



Impact of Image Preprocessing on OCR Performance for English Printed Receipts: A Comparative Study of Tesseract and Easy-OCR

Alya M. Ali *

Department of Software Engineering, Faculty of Information Technology,
University of Tripoli, Tripoli, Libya

أثر المعالجة المسبقة للصور على أداء التعرف الضوئي على الحروف للإيصالات المطبوعة باللغة
الإنجليزية: دراسة مقارنة بين محركي Tesseract و EasyOCR

علياء مفتاح علي *

قسم هندسة البرمجيات، كلية تقنية المعلومات، جامعة طرابلس، طرابلس، ليبيا

*Corresponding author: aly.issa@uot.edu.ly

Received: July 19, 2026

Accepted: October 02, 2026

Published: October 15, 2026

Abstract:

Optical Character Recognition (OCR) plays an important role in extracting machine-readable information from scanned documents and receipts; however, recognition performance can be affected by image quality and document characteristics. Although image preprocessing is commonly used to improve OCR input quality, its effectiveness may vary across OCR systems. This study experimentally investigates the effect of image preprocessing on the performance of two open-source OCR engines, Tesseract and EasyOCR, for English printed receipt images. A controlled comparative experiment was conducted using 200 images from the Scanned Receipts OCR and Information Extraction (SROIE) dataset. Each OCR engine was evaluated under four image conditions: the original images as a baseline, Gaussian Blur, Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Non-Local Means (Fast NL-Means) denoising. Recognition performance was assessed using Character Error Rate (CER) and Word Error Rate (WER), while OCR processing time was used to evaluate computational efficiency. The results demonstrate that preprocessing does not provide a uniform improvement across OCR systems. CLAHE produced the best recognition performance for EasyOCR, reducing CER from 27.10% to 26.23% and WER from 61.17% to 58.38%. For Tesseract, Gaussian Blur produced the lowest WER, reducing it from 49.46% to 48.54%, whereas the original images yielded the lowest CER of 26.19%. Fast NL-Means did not provide an overall recognition benefit for either system. Tesseract also exhibited substantially lower OCR execution time than EasyOCR under the tested conditions. These findings indicate that image preprocessing should not be regarded as a universally beneficial OCR step; instead, its selection should be considered in relation to the OCR engine, document type, and image characteristics.

Keywords: Optical Character Recognition (OCR), Image Preprocessing, Tesseract, EasyOCR, SROIE, Character Error Rate (CER), Word Error Rate (WER).

المخلص

يؤدي التعرف الضوئي على الحروف (OCR) دورًا مهمًا في استخراج المعلومات القابلة للمعالجة آليًا من المستندات والإيصالات المسحوبة ضوئيًا، إلا أن أداء التعرف قد يتأثر بجودة الصور وخصائص المستندات. وعلى الرغم من شيوع استخدام المعالجة المسبقة للصور لتحسين جودة المدخلات لأنظمة التعرف الضوئي على الحروف، فإن فاعليتها قد تختلف باختلاف النظام المستخدم. تهدف هذه الدراسة إلى التحقق تجريبيًا من أثر المعالجة المسبقة للصور في أداء محركي مفتوح المصدر للتعرف الضوئي على الحروف، هما Tesseract و EasyOCR، باستخدام صور الإيصالات المطبوعة باللغة الإنجليزية. أجريت تجربة مقارنة مضبوطة باستخدام 200 صورة من مجموعة بيانات الإيصالات المسحوبة ضوئيًا

