22 - Final Project Report

Prashanth N

ai19btech11010@iith.ac.in

Praveen M

ai19btech11018@iith.ac.in

Jayanth Bhaskar M

ai19btech11008@iith.ac.in

Abstract

*This is a thorough visit back to problem of **Document Enhancement** and analyzing various methods that were proposed to address it ranging from simple Binarization [2] to GANs [11] for cleaning up the image affected by various kinds of degradation (blur, leak, shadow, etc.) specifically for producing more robust results incrementally from the already matured OCR technology, to eliminate the confusion of choosing from myriad of techniques for interesting observations which could potentially bud a new technique which could possibly be flexible and adaptable to various kinds of issues and artifacts, more accurate and faster load and inference times for the Document Enhancement problem*

1. Introduction

There is huge interest in people to make all kinds of knowledge accessible online but there are many challenges which hinder that process one of the main one being documenting the written texts which are often occluded or comes with damaged versions, maybe coffee stains, overtime degradation and many other natural scenarios,

So if we want to have our knowledge repository filled with newer ways of obtaining documents like scanned images from smart-phones problem persists there too in forms such as shadow 2, blur 5, bleed-through 6, different lighting conditions(occlusions and overexposure) and distorted images and also other well-known problems of rolling-shutter. Even after the process of obtaining the image digitally there is no guarantee that the image is free from damage there might be some watermark, annotations and stamps which might hinder in obtaining the complete data through these images, So there is a need to adopt to various kinds of algorithms based on different use cases

1. Mobile: Document scanning - Quick inference even though a little accuracy is compromised

2. Knowledge base: Preserving for online libraries - The prediction should be more accurate even though time intense

We can clearly see the need of the hour for classify-

Figure 1. Various kinds of degradations currently under analysis

ing existing methodologies based on the results and performances of the methods employed to tackle various degradations

One such idea is to analyze various existing solutions and contrasting various algorithms with and without the preprocessing step of image visibility detection which costs $O(nlogn)$ for n pixels [5] and proposing which model combination suits best for a specific use case

2. Previous works

Although the problem has been studied in various magnifications of scientific lenses some narrowed to specific artifacts like shadows. For example, Bako et al. [1] which showed very appreciable results owing to their very simple method of eliminating shadows by scaling down by a factor of α_i which is the ratio of local and global background. They made use of the fact that the background stays the same color

In contrast there are very novel latest methods which take

even the information given by **text** present in the document to help its enhancement. For example, Khamekhem Jemni et al. [4] used the UNet and encoded Convolutional Recurrent Neural Nets(CRNNs) embeddings into Gated Recurrent Units(GRUs) (for text recognition) and made the unavailable degraded datasets of KHATT and IAM (with strokes and other artifacts added) All these approach the problem of Document Enhancement in a fantastic manner but for a amateur interested in enhancing certain documents with some set of constraints they bring in the paradox of choice and tremendous load of research

So, our work mainly is to reduce the burden and let the amateur steer towards a method well-suited to his use-case

3. Problem Statement

The problem boils down to classifying various combinations of models aiming to provide ready-made models suited for different needs of the document enhancer on the fly

4. Literature Review

[11] studies the same problem and giving a very heavy solution of GAN called DE-GAN [11] which would train up a Generator which tries to fool the Discriminator by producing images as close to the original as the Generator learns to remove the damage in the conditioned image while preserving the required text This paper mainly address two key issues in depth

1. Image blurring
2. Watermark removal

[5] mainly identifies a novel way of representing images to remove the occlusions that are formed due to intermediary intensities in the image and then binarizing with a new technique called **Lowlights Map** which provides much better results compared to standard one step separation of foreground with background

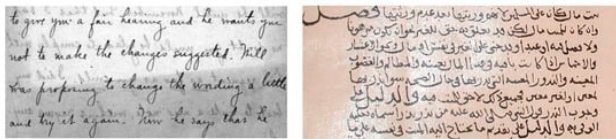## 5. Handling various degradations

5.1. Uneven Illumination

Figure 2. Uneven illumination in historical handwritten document from DIBCO (**left**) and Arabic Databases (**right**)

Uneven illumination occurs when texts(or)images are scanned poorly, when photos taken by the digital camera. Bako et al. [1] which showed very appreciable results owing to their very simple method of eliminating shadows by

scaling down by a factor of α_i which is the ratio of local and global background. They made use of the fact that the background stays the same color [5] also made good efforts and results are very similar to Human Visual System through SSIM index when compared with former
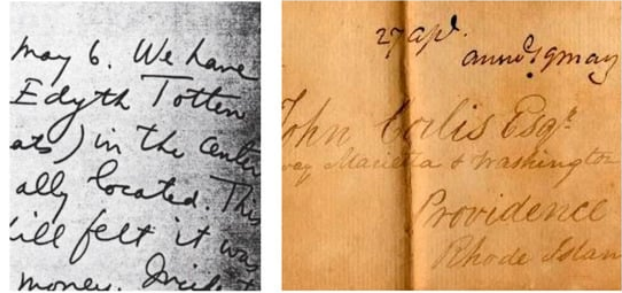
5.2. Contrast Variation

Figure 3. Degraded document images showing variation of contrast: crease (**right**)

Contrast is a variation in brightness.Factors that causes contrast variations in documents are noisy environment, sunlight, illumination and occlusion. The most common enhancement method to correct the contrast variation is **Histogram Equalization**(HE).The adaptive HE is an improvement in HE-based contrast enhancement.It treats the image globally.

5.3. Show-through

Figure 4. Show-through/smear effects depicted.

It is a problem appears when low-quality ink impressions available in one side of the document start to pass through and appear on the other side,and it's difficult read that document. In this paper [8] A clear background can be obtained if the show-through on this image were to be removed. [10] So here well developed system created for degraded old documents was quite capable of dealing with degradations occurring as a result of noise, low contrast, non-uniform illumination, or shadows. [8] they also came up with a novel approach to improve ancient historical documents with a two-way approach.

5.4. Blur

Figure 5. Examples of images of degraded documents with blurring effect.

[6,9,12] In these paper they mainly discussed about the various de-blurring methods,and all of them are have proven to be effective on natural images, but completely the opposite on degraded document images that contain text, because of the ringing artefacts that emerge due to the de-blurring process. [13]In the paper they devised a way to remove the blur that is present in degraded documents containing textual information, and further built it into OCR to improve accuracy. However, the later results were not good enough as the ringing effect degraded the image quality. But it is still difficult to restore the blurry image of the document. RNN-based methods have also been taking traction which is fostering further research on their utilities.

5.5. Bleed Through

Figure 6. Examples of ink-bleed degradation from DIBCO database (**left**) from 2016 and (**right**) from 2017 respectively.

This is caused when an ink oozes through one side and spreads across the other and ruining the text. This issue can be solved by either preparing an image of certain degradation quality based or the corresponding ground truth or by forming an image as ground truth by knowing the original degraded images [7]. Why can't we just use thresholding methods to solve this? In this paper, the authors have shown by comparing the three global thresholding techniques(Otsu's, Kapur's entropy, and Solihin's quadratic integral ratio (QIR)) which may be used as the first step of a multi-stage thresholding algorithm for degraded document images. It is found that Otsu's and Kapur's algorithms are ineffective for complex documents because they over-threshold the image, resulting in loss of most of the useful information.The QIR algorithm is more accurate in separating the foreground and background in these images, leaving a range of undecided, fuzzy, pixels for later processing in a subsequent stage.

6. Results

We summarize the methods and techniques used to solved various degradation problems based on their efficacy through various metrics like (Peak Signal-to-Noise Ratio (PSNR), F-measure, Pseudo-F-Measure, Distance Reciprocal Distortion (DRD), Multi-Classification Penalty Metric (MPM) and on dataset DIBCO [3] mentioned in their respective papers Here the Bako et.al [1] takes very less time overall to learn a local scale that could be applied to local regions to obtain the appropriate scale for that region to enhance the shadow Below are the tables for delineating the degradation correction methods and methods that have potential to become mobile (low latency)

Table 1. Low latency

Degradation	Good Time Method
Uneven Illumination	[1]
Contrast Variation	–
Show Through	[8]
Blur	[13]
Bleed Through	–

Table 2. Enhancement methods for various degradation issues

Degradation	Good Correction Method
Uneven Illumination	Visibility Detection [5]
Contrast Variation	Histogram Equalization(HE)
Blur	RNN Based methods
Bleed Through	DE-GAN [11]

7. Discussion

In this paper we have studied and analysed some of the methods given by authors from various backgrounds in solving the degradation issues of documents we have come across so far by crawling the references and picking the good methods and creating a helpful table1 for a neophyte who quickly wants to solve one or several of the degradations, although most of the studied problems are old and mostly based on binarization often taking in some heuristics but with the recent advances in GANs and other Neural network based methods which can be used in unifying the solving ability of these discrete methods for various kinds of degradations with enhanced speed and accuracy by making lesser assumptions/ implicit assumptions Also a classifier for classifying the type of degradation and automatically assigning the best method currently in field could be a viable add-on so that and end-to-end automation for enhancing the documents could happen

8. Contributions

AI19BTECH11010(Prashanth Reddy Nimmakayala) - Code testing, problem statement, Paper collection, idea, outline

AI19BTECH11018 - Handling Degradations,results

AI19BTECH11008 - Reading Abstracts

References

[1] Steve Bako, Soheil Darabi, Eli Shechtman, Jue Wang, Kalyan Sunkavalli, and Pradeep Sen. Removing shadows from images of documents. *Asian Conference on Computer Vision (ACCV 2016)*, 2016. 1, 2, 3

[2] Sheng He and Lambert Schomaker. Deepotsu: Document enhancement and binarization using iterative deep learning. *CoRR*, abs/1901.06081, 2019. 1

[3] Xenofon Karagiannis Ioannis Pratikakis, Konstantinos Zagoris. Icdar 2019. https://vc.ee.duth.gr/dibco2019/, Sep 2019. 3

[4] Sana Khamekhem Jemni, Mohamed Ali Souibgui, Yousri Kessentini, and Alicia Fornés. Enhance to read better: A multi-task adversarial network for handwritten document image enhancement. *Pattern Recognition*, 123:108370, Mar 2022. 2

[5] Netanel Kligler, Sagi Katz, and Ayellet Tal. Document enhancement using visibility detection. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2374–2382, 2018. 1, 2, 3

[6] RL Lagendijk and J Biemond. Basic methods for image restoration and identification. *Handbook of Image an Video Processing, AC Bovik and JD Gibson, eds*, 2000. 3

[7] G. Leedham, S. Varma, A. Patankar, and V. Govindaraju. Separating text and background in degraded document images - a comparison of global thresholding techniques for multi-stage thresholding. In *Proceedings Eighth International Workshop on Frontiers in Handwriting Recognition*, pages 244–249, 2002. 3

[8] Md. Iqbal Quraishi, Mallika De, Krishna Gopal Dhal, Saheb Mondal, and Goutam Das. A novel hybrid approach to restore historical degraded documents. In *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*, pages 185–189, 2013. 2, 3

[9] Filip Rooms, Aleksandra Pizurica, and Wilfried Philips. Estimating image blur in the wavelet domain. In *Ieee international conference on acoustics speech and signal processing*, volume 4, pages 4190–4190. IEEE; 1999, 2002. 3

[10] K. Shirai, Y. Endo, A. Kitadai, S. Inoue, N. Kurushima, H. Baba, A. Watanabe, and M. Nakagawa. Character shape restoration of binarized historical documents by smoothing via geodesic morphology. In *2013 12th International Conference on Document Analysis and Recognition*, pages 1285–1289, 2013. 2

[11] Mohamed Ali Souibgui and Yousri Kessentini. Degan: A conditional generative adversarial network for document enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 1–1, 2021. 1, 2, 3

[12] Paolo Vivirito, Sebastiano Battiato, Salvatore Curti, M. La Cascia, and Roberto Pirrone. Restoration of out-of-focus images based on circle of confusion estimate. In Andrew G. Tescher, editor, *Applications of Digital Image Processing XXV*, volume 4790, pages 408 – 416. International Society for Optics and Photonics, SPIE, 2002. 3

[13] Shiqian Wu, Weisi Lin, Lijun Jiang, Wei Xiong, Lihao Chen, and Sim Heng Ong. An objective out-of-focus blur measurement. In *2005 5th International Conference on Information Communications & Signal Processing*, pages 334–338. IEEE, 2005. 3