Kang, S., Iwana, B. K., & Uchida, S. (2019). Cascading Modular U-Nets for Document Image Binarization. 2019 International Conference on Document Analysis and Recognition (ICDAR),

Karpinski, R., Lohani, D., & Belaid, A. (2018). Metrics for Complete Evaluation of OCR Performance.

Katsigiannis, S., Scovell, J., Ramzan, N., Janowski, L., Corriveau, P., Saad, M. A., & Van Wallendael, G. (2018). Interpreting MOS scores, when can users see a difference? Understanding user experience differences for photo quality. *Quality and User Experience, 3*(1), 6.

Kaur, M., Kaur, J., & Kaur, J. (2011). Survey of contrast enhancement techniques based on histogram equalization. *International Journal of Advanced Computer Science and Applications, 2*(7).

Kavallieratou, E. (2005). A binarization algorithm specialized on document images and photos. Eighth International Conference on Document Analysis and Recognition (ICDAR'05),

Kavallieratou, E., & Antonopoulou, H. (2005). Cleaning and enhancing historical document images. International Conference on Advanced Concepts for Intelligent Vision Systems,

Kim, T. K., Paik, J. K., & Kang, B. S. (1998). Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering. *IEEE Transactions on Consumer Electronics, 44*(1), 82-87.

Kim, Y.-T. (1997). Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Transactions on Consumer Electronics, 43*(1), 1-8.

Kluzner, V., Tzadok, A., Shimony, Y., Walach, E., & Antonacopoulos, A. (2009). Word-based adaptive OCR for historical books. 2009 10th International Conference on Document Analysis and Recognition,

Krbcova, Z., & Kukal, J. (2017). Relationship between entropy and SNR changes in image enhancement. *EURASIP Journal on Image and Video Processing, 2017*(1), 1-8.

Li, J., Mei, Z., & Zhang, T. (2020). A method for document image enhancement to improve template-based classification. Proceedings of the 2020 4th High Performance Computing and Cluster Technologies Conference & 2020 3rd International Conference on Big Data and Artificial Intelligence, Likforman-Sulem, L., Darbon, J., & Smith, E. H. B. (2011). Enhancement of historical printed document images by combining total variation regularization and non-local means filtering. *Image and vision computing, 29*(5), 351-363.

Ling, Z., Liang, Y., Wang, Y., Shen, H., & Lu, X. (2015). Adaptive extended piecewise histogram equalisation for dark image enhancement. *IET Image Processing, 9*(11), 1012-1019.

Liu, Y. (2013). Noise reduction by vector median filtering. *Geophysics, 78*(3), V79-V87.

Lu, D., Huang, X., & Sui, L. (2018). Binarization of degraded document images based on contrast enhancement. *International Journal on Document Analysis and Recognition (IJDAR), 21*(1-2), 123-135.

Mafi, M., Martin, H., Cabrerizo, M., Andrian, J., Barreto, A., & Adjouadi, M. (2019). A comprehensive survey on impulse and Gaussian denoising filters for digital images. *Signal Processing, 157*, 236-260.

Moghaddam, R. F., & Cheriet, M. (2009). A variational approach to degraded document enhancement. *IEEE transactions on pattern analysis and machine intelligence, 32*(8), 1347-1361.

Motl, J. (2013). Niblack local thresholding. *MATLAB Cent. File Exch*.

Najafi, M. H., & Salehi, M. E. (2015). A fast fault-tolerant architecture for sauvola local image thresholding algorithm using stochastic computing. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 24*(2), 808-812.

Navarro, G. (2001). A guided tour to approximate string matching. *ACM computing surveys (CSUR), 33*(1), 31-88.

Nguyen, T.-A., Song, W.-S., & Hong, M.-C. (2010). Spatially adaptive denoising algorithm for a single image corrupted by Gaussian noise. *IEEE Transactions on Consumer Electronics, 56*(3), 1610-1615.

Niblack, W. (2003). An introduction to digital image processing. 1986. G. Leedham, C. Yan, K. Takru, J. Tan and L. Mian. Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images. In Proceedings of the Seventh International Conference on Document Analysis and Recognition,

Ooi, C. H., & Isa, N. A. M. (2010). Adaptive contrast enhancement methods with brightness preserving. *IEEE Transactions on Consumer Electronics, 56*(4), 2543-2551.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics, 9*(1), 62-66.

Papadopoulos, C., Pletschacher, S., Clausner, C., & Antonacopoulos, A. (2013). The IMPACT dataset of historical document images. Proceedings of the 2Nd international workshop on historical document imaging and processing, Peng, X., Cao, H., & Natarajan, P. (2017). Using convolutional encoder-decoder for document image binarization. 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR),

Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE transactions on pattern analysis and machine intelligence, 12*(7), 629-639.

Pham, B. T., Qi, C., Ho, L. S., Nguyen-Thoi, T., Al-Ansari, N., Nguyen, M. D., Nguyen, H. D., Ly, H.-B., Le, H. V., & Prakash, I. (2020). A novel hybrid soft computing model using random forest and particle swarm optimization for estimation of undrained shear strength of soil. *Sustainability, 12*(6), 2218.

