# PERFORMANCE OPTIMIZATION OF TEXT EXTRACTION TECHNIQUE FOR COMPLEX DEGRADED IMAGES

Abhishek Gupta*, Ramapati Mishra**, Ashutosh Kumar Singh***
*Research Scholar, **Professor, ***Associate Professor
Department of Electronics and Communication Engineering
Institute of Engineering and Technology, Dr. Ram Manohar Lohia Avadh University, Ayodhya, India

***Abstract-***

***In recent years, there has been a growing demand to preserve historical documents, books and convert them into digital format. Moreover, the fast development of data innovation and the speedy propagation of the Internet have also lead to the huge amount of image and video data. The texts present in the image and video help us in the analysis of those images and videos as well as used in indexing, archiving, and retrieval. An image can be easily affected by various noises like Gaussian noise, salt and pepper noise, speckle noise etc. To remove these different noises from images various image filtering algorithms such as Gaussian filter, mean filter, median filter etc. are used. The effect of various pre-processing techniques like Thresholding, Morphology, and Blurring processes to optimize the text extraction technique are analysed in this article. The results obtained from the experiment show that the pre-processing techniques certainly enhance the document's visual and structural quality.***

***Keywords- Complex degraded image, Thresholding, Morphology, Edge detection, OCR***

## I. INTRODUCTION

The advancement of the Internet led to the huge increment in images and videos database. Most of the images and videos consist of numerous texts information. Moreover, in recent times there has been an increase in demand to preserve historical documents, books and convert them into digital format. However, extracting texts from these natural images, scanned documents or videos is a tedious work. Images are usually distorted by noise during its acquisition, processing, transmission and reproduction. Restoring the original image after removing the noise is a fundamental objective of image processing [4]. Despite numerous algorithms proposed in the past for text extraction from images, there still needs some work to be done to improve the text extraction from images and videos [10]. The main challenges in text extraction from complex degraded images are orientation of texts, font size, diversity of

background, low quality, difference in colour of texts and interference of noise etc. So, pre-processing techniques are the most key thing to enhance the quality of text extraction [13].

In this work, we look at how fundamental pre-processing approaches affect image quality and as a result, how they affect text extraction.

### *A. Noise*

Noise is a random variation in digital images that may cause different intensity value of pixels instead of true pixel values. There are numerous types of noises in image processing that affects the image in various manners. Some of the most common types of noises are explained as follows:

#### *1) Gaussian Noise*

This noise primarily occurs in images during its acquisition. It is additive in nature. These noise models are popular due to their adaptability in both the spatial and frequency domains. The PDF of a Gaussian random variable may be calculated using the equation below.

$$p(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}}$$

Where, z = Intensity

$\bar{z}$ = Mean value of z

σ = Standard deviation

σ2 = Variance of z

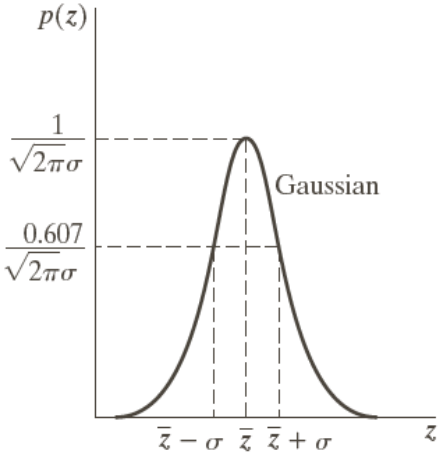



**Figure 1.** PDF of Gaussian noise

#### *2) Impulse Noise*

It is also known as salt and pepper noise. As a result of this noise, black dots appear in bright areas and white dots emerge in dark areas. This noise primarily occurs during transmission or conversion process. Salt noise has a pixel value of 255 in an 8-bit picture, while pepper noise has a pixel value of 0. The following is the PDF for Impulse noise:

$$p(z) = \begin{cases} P_a & for \quad z = a \\ P_b & for \quad z = b \\ 0 & otherwise \end{cases}$$

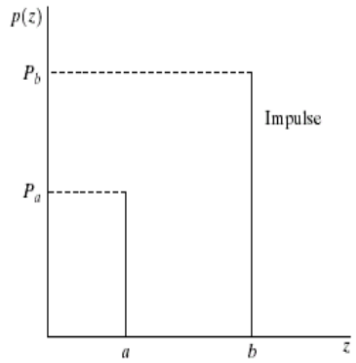



**Figure 2.** PDF of impulse noise

*B. Pre-processing Methods*

The image pre-processing technique may be used to a wide range of advanced image processing tasks. The nature of these algorithms and how they may be utilised to accelerate the development of image processing is explained in [10]. There are either generic or specific pre-processing processes utilised in various text extraction approaches.