والتعرف الضوئي على الحروف واستخراج المعلومات (SROIE). جرى تقييم كل محرك في أربع حالات للصور، شملت الصور الأصلية بوصفها الحالة المرجعية، والتمويه الغاوسي (Gaussian Blur)، ومعادلة المدرج التكراري التكيفية محدودة التباين (CLAHE)، وإزالة الضوضاء باستخدام خوارزمية المتوسطات غير المحلية السريعة. (-Fast NL-Means) (Means) فُيَم أداء التعرف باستخدام معدل الخطأ على مستوى الحروف (CER) ومعدل الخطأ على مستوى الكلمات (WER)، بينما استُخدم زمن معالجة التعرف الضوئي على الحروف لتقييم الكفاءة الحسابية. أظهرت النتائج أن المعالجة المسبقة لا تحقق تحسناً موحداً في أداء أنظمة التعرف الضوئي على الحروف. فقد حققت تقنية CLAHE أفضل أداء للتعرف باستخدام EasyOCR، إذ انخفض معدل الخطأ على مستوى الحروف من 27.10% إلى 26.23%، ومعدل الخطأ على مستوى الكلمات من 61.17% إلى 58.38%. أما بالنسبة إلى Tesseract، فقد حقق التمهويه الغاوسي أدنى معدل للخطأ على مستوى الكلمات، حيث انخفض من 49.46% إلى 48.54%، في حين حققت الصور الأصلية أدنى معدل للخطأ على مستوى الحروف بلغ 26.19%. ولم تحقق تقنية Fast NL-Means تحسناً إجمالياً في أداء التعرف لأي من المحركين. كما أظهر Tesseract زمن تنفيذ أقل بكثير من EasyOCR في ظل ظروف الاختبار المستخدمة. وتشير هذه النتائج إلى أن المعالجة المسبقة للصور لا ينبغي اعتبارها خطوة مفيدة بصورة مطلقة في أنظمة التعرف الضوئي على الحروف، بل ينبغي اختيار تقنياتها وفقاً لمحرك التعرف المستخدم، ونوع المستند، وخصائص الصور.

الكلمات المفتاحية: التعرف الضوئي على الحروف (OCR)، المعالجة المسبقة للصور، Tesseract، EasyOCR، مجموعة بيانات SROIE، معدل الخطأ على مستوى الحروف (CER)، معدل الخطأ على مستوى الكلمات (WER).

1.Introduction

Optical Character Recognition (OCR) has become an important technology for automatically converting text contained in images and scanned documents into machine-readable text. OCR systems are widely used in document digitization, information retrieval, automated data entry, and document processing applications. In particular, receipts and invoices contain valuable textual information, including product names, quantities, prices, dates, and transaction details, which can be difficult and time-consuming to extract manually. Automating the extraction of such information can reduce manual data entry and support the development of intelligent document-processing applications. Therefore, the accuracy and efficiency of OCR systems are important factors in determining the reliability of applications that depend on automatically extracted textual information [1].

Despite the advantages of OCR technology, the recognition of text from scanned receipts remains challenging because of the physical and visual characteristics of these documents. Receipt images may contain low-contrast text, background noise, blurred characters, uneven illumination, printing imperfections, and variations in text layout. The SROIE benchmark was specifically established to investigate OCR and information extraction from scanned receipts, reflecting the practical importance of this document type in document analysis applications [2]. These factors can make characters less distinguishable and consequently increase recognition errors. Therefore, applying suitable image preprocessing techniques before OCR can play an important role in improving the visual quality of the input image and potentially enhancing text recognition performance.

Image preprocessing techniques are commonly used to enhance text visibility and reduce unwanted image variations before the OCR stage. Among these techniques, noise reduction and contrast enhancement are frequently used in document image processing. Gaussian filtering is a conventional smoothing operation that can reduce high-frequency variations, while CLAHE provides localized contrast enhancement with a limitation on contrast amplification [3], [4]. Non-Local Means (NL-Means) provides another approach to image denoising by exploiting similarities between image regions rather than relying only on local neighborhoods [5]. Previous OCR-related studies have also reported that preprocessing can affect recognition performance, although the magnitude and direction of the effect may vary according to the document characteristics and OCR system [6]–[8].

Among widely used open-source OCR solutions, Tesseract and EasyOCR provide different approaches to text recognition. Tesseract is an established open-source OCR engine whose development and recognition architecture have been documented extensively [9]. EasyOCR is a ready-to-use deep-learning-based OCR library that integrates text detection and recognition components and supports multiple languages [10]. These differences make the two systems suitable candidates for comparative evaluation under different image preprocessing conditions. Comparing them using the same images, preprocessing procedures, and evaluation criteria can provide a clearer understanding of whether a particular preprocessing technique consistently benefits OCR performance or whether its effectiveness depends on the OCR engine being used.