Pratikakis, I., Gatos, B., & Ntirogiannis, K. (2013). ICDAR 2013 document image binarization contest (DIBCO 2013). 2013 12th International Conference on Document Analysis and Recognition,

Pratikakis, I., Zagoris, K., Barlas, G., & Gatos, B. (2017). ICDAR2017 competition on document image binarization (DIBCO 2017). 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR),

Primaresearch. (2021). *Pattern Recognition & Image Analysis Research Lab*. Retrieved April,5,2021 from https://www.primaresearch.org/

Qiu, Y., Gan, Z., Fan, Y., & Zhu, X. (2011). An adaptive image denoising method for mixture gaussian noise. 2011 International Conference on Wireless Communications and Signal Processing (WCSP),

Rani, R., Verma, A., Verma, S. K., & Bansal, U. (2014). AHybrid APPROACH OF STRETCHING AND FILTERING FOR ENHANCING UNDERWATER GRAY SCALE IMAGE. *Red, 685*, 5m.

Rao, B. S. (2020). Dynamic histogram equalization for contrast enhancement for digital images. *Applied Soft Computing, 89*, 106114.

Reza, A. M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology, 38*(1), 35-44.

Santhi, K., & Banu, R. W. (2015). Adaptive contrast enhancement using modified histogram equalization. *Optik-International Journal for Light and Electron Optics, 126*(19), 1809-1814.

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

Sengee, N., & Choi, H. K. (2008). Brightness preserving weight clustering histogram equalization. *IEEE Transactions on Consumer Electronics, 54*(3), 1329-1337.

Shahkolaei, A., Beghdadi, A., & Cheriet, M. (2019). Blind quality assessment metric and degradation classification for degraded document images. *Signal Processing: Image Communication, 76*, 11-21.

Shahkolaei, A., Nafchi, H. Z., Al-Maadeed, S., & Cheriet, M. (2018). Subjective and objective quality assessment of degraded document images. *Journal of Cultural Heritage, 30*, 199-209.

Shi, Z., & Govindaraju, V. (2004). Historical document image enhancement using background light intensity normalization. Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.,

Smith, S. M., & Brady, J. M. (1997). SUSAN—a new approach to low level image processing. *International journal of computer vision, 23*(1), 45-78.

Su, H., & Jung, C. (2017). Low light image enhancement based on two-step noise suppression. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),

Sulaiman, A., Omar, K., & Nasrudin, M. F. (2019). Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. *Journal of imaging, 5*(4), 48.

Sun, B., Li, S., Zhang, X.-P., & Sun, J. (2016). Blind bleed-through removal for scanned historical document image with conditional random fields. *IEEE Transactions on Image Processing, 25*(12), 5702-5712.

Tensmeyer, C., & Martinez, T. (2017). Document image binarization with fully convolutional neural networks. 2017 14th IAPR international conference on document analysis and recognition (ICDAR),

Tomasi, C., & Manduchi, R. (1998). Bilateral filtering for gray and color images. Sixth international conference on computer vision (IEEE Cat. No. 98CH36271),

Van Kempen, G., Van Vliet, L., Verveer, P., & Van Der Voort, H. (1997). A quantitative comparison of image restoration methods for confocal microscopy. *Journal of Microscopy, 185*(3), 354-365.

Vo, Q. N., Kim, S. H., Yang, H. J., & Lee, G. (2018). Binarization of degraded document images based on hierarchical deep supervised network. *Pattern Recognition, 74*, 568-586.

Wang, M., Zhou, S., & Yan, W. (2018). Blurred image restoration using knife-edge function and optimal window Wiener filtering. *PloS one, 13*(1), e0191833.

Wang, Z., Bovik, A. C., & Simoncelli, E. P. (2005). Structural approaches to image quality assessment. *Handbook of Image and Video Processing, 7*, 18.

Wellner, P. D. (1993). Adaptive thresholding for the DigitalDesk. *Xerox, EPC1993-110*, 1-19.

Westphal, F., Lavesson, N., & Grahn, H. (2018). Document image binarization using recurrent neural networks. 2018 13th IAPR International Workshop on Document Analysis Systems (DAS),

Winiarti, S., Ismi, D. P., & Prahara, A. (2017). Image enhancement using piecewise linear contrast stretch methods based on unsharp masking algorithms for leather image processing. 2017 3rd International conference on science in information technology (ICSITech),

Xiong, W., Zhou, L., Yue, L., Li, L., & Wang, S. (2021). An enhanced binarization framework for degraded historical document images. *EURASIP Journal on Image and Video Processing*, *2021*(1), 13.

Xu, L., Yan, Q., Xia, Y., & Jia, J. (2012). Structure extraction from texture via relative total variation. *ACM transactions on graphics (TOG)*, *31*(6), 1-10.

Yagoubi, M. R., Serir, A., & Beghdadi, A. (2015). A new automatic framework for document image enhancement process based on anisotropic diffusion. 2015 13th International Conference on Document Analysis and Recognition (ICDAR),

Yang, F., & Wu, J. (2010). An improved image contrast enhancement in multiple-peak images based on histogram equalization. 2010 International Conference on Computer Design and Applications,

Ye, P., & Doermann, D. (2013). Document image quality assessment: A brief survey. 2013 12th International Conference on Document Analysis and Recognition,

Zagoris, K., & Pratikakis, I. (2017). Bio-inspired modeling for the enhancement of historical handwritten documents. 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR),

Zhu, Y., & Huang, C. (2012). An improved median filtering algorithm for image noise reduction. *Physics Procedia*, *25*, 609-616.

Zhu, Y., Sun, J., & Naoi, S. (2011). Recognizing natural scene characters by convolutional neural network and bimodal image enhancement. International Workshop on Camera-Based Document Analysis and Recognition,