*1) Blurring*

Blurring removes high-frequency components in the image, such as noise [1]. There are various filters used for blurring but only two of them i.e. Gaussian filter and median filter are examined in this work.

*i) Gaussian Filter*

The Gaussian filter blurs images using a Gaussian kernel. It has the ability to eliminate the Gaussian noise. The expression of a Gaussian function is defined as follows:

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

Where, $\sigma$ = standard deviation

x and y = distance along the horizontal and vertical axis respectively

*ii) Median Filter*

The median filter replaces the centre element with the median value after taking the median of all pixels within the given region. It minimises noise while maintaining the edges. This approach eliminates the salt and pepper noise. The median filter can be defined by the expression as below

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\text{median}} \{g(s,t)\}$$

*2) Thresholding*

Thresholding methods try to make a grayscale picture into binary form using pixel density as a criterion. Thresholding is a basic method of doing picture segmentation into two regions i.e. foreground and background. The output is determined by intensity threshold [12]. If the pixel intensity value exceeds the threshold, the pixel is replaced with a white pixel and if it is less than the threshold value, it is replaced by a black pixel. The most common and widely used method for thresholding is discussed as below.

*i) Global Thresholding*

The global thresholding technique helps us to classify the image pixels and background pixels of an object. The global binarization is a method that uses a single threshold value for the entire document. While in the local binarization method instead of a single threshold value, different threshold values are selected for every pixel in the entire image [6].

Otsu's method is a technique for image binarization that uses adaptive thresholding. It selects an optimal threshold value from a possible range of threshold values i.e. from 0 to 255. This method is based on global thresholding and it is used to carry out comprehensive image thresholding [9]. It is also used to convert a grayscale image to a binary image.

The method considers two types of pixels that an image consists of i.e. foreground and background pixels, it then evaluates the optimum threshold value that distinguishes the two types of pixels in order to minimize their mixed spread as well as the intra-class variance or in an equivalent manner their inter-class variance is maximised [8].

*3) Morphological Operations*

Morphology refers to a broad range of image processing procedures that alter pictures depending on their forms or shapes [5]. It is one of the data processing methods that may be used in picture processing. It has a wide range of applications, including texture analysis, noise reduction and boundary extraction [15].

*i) Dilation*

Dilation is one of the fundamental operations in morphology. It is most commonly used on binary images, although it may also be used on grayscale images. The objects enlarge as a result of dilation. This procedure has the effect of gradually increasing the borders of foreground pixels, resulting in larger regions and smaller gaps in that region [3].

*ii) Erosion*

Erosion causes objects to decrease in size. Erosion essentially erodes the foreground's borders, causing parts of those pixels to diminish in size and gaps in those regions to become wider [2]. Erosion is the reverse of dilatation. While, dilatation widens limits and fills holes, erosion narrows boundaries and widens holes [7]. If the structuring element does not entirely overlap ON-valued pixels, it sets an ON pixel to OFF [11].

*iii) Opening and Closing*

More complicated sequences can be created by combining the two primary procedures of dilation and erosion. The most helpful of them for morphological filtering are opening and closing. An opening operation is described as erosion followed by dilation, both of which are performed with the same structural element [14].

## II. METHODOLOGY

In our proposed method, the input complex degraded image is first filtered to remove noise by applying a suitable filter and if the given image is in colour, it is transformed to grayscale. Later, the Otsu's thresholding technique is applied to the gray scale image and horizontal as well as vertical gradient are calculated and on the basis of these gradients the edges are detected. The next stage involves the morphological operations in which the dilation operation is performed first to enlarge the boundaries of the foreground pixels and then the erosion operation is used on the masked image to narrow the boundaries and finally the closing operation is performed to improve the edges. After that, a post-processing step is involved to further improve the image quality by reducing uneven illumination. Finally, the text is recognized with the help of OCR process. The flowchart of the proposed methodology is depicted in Figure 3.