Previous studies have investigated OCR performance on scanned documents and receipts and have examined the use of image preprocessing to improve recognition results. For example, Sporici et al. investigated preprocessing methods for improving the accuracy of Tesseract, while Ramadhan et al. examined image-quality enhancement techniques for Tesseract-based extraction from Indonesian identity cards [6], [7]. More recently, Puteri et al. compared Tesseract and EasyOCR on receipt images and reported differences in OCR performance under different preprocessing conditions [8]. However, differences in datasets, preprocessing procedures, OCR configurations, and evaluation measures make direct comparison between previous findings difficult.

Accordingly, this study investigates the effect of three image preprocessing techniques Gaussian Blur, Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Non-Local Means (Fast NL-Means) on the performance of Tesseract and EasyOCR for English printed receipt images. The study uses 200 images from the Scanned Receipts OCR and Information Extraction (SROIE) dataset and evaluates each OCR engine under four conditions: the original images as a baseline and the three preprocessing conditions. OCR performance is assessed using Character Error Rate (CER), Word Error Rate (WER), and OCR processing time [2].

The main contributions of this study are as follows:

1. A controlled comparative evaluation of Tesseract and EasyOCR using the same set of 200 SROIE receipt images and identical evaluation procedures.
2. An empirical assessment of Gaussian Blur, CLAHE, and Fast NL-Means to examine whether their effects are consistent across different OCR engines.
3. A multi-dimensional evaluation of OCR performance using character-level errors, word-level errors, and processing time.
4. An analysis of the engine-dependent effects of image preprocessing, providing empirical evidence on whether preprocessing techniques should be selected according to the OCR engine and document characteristics rather than applied as universally beneficial operations.

2. LITERATURE REVIEW

2.1 OCR in Scanned Receipts and Financial Documents

Optical Character Recognition (OCR) is an important component of document digitization, particularly for receipts, invoices, and other financial documents. OCR converts text contained in images into machine-readable representations, enabling subsequent processing, indexing, and information extraction [1]. However, OCR performance can be affected by the visual characteristics of the input images. Receipt images may contain noise, low contrast, faded text, shadows, folds, and irregular layouts, which can make text detection and recognition more challenging.

Recent research reviewing OCR methods for invoices and receipts has highlighted the continued use of open-source OCR systems, with Tesseract being among the most widely used tools in this area [11].

The importance of receipt OCR has been demonstrated by the ICDAR2019 Scanned Receipt OCR and Information Extraction (SROIE) competition, which established a dataset and evaluation tasks for receipt text localization, OCR, and key information extraction [2].

The availability of such a benchmark has provided a standardized basis for investigating OCR performance on receipt images.

Previous studies have examined OCR performance on different types of real-world documents. Ramadhan et al. examined OCR-based text extraction from Indonesian identity cards and found that degraded image quality and document characteristics can affect Tesseract-based recognition [7]. Puteri et al. specifically compared Tesseract and EasyOCR on receipt images and observed differences in recognition performance under different preprocessing conditions [8]. These studies indicate that OCR performance should be considered in relation to both the characteristics of the input document and the OCR system.

Receipt-specific OCR applications have also been investigated in practical information extraction systems. Indrakusuma et al. [12] developed a system for recognizing and classifying text from material purchase receipts using Tesseract OCR, followed by text processing and classification.

2.2 Comparative Evaluation of OCR Engines

Different OCR engines employ different approaches for text detection and recognition and may consequently respond differently to the same input images and preprocessing operations. Tesseract is a widely used open-source OCR engine whose development and architecture have been documented extensively [9]. EasyOCR, in contrast, provides a ready-to-use deep-learning-based OCR pipeline incorporating text detection and recognition components [10].

Previous comparative studies have reported differences in OCR performance across engines and document types. Puteri et al. [8] compared the performance of Tesseract and EasyOCR on receipt images using 50 invoices selected from the Find It Again Dataset available on Kaggle. The study evaluated the two OCR engines under four image conditions: the original images, grayscale conversion, Otsu thresholding, and Gaussian Blur. OCR performance was evaluated using Character Error Rate (CER), calculated based on the Levenshtein distance between the recognized text and the corresponding ground-truth text. The results showed that the relative performance of Tesseract and EasyOCR varied across the preprocessing conditions, indicating that the choice of preprocessing

could affect the performance of the two OCR engines differently [8]. More broadly, comparative OCR research has demonstrated that the choice of OCR engine can affect recognition outcomes when the same document images are processed under different configurations.

The comparison therefore indicates that OCR performance depends on both the input image and the OCR engine used for recognition. Therefore, evaluating preprocessing techniques across more than one OCR engine provides a more informative assessment than evaluating a preprocessing method with a single system.

2.3 Image Preprocessing for OCR

Image preprocessing is commonly applied before OCR to improve the visual characteristics of document images and facilitate text detection and recognition. Common preprocessing operations include smoothing, denoising, contrast enhancement, and thresholding. General image-processing references describe these operations as fundamental techniques for modifying image characteristics before subsequent analysis [3].

Experimental evidence has also shown that preprocessing and binarization can substantially affect OCR results. Malinski and Okarma [13] evaluated several image preprocessing and binarization methods for OCR-based recognition of integrated-circuit labels and found that the average Levenshtein distance decreased from 27.63 with the initial Tesseract configuration to 19.32 after applying the evaluated preprocessing pipeline.

Previous experimental studies have shown that preprocessing can affect OCR performance in different ways. Sporici et al. [6] investigated convolution-based preprocessing for Tesseract and found improvements in recognition accuracy under their experimental conditions. Ramadhan et al. [7] also observed improvements in Tesseract performance after applying image enhancement techniques to Indonesian identity-card images. In comparison, Puteri et al. [8] examined Tesseract and EasyOCR on receipt images and identified differences in performance before and after preprocessing. These findings indicate that the effect of preprocessing may vary across OCR engines, preprocessing techniques, and document types.

Sutriawan et al. [14] examined the use of image enhancement in OCR-based recognition of scanned shopping receipts. Their results showed improvements in character and word recognition using a system that combined image enhancement with a neural-network-based approach. These findings provide further evidence of the potential benefits of image enhancement for receipt OCR.

2.3.1 Noise Reduction

Denoising techniques aim to reduce unwanted image variations while preserving important structures. Gaussian filtering is a conventional image-smoothing operation that can be used to reduce high-frequency variations before subsequent image analysis [3]. However, smoothing may also alter fine image details that are relevant to character recognition.

Fast Non-Local Means (Fast NL-Means) represents another denoising approach. The original Non-Local Means method introduced by Buades et al. exploits similarities between different image regions to perform denoising while preserving image structures [5]. The effectiveness of denoising for OCR may therefore depend on the characteristics of the input images and the recognition system.

Kumar et al. [15] examined the use of image preprocessing techniques, including image scaling, contrast enhancement, binarization, noise removal, and deskewing, to improve OCR recognition from low-quality images.

Previous experimental studies have demonstrated that preprocessing can produce substantial but different effects on OCR performance. Sporici et al. [6] found that their convolution-based preprocessing approach increased character-level accuracy for Tesseract 4.0 from 0.134 to 0.616, while the F1 score increased from 0.163 to 0.729. Ramadhan et al. [7] similarly observed an increase in average correctly recognized characters from 30.82% to 64.97% after applying their preprocessing pipeline to Indonesian identity-card images. In a receipt-specific study, Puteri et al. [8] evaluated Tesseract and EasyOCR using 50 receipt images, with CER values ranging from approximately 18% to 25%. Their results showed that Tesseract achieved its best CER under denoising, whereas EasyOCR achieved its best result with grayscale preprocessing; thresholding produced the largest reduction in recognition performance.

These findings demonstrate that the effect of preprocessing varies across OCR systems, preprocessing methods, and document types. However, the studies used different datasets, preprocessing configurations, OCR systems, and evaluation procedures, making direct comparison of their reported results difficult.

2.3.2 Contrast Enhancement

Contrast enhancement is another preprocessing approach used to improve the visibility of text in document images. Contrast Limited Adaptive Histogram Equalization (CLAHE) enhances local contrast by applying histogram equalization to localized image regions while limiting excessive contrast amplification [4]. The original

CLAHE method was introduced by Zuiderveld as a contrast-enhancement technique based on adaptive histogram equalization with contrast limitation [4].

CLAHE can therefore be useful when text-background contrast varies across an image. However, contrast enhancement may also modify or amplify image variations that are not useful for recognition. Consequently, its effect on OCR accuracy should be evaluated experimentally rather than assumed to be universally beneficial.

Ramadhan et al. investigated CLAHE as part of a sequential preprocessing pipeline for Tesseract-based OCR on degraded Indonesian e-KTP images [7]. Their experiment involved 10 e-KTP images, each containing 15 text attributes. After applying the complete preprocessing pipeline, the reported OCR accuracy increased from 39.55% to 64.97%. In addition, the average number of correctly recognized attributes increased from approximately 4 to 8 out of 15 attributes per e-KTP. These results indicate that the image-quality enhancement pipeline, which included CLAHE, was associated with a substantial improvement in Tesseract recognition performance under the tested conditions [7].

2.4 Synthesis of Previous Studies

The reviewed literature demonstrates three relevant observations.

First, OCR performance on real-world document images is affected by image quality and document characteristics [2], [6]– [8].

Second, different OCR engines may exhibit different levels of performance under the same document conditions. In particular, studies comparing Tesseract and EasyOCR on receipt images have shown that their performance can vary according to the preprocessing condition [8].

Third, image preprocessing can improve or reduce OCR performance depending on the preprocessing operation and the OCR system being used [6]– [8].

The original literature on image enhancement and denoising also indicates that preprocessing operations have different objectives and may affect image structures differently [3]– [5].

The studies differ in their datasets, preprocessing procedures, OCR configurations, and evaluation measures. These differences limit direct comparison of their reported results. Using the same receipt images, ground-truth references, preprocessing conditions, OCR systems, and evaluation measures in one experiment can reduce these sources of variation and allow the preprocessing effects to be compared under consistent conditions.

2.5 Research Gap

Previous studies have examined preprocessing effects using different document types, OCR engines, preprocessing operations, and evaluation measures [6]– [8]. The available results therefore cannot always be compared directly, since differences in datasets, OCR configurations, and experimental procedures may contribute to the reported performance changes.

The comparison by Puteri et al. [8] is particularly relevant because it evaluates Tesseract and EasyOCR using receipt images. However, differences remain in the experimental conditions and evaluation framework considered in the present study. A comparison using the same receipt images, the same three preprocessing operations, and the same evaluation procedure for both OCR engines can provide a more controlled assessment of these effects. The present study therefore evaluates Gaussian Blur, CLAHE, and Fast Non-Local Means using CER, WER, and OCR processing time. Considering character-level errors, word-level errors, and processing time together allows the preprocessing effects to be examined from both recognition and computational perspectives.

2.6 Aim and Contributions of the Study

The aim of this study is to experimentally investigate the effect of image preprocessing on OCR performance for English printed receipt images. To achieve this aim, the study conducts a controlled comparative evaluation of two open-source OCR engines, Tesseract and EasyOCR, using 200 images from the Scanned Receipts OCR and Information Extraction (SROIE) dataset [2].

Each OCR engine is evaluated under four image conditions: the original image condition, used as the baseline, and three preprocessing conditions involving Gaussian Blur, Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Non-Local Means denoising. The same image set, ground-truth text, and evaluation procedures are used for both OCR engines to ensure a consistent basis for comparison.

The performance of the OCR systems is evaluated using three complementary measures: Character Error Rate (CER), Word Error Rate (WER), and processing time. CER is used to assess character-level recognition errors, WER to assess word-level recognition errors, and processing time to examine the computational efficiency associated with OCR inference [16].

The main contributions of this study are therefore:

1. A controlled comparative evaluation of Tesseract and EasyOCR under the same receipt-image dataset and experimental conditions.

2. An empirical assessment of three preprocessing techniques Gaussian Blur, CLAHE, and Fast Non-Local Means to determine whether their effects are consistent across different OCR engines.
3. A multi-dimensional evaluation of OCR performance using CER, WER, and processing time rather than relying on a single recognition measure.
4. An analysis of the practical effect of preprocessing on OCR recognition performance, providing evidence on whether preprocessing should be considered universally beneficial or selected according to the OCR engine and document characteristics.

Overall, the study seeks to provide a controlled and reproducible comparison of preprocessing effects on OCR performance for scanned receipt images, while avoiding the assumption that a particular preprocessing technique will improve all OCR systems equally.

3. METHODOLOