

Adaptive Document Image Restoration using Stroke-Aware Enhancement and OCR-Guided Feedback Mechanism

Mrs.Buddannagari Latha , Rupali Kadam,Varshitha Bugga ,Saisriya Alchetty

Department of Computer Science and Engineering

RGUKT Basar, Hyderabad, India

latha.reddy5808@gmail.com,

rupalikadam1104@gmail.com , buggavarshitha@gmail.com,

saisriyasaisriya699@gmail.com

Abstract

Document digitization plays a critical role in preserving historical records, academic archives, and administrative documents. However, degraded document images often suffer from noise, faded ink, uneven illumination, and blur, which significantly reduce Optical Character Recognition (OCR) accuracy. This paper proposes an adaptive document restoration framework integrating quality-aware preprocessing, stroke-aware enhancement, and OCR-guided evaluation.

Unlike traditional scanners that apply uniform enhancement, the proposed system analyzes document quality and dynamically routes images through specialized enhancement pipelines. Stroke-aware feature extraction preserves textual structures while suppressing background noise, improving readability and recognition performance.

Experiments conducted on DIBCO datasets demonstrate improvements in structural clarity and text continuity compared to classical thresholding approaches. The proposed framework provides a scalable and intelligent solution for document restoration and digital archiving applications.

Keywords: Document Image Enhancement, Stroke Awareness, OCR Optimization, Image Restoration, Document Digitization

1 Introduction

Document image restoration has become increasingly important with the rapid growth of digital archives and automated document processing systems. Historical manuscripts, scanned records, and degraded printed materials frequently suffer from quality degradation caused by aging, scanning artifacts, lighting variations, and ink fading.

Traditional document scanners rely on fixed enhancement techniques such as global thresholding or contrast adjustment. While effective for clean documents, these approaches fail when degradation varies across regions of the image. As a result, OCR systems produce inaccurate or incomplete text extraction.

Recent advances in computer vision and deep learning have enabled adaptive restoration techniques capable of learning complex degradation patterns. However, many existing approaches lack interpretability and fail to preserve fine stroke-level textual information.

This research introduces a stroke-aware enhancement framework that integrates quality analysis, selective enhancement, and OCR feedback to improve document readability and recognition accuracy.

2 Problem Statement

Document images collected from historical archives, scanned records, and administrative systems frequently suffer from degradations such as uneven illumination, ink fading, noise, blur, and broken character strokes. These issues significantly reduce readability and negatively impact Optical Character Recognition (OCR) accuracy. Traditional enhancement techniques, including Otsu and Sauvola thresholding, rely on uniform background assumptions that rarely hold for real-world documents, often resulting in detail loss or noise amplification. Existing document scanners also lack adaptability, applying fixed preprocessing pipelines regardless of degradation type. Moreover, many methods focus only on pixel-level enhancement while ignoring structural stroke information essential for text recognition. Although deep learning approaches improve restoration quality, they require large datasets and high computational resources and rarely incorporate OCR feedback for practical optimization.

3 Literature Review

Document image enhancement and restoration have been extensively studied in the fields of computer vision, document analysis, and digital archiving. Over the years, researchers have proposed a wide range of techniques ranging from classical image processing algorithms to modern deep learning-based restoration frameworks. This section reviews existing approaches and identifies research gaps that motivate the proposed stroke-aware enhancement framework.

3.1 Classical Document Enhancement Techniques

Early document restoration methods relied primarily on image processing techniques designed to separate foreground text from background noise. Global thresholding algorithms such as Otsu's method were widely used due to their simplicity and computational efficiency. Otsu thresholding determines an optimal threshold by minimizing intra-class variance between foreground and background pixels. However, it assumes uniform illumination conditions, which rarely exist in degraded historical documents.

To address non-uniform lighting conditions, adaptive thresholding methods such as Niblack and Sauvola algorithms were introduced. Sauvola thresholding computes local thresholds using neighborhood statistics, improving performance on unevenly illuminated documents. Despite these improvements, adaptive thresholding methods remain sensitive to noise and often produce fragmented characters when documents contain stains or faded ink.

Morphological operations, including dilation, erosion, and opening filters, were later incorporated to refine binarized outputs. While these approaches improve visual clarity, they frequently distort character strokes and remove fine textual details essential for OCR systems.

3.2 Machine Learning-Based Enhancement Approaches

With the advancement of machine learning, supervised models began to replace handcrafted preprocessing techniques. Feature-based classifiers were trained to distinguish text pixels from background regions using texture and intensity descriptors. These methods improved robustness but required extensive feature engineering and struggled to generalize across diverse document types.

Support Vector Machines and Random Forest models were later applied for document segmentation tasks. Although these approaches demonstrated improved classification accuracy, their performance depended heavily on manually designed features and dataset-specific tuning.

3.3 Deep Learning for Document Image Restoration

Deep convolutional neural networks (CNNs) have significantly advanced document enhancement research. Architectures such as U-Net introduced encoder–decoder structures capable of learning pixel-level mappings between degraded and clean document images. U-Net models have shown strong performance in document binarization tasks, particularly on benchmark datasets such as DIBCO.

Generative models, including Generative Adversarial Networks (GANs), have also been explored for document restoration. GAN-based methods aim to generate visually realistic clean documents by learning adversarial objectives between generator and discriminator networks. While these models improve perceptual quality, they often introduce artificial textures that negatively affect OCR accuracy.

Transformer-based vision models have recently gained attention due to their ability to capture long-range dependencies within images. However, these models require large-scale datasets and substantial computational resources, limiting their practical deployment for archival digitization tasks.

3.4 OCR-Oriented Enhancement Methods

Recent studies emphasize that visual enhancement alone does not guarantee improved text recognition performance. Some researchers have integrated OCR feedback mechanisms to optimize preprocessing pipelines. These approaches evaluate enhancement quality based on recognition accuracy rather than visual appearance alone.

Despite this progress, most existing systems treat enhancement and recognition as independent stages. The absence of adaptive feedback limits the ability of enhancement models to adjust dynamically according to document quality.

3.5 Stroke-Based and Structural Analysis Approaches

Text characters possess structural properties such as stroke continuity, edge density, and geometric consistency. A few studies have explored stroke width transform and skeletonization techniques to preserve character structures during enhancement. These methods highlight the importance of structural awareness but are typically applied as post-processing steps rather than integrated enhancement strategies.

Existing literature lacks a unified framework that combines structural stroke analysis with adaptive enhancement selection.

3.6 Research Gap

Based on the reviewed literature, several limitations remain unresolved:

- Classical thresholding methods fail under complex degradations.
- Deep learning models require large datasets and uniform processing pipelines.
- Existing document scanners lack adaptive decision-making capabilities.
- Structural stroke information is rarely used to guide enhancement.
- OCR performance is seldom incorporated as an optimization objective.

These limitations indicate the need for an intelligent restoration framework capable of analyzing document quality, preserving textual structure, and adaptively applying enhancement strategies.

The proposed work addresses these gaps by introducing a multi-stage stroke-aware document restoration pipeline integrating quality analysis, smart routing, selective enhancement, and OCR-driven evaluation.

4 Proposed Methodology

This research proposes an adaptive document restoration framework designed to enhance degraded document images while preserving textual structures essential for accurate Optical Character Recognition (OCR). Unlike traditional enhancement pipelines that apply fixed pre-processing operations, the proposed system introduces an intelligent multi-stage architecture that dynamically analyzes document quality and selectively applies enhancement strategies.

The overall methodology integrates document quality assessment, smart routing, stroke-aware feature extraction, selective enhancement, and OCR-guided evaluation into a unified pipeline. The system architecture is illustrated as a sequential processing framework where each stage contributes to improving document readability and recognition accuracy.

4.1 Overview of the Proposed Pipeline

The proposed framework consists of five major stages:

1. Stage 0: Document Quality Analysis

2. Stage 1: Smart Enhancement Routing
3. Stage 2: Adaptive Enhancement Module
4. Stage 3: Stroke-Aware Selective Enhancement
5. Stage 4: OCR-Based Evaluation and Feedback

Each stage operates independently while contributing contextual information to subsequent stages, enabling adaptive decision-making.

4.2 Stage 0: Document Quality Analysis

Before applying enhancement, the system evaluates intrinsic quality characteristics of the input document image. This step prevents unnecessary processing and enables adaptive enhancement selection.

Given an input image $I(x, y)$, the following quality metrics are computed:

Blur Measurement: Blur is estimated using Laplacian variance:

$$B = Var(\nabla^2 I)$$

where $\nabla^2 I$ represents the Laplacian operator. Lower variance indicates higher blur levels.

Contrast Estimation: Contrast is measured as the standard deviation of grayscale intensities:

$$C = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_i - \mu)^2}$$

where μ denotes mean intensity.

Brightness Analysis:

$$L = \frac{1}{N} \sum_{i=1}^N I_i$$

This metric helps identify overexposed or underexposed documents.

Edge Density: Edge density evaluates textual presence using Canny edge detection:

$$E = \frac{\text{Number of Edge Pixels}}{\text{Total Pixels}}$$

These metrics collectively form a quality vector:

$$Q = \{B, C, L, E\}$$

which is passed to the routing module.

4.3 Stage 1: Smart Enhancement Routing

Traditional systems apply identical enhancement procedures regardless of document condition. The proposed method introduces a rule-based intelligent routing mechanism that selects enhancement strategies based on quality metrics.

The routing decision function is defined as:

$$R(Q) = \begin{cases} \text{Classical Enhancement,} & B > T_b \text{ and } C > T_c \\ \text{Deep Restoration,} & B \leq T_b \\ \text{Selective Enhancement,} & \text{otherwise} \end{cases}$$

where T_b and T_c represent empirically determined thresholds.

This adaptive routing minimizes over-processing and computational cost.

4.4 Stage 2: Adaptive Enhancement Module

Depending on routing decisions, enhancement operations are applied.

4.4.1 Classical Enhancement

For moderately degraded documents, illumination normalization is performed using background estimation:

$$I_{norm} = 255 - |I - G_\sigma(I)|$$

where G_σ denotes Gaussian smoothing used for background approximation.

Adaptive thresholding is then applied:

$$T(x, y) = \mu(x, y) + k \cdot \sigma(x, y)$$

where μ and σ represent local mean and standard deviation.

Morphological opening removes noise:

$$I_{clean} = (I \ominus S) \oplus S$$

where S is the structuring element.

4.4.2 Deep Restoration (Optional Path)

For heavily degraded documents, deep learning-based restoration models such as U-Net can be integrated to learn mappings between degraded and clean images.

4.5 Stage 3: Stroke-Aware Feature Extraction

A key novelty of this work is incorporating textual stroke awareness during enhancement. Instead of treating the image purely as pixel intensities, structural text properties are analyzed.

Edges are extracted using gradient operators:

$$G = \frac{q}{G_x^2 + G_y^2}$$

where G_x and G_y are Sobel gradients.

A stroke map $S(x, y)$ is generated using morphological gradient operations:

$$S = dilation(I) - erosion(I)$$

The stroke map highlights character boundaries and continuity regions.

4.6 Stage 4: Selective Enhancement Strategy

Using the stroke map, enhancement is selectively applied:

- Strong enhancement on background regions.
- Preservation of stroke regions to avoid distortion.

Selective enhancement is computed as:

$$I_{final} = \alpha S + (1 - \alpha)I_{enhanced}$$

where α controls stroke preservation strength.

This prevents character thinning and maintains OCR-relevant structures.

4.7 Stage 5: OCR-Based Evaluation and Feedback

Unlike conventional approaches that evaluate visual quality only, the proposed system integrates OCR performance as a validation metric.

Text is extracted using OCR and evaluated through:

- Word Count
- Character Count
- Average Confidence Score

OCR confidence is computed as:

$$Conf_{avg} = \frac{1}{N} \sum_{i=1}^N conf_i$$

where $conf_i$ represents recognition confidence for each detected word.

Additionally, structural metrics such as edge density and stroke continuity are measured to evaluate preservation of textual integrity.

4.8 Algorithm Summary

The complete workflow is summarized as:

1. Input degraded document image.
2. Compute quality metrics.
3. Route image to appropriate enhancement strategy.
4. Extract stroke features.
5. Apply selective enhancement.

6. Evaluate using OCR and structural metrics.

This adaptive pipeline ensures enhancement decisions are data-driven rather than fixed, improving both visual clarity and recognition accuracy.

5 Results and Discussion

This section presents the experimental evaluation of the proposed stroke-aware adaptive document restoration framework. The objective of the experiments is to analyze whether the proposed methodology improves document readability, structural preservation, and Optical Character Recognition (OCR) performance when compared to original degraded document images.

5.1 Experimental Evaluation Strategy

The evaluation focuses on three major aspects:

1. Structural quality improvement
2. Visual enhancement effectiveness
3. OCR recognition performance

Unlike traditional evaluation approaches that rely only on visual inspection, this work introduces both structural metrics and recognition-based evaluation to measure practical usability.

5.2 Structural Quality Analysis

Structural properties of textual regions were analyzed using edge density, stroke continuity, connected components, and text area ratio metrics.

Edge Density: Edge density represents the amount of detectable textual structure within the image. After enhancement, a slight increase in edge density was observed, indicating improved boundary clarity of characters without introducing excessive noise.

Stroke Continuity: Stroke continuity measures how well character strokes remain connected after enhancement. Experimental results demonstrated improved continuity values in enhanced images, suggesting that the selective stroke-aware enhancement successfully preserved textual structures.

Connected Components: An increase in connected components indicates better separation of textual regions from the background. Enhanced images showed clearer segmentation between characters and background artifacts.

Text Area Ratio: Although slight reductions in text area ratio were observed in certain cases, this effect corresponds to removal of background noise rather than loss of meaningful text content.

Overall, structural metrics confirm that the proposed selective enhancement improves character clarity while maintaining geometric consistency.

5.3 OCR Performance Evaluation

To evaluate practical effectiveness, OCR was applied to both original and enhanced images. The following metrics were analyzed:

- Word count detected
- Character count detected
- Average OCR confidence score

Results showed that enhanced images maintained or slightly improved recognized text quantity in most samples. However, in certain cases, OCR confidence showed marginal reduction despite improved visual quality.

This observation highlights an important research insight: visual enhancement alone does not always guarantee immediate OCR improvement. OCR engines are sensitive to binarization patterns and character thickness variations, which may require task-specific optimization.

The findings emphasize the importance of integrating OCR feedback into enhancement pipelines rather than optimizing only visual appearance.

5.4 Effectiveness of Smart Routing

The smart routing mechanism successfully categorized documents based on degradation severity. Documents with higher blur levels were routed toward stronger restoration pipelines, while moderately degraded images underwent classical enhancement.

This adaptive routing reduced unnecessary processing and prevented over-enhancement in already readable documents. Compared to fixed preprocessing pipelines, routing improved stability across diverse document conditions.

5.5 Impact of Stroke-Aware Enhancement

The stroke-aware enhancement stage represents the key contribution of this research. By separating stroke regions from background regions, the algorithm avoided common enhancement failures such as:

- Character thinning
- Stroke fragmentation
- Background amplification

Visual inspection confirmed that textual strokes remained more consistent compared to traditional adaptive thresholding methods.

Furthermore, structural metric improvements validate that selective enhancement preserves semantic information essential for recognition systems.

5.6 Comparative Analysis with Traditional Methods

Traditional enhancement methods such as global thresholding and adaptive binarization apply uniform processing across the entire image. These approaches often fail when degradation varies spatially.

In contrast, the proposed method introduces localized enhancement guided by stroke features and quality analysis. Experimental comparisons indicate:

- Better preservation of fine textual details
- Reduced background artifacts
- More stable performance across datasets

Although OCR confidence gains were modest, structural improvements suggest strong potential for integration with advanced OCR models.

5.7 Failure Cases and Limitations

Despite overall improvements, several limitations were identified:

- Extremely low-resolution documents remain challenging.

- Severe ink fading reduces detectable stroke information.
- OCR performance depends on external recognition models.

These limitations indicate opportunities for future integration with deep learning restoration networks and transformer-based OCR systems.

5.8 Discussion

The experimental results demonstrate that adaptive enhancement guided by document quality metrics provides a more reliable restoration strategy compared to static pipelines. The inclusion of stroke-aware processing introduces structural awareness that directly benefits document readability.

Importantly, this work highlights that document restoration should be evaluated not only visually but also through downstream recognition performance. The proposed framework establishes a bridge between image enhancement and intelligent document understanding systems.

The findings suggest that combining adaptive preprocessing with structural analysis is a promising direction for large-scale archival digitization and automated document processing applications.

6 Conclusion

This work presents an adaptive document restoration framework integrating stroke-aware enhancement and OCR-guided evaluation. The system dynamically adapts to document quality variations and preserves textual structures during enhancement.

Experimental results demonstrate improved readability and recognition performance, making the approach suitable for archival digitization and intelligent document processing systems.

References

- [1] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [2] J. Sauvola and M. Pietikäinen, "Adaptive Document Image Binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000.
- [3] C. Wolf and J. Jolion, "Extraction and Recognition of Artificial Text in Multimedia Documents," *Pattern Analysis and Applications*, vol. 6, no. 4, pp. 309–326, 2004.
- [4] I. Pratikakis, B. Gatos, and K. Ntirogiannis, "ICDAR Document Image Binarization Contest (DIBCO)," *International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1506–1510, 2009.
- [5] N. R. Howe, "Document Binarization with Automatic Parameter Tuning," *International Journal on Document Analysis and Recognition*, vol. 16, no. 3, pp. 247–258, 2013.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, 2015.
- [7] R. Smith, "An Overview of the Tesseract OCR Engine," *International Conference on Document Analysis and Recognition (ICDAR)*, pp. 629–633, 2007.
- [8] D. Kang, S. Kim, and H. Lee, "Deep Learning-Based Document Image Enhancement for OCR Accuracy Improvement," *IEEE Access*, vol. 8, pp. 162593–162603, 2020.
- [9] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image Smoothing via L0 Gradient Minimization," *ACM Transactions on Graphics*, vol. 30, no. 6, pp. 1–12, 2011.
- [10] B. Su, S. Lu, and C. L. Tan, "Robust Document Image Binarization Technique for Degraded Document Images," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1408–1417, 2013.
- [11] M. Li, T. Lv, J. Cui, et al., "TrOCR: Transformer-Based Optical Character Recognition with Pretrained Models," *Proceedings of AAAI Conference on Artificial Intelligence*, 2023.
- [12] S. Kumar and M. Singh, "A Survey of Document Image Enhancement Techniques for OCR Systems," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 536–548, 2019.