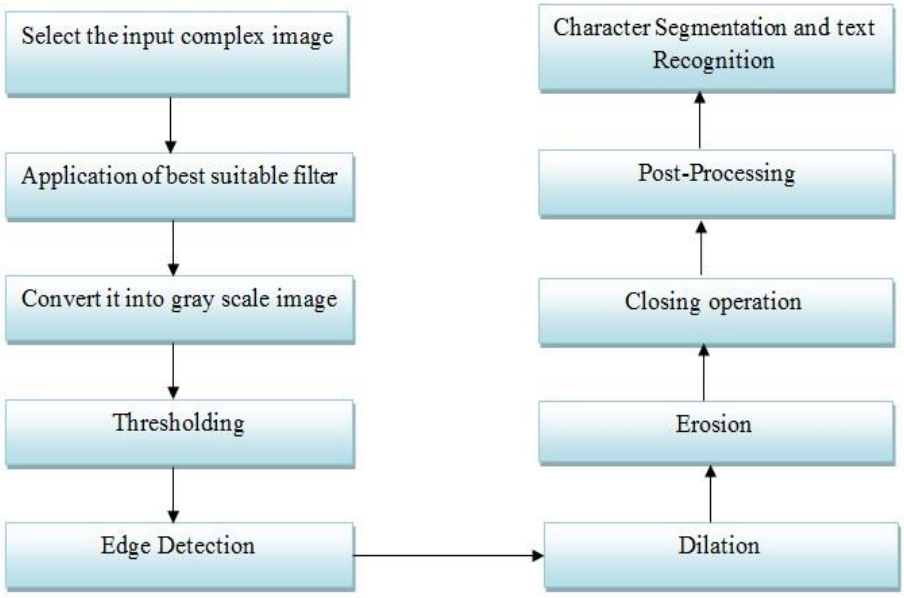



**Figure 3.** Flowchart of the proposed methodology

## III. EXPERIMENTAL ANALYSIS

To test our text extraction technique from complex degraded images, we have used the random images from various Internet sources that contain a wide range of datasets having different background illumination, colour, image quality, font sizes, orientation etc. The

proposed algorithm is implemented at the MATLAB R2018a platform. All the experiments are executed on a standard computer (Intel Core i7- 4770, 3.40 GHz CPU, 4 GB RAM and Windows 8.1 Pro, 64-bit OS). Some of the complex degraded images and the extracted texts are shown in Table 1 and Table 2.

**Table 1.** Results obtained after applying the proposed algorithm on various complex degraded images with the Gaussian noise

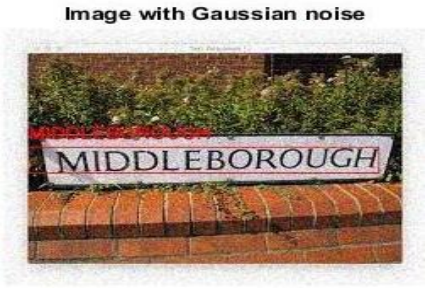
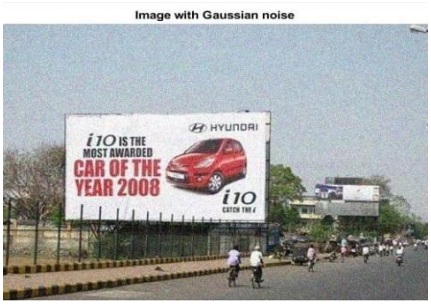
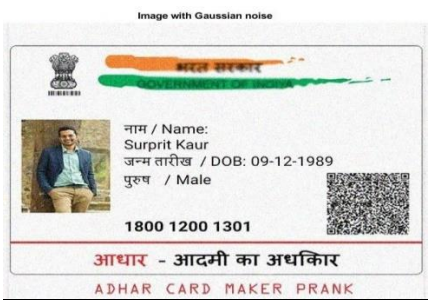
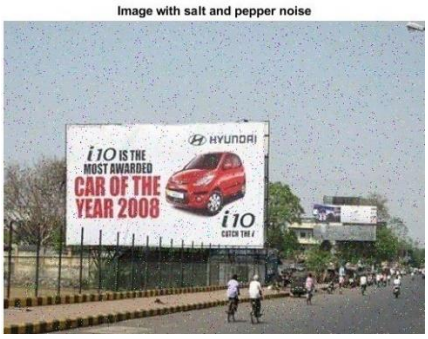
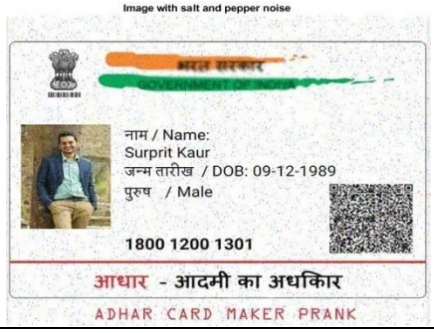
| Original Image | Image with the Gaussian noise | Extracted Text |
|---|---|---|
|   |   | MIDDLEBOROUGy |
|   |   | HYUNDAI<br>I10 IS THE<br>MOST AWARDED<br>CAR OF TH3<br>YEAR 2008<br>I 10 |
|   |   | INCOME TAX DEPARTMENT<br>GOVT. OF INDIA<br>Permanent Account Number Card<br>ABCDE1234F<br>APPLICANT NAME<br>/Fathers Name<br>APPLICANT'S FAlHER NAME<br>01/06/1995 Signature |
|   |   | Name:<br>Surprit Kaur<br>/ DOB: 09-12-1989<br>/ Male<br>18oo.12oo.13o1<br>ADHAR CARD MAKER FRANK |

**Table 2**.Results obtained after applying the proposed algorithm on various complex degraded images with the salt and pepper noise

| Original Image | Image with salt and pepper noise | Extracted Text |
|---|---|---|
|   |   | MIDDLEBOROUGJ |
|   |   | HYUNDAI<br>I10 IS THE<br>MOST AWARDED<br>CAR OF TH3<br>YEAR 2008 |
|   |   | INCOME TAX DEPARTMENT GOVT. OF INDIA<br>Permanent Account Number Card<br>ABCDE1234F<br>APPLICANT NAME<br>/Fathers Name<br>APPLICANTS FAlHER NAME<br>01/06/1995 Signature |
|   |   | Name:<br>Surprit Kaur<br>/DOB: 09-12-1989<br>/ Male<br>18oo.12oo.13o1<br>ADHAR CARD MAKER FRANK |

## IV. RESULTS AND DISCUSSION

The proposed method has its own advantages and disadvantages. It works properly on a wide range of datasets such as screenshots, invoices, documents and hoarding banners etc. However, it might not function very accurately on complex degraded images with poor image quality and uneven edges. A comparative analysis of accuracy for various images is illustrated in Tables 3, 4 and figures 4, 5.

**Table 3.** Accuracy for different images with Gaussian noise

| Image Id | Accuracy (%age) | |
|---|---|---|
| | Original Image (Without Noise) | Image with Gaussian noise |
| 1 | 100.00 | 92.3077 |
| 2 | 99.6154 | 99.2308 |
| 3 | 99.0991 | 98.7988 |
| 4 | 98.5714 | 97.8751 |

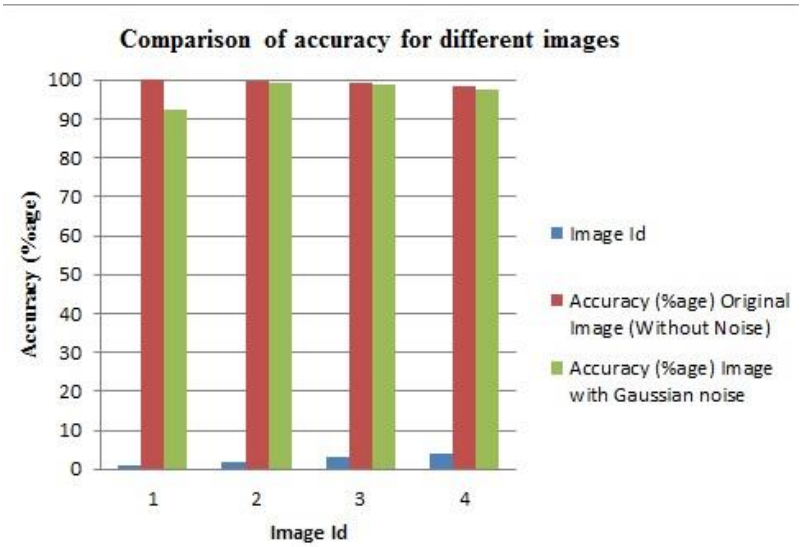



**Figure 4.** Comparison of accuracy for different images with Gaussian noise

**Table 4.** Accuracy for different images with salt and pepper noise

| Image Id | Accuracy (%age) | |
|---|---|---|
| | Original Image (Without Noise) | Image with salt and pepper noise |
| 1 | 100.00 | 92.3077 |
| 2 | 99.6154 | 99.2308 |
| 3 | 99.0991 | 96.3964 |
| 4 | 98.5714 | 97.5000 |

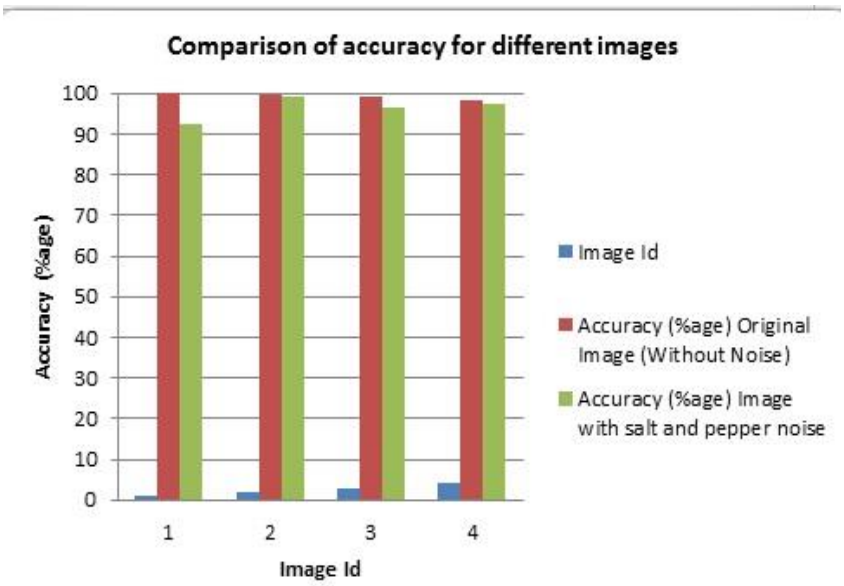



**Figure 5.** Comparison of accuracy for different images with salt and pepper noise

## V. CONCLUSION AND FUTURE WORK

A performance optimization technique of text extraction from complex degraded images is discussed in this paper. The suggested techniques' performance and efficacy are demonstrated by tests on a wide range of datasets. The accuracy of the proposed method is quite good and in some cases it is more than 99%. However, the proposed algorithm does not perform well to all types of complex degraded images, documents etc. so, there are two components to the future work:

i) Addressing the method's existing shortcomings such as variation of light, unconnected components and orientation of the texts present in the complex degraded images.

ii) We have used only two kinds of noise in the original image i.e. Gaussian noise and impulse or salt and pepper noise and to remove these noises only Gaussian filter and median filter are used therefore, future work can be extended to other types of noise and filtering algorithms as well.

iii) We have only trained and evaluated text detection from complex images for English; hence, multi-language text extraction technique can also be used in future research.

## REFERENCES

1. A. Koshy, N. B. M.J., S. A. and A. John, "Preprocessing Techniques for High Quality Text Extraction from Text Images," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), 2019, pp. 1-4, doi: 10.1109/ICIICT1.2019.8741488.
2. Chudasama, Diya & Patel, Tanvi & Joshi, Shubham & Prajapati, Ghanshyam. (2015). Image Segmentation using Morphological Operations. International Journal of Computer Applications. 117.

3. De Baets, Bernard, Kerre, E., & Gupta, M. (1995). The fundamentals of fuzzy mathematical morphology, part 1 : basic concepts. International Journal of General Systems, 23(2), 155–171.
4. Gupta, A., Mishra, R., & Singh, A. P. (2021). A Comprehensive Analysis of Various Text Detection and Extraction Techniques for Complex Degraded Images. Turkish Online Journal of Qualitative Inquiry (TOJQI), 12(7), 6186–6196.
5. H. Kawano, H. Orii, H. Maeda, N. Ikoma, "Text Extraction from Degraded Document Image Independent of Character Color Based on MAP-MRF Approach", IEEE, Jeju Island, Aug. 20-24 (2009) 165-168.
6. Jagroop Kaur D. and Mahajan R. 2014. A Review of Degraded Document Image Binarization Techniques. Changes. 3(5).
7. Jamil, N., Sembok, T. M. T., & Bakar, Z. A. (2008). Noise removal and enhancement of binary images using morphological operations. Proceedings - International Symposium on Information Technology 2008, ITSim, 3, 2–7. https://doi.org/10.1109/ITSIM.2008.4631954
8. Kurugollu, F., Sankur, B., & Harmanci, A. E. (2001). Color image segmentation using histogram multithresholding and fusion. Image and Vision Computing, 19(13), 915–928.
9. N. Otsu, "A Threshold Selection Method from Gray-level Histogram", IEEE Trans. Syst. Man Cybern. 8, 62-66 (1978).
10. Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Second Edition, 2002.
11. Raid, A., Khedr, W., El-dosuky, M., & Aoud, M. (2014). Image Restoration Based on Morphological Operations. International Journal of Computer Science, Engineering and Information Technology, 4(3), 9–21. https://doi.org/10.5121/ijcseit.2014.4302.
12. S. Grover, K. Arora, and S. Mitra, "Text Extraction from Document Images using Edge Information", in Annual IEEE India Conference (INDICON), Vol. 1-4, IEEE, Gujarat, India (2009).
13. Sheik, H. and Bovik, A., "Image information and visual quality, Image Processing", IEEE Transactions on, 15, 430-444, 2006.
14. Wang, Z., Bovik, A., Sheik, H., and Simoncelli, E,. "Image quality assessment: From error visibility to structural similarity", Image Processing, IEEE Transactions on, 13,600-612, 2004.
15. Young, Ian. (1983). Image analysis and mathematical morphology, by J. Serra. Academic Press, London, 1982.

AUTHORS

**Abhishek Gupta** received B.Tech degree in Electronics Engineering from Institute of Engineering and Technology, Lucknow in 2013. Presently, he is working in the Department of Technical Education, Uttar Pradesh. Currently, pursuing his M.Tech degree from Institute of Engineering and Technology, Dr. Ram Manohar Lohia Avadh University, Ayodhya. His area of interest includes Image Processing and Digital Signal Processing.
E-mail: abhishekdte@gmail.com

**Ramapati Mishra** received B.Tech degree in Electronics and Communication Engineering in 1990, M.Tech degree in Electronics (optical communication) in 2000 and Ph.D. in Electronics (optical Amplifiers in optical communication) in 2010. He is a member of IEEE, IETE and IEI. He has published around 20 research papers in various national and international journals and conferences. He has guided about 12 M.Tech. Thesis and is also supervising Ph.D. work. His research interest includes Radar Technology, Optical communication and Digital Signal Processing. (CORRESPONDING AUTHOR)
E-mail: director.rpm@gmail.com

**Ashutosh Kumar Singh** received B.Tech degree in Electronics and Communication Engineering from P.S.I.T., Kanpur in 2008, M.Tech degree in Digital Systems from M.M.M. Engineering College in 2011, Gorakhpur and Ph.D. from University of Allahabad in 2018. He has published around 25 research papers in various national and international journals and conferences. His research interest includes Digital Systems, Optical Communication, Optical Networks, Survivable Optical Networks and Digital Signal Processing.
E-mail: aksinghelectronics@gmail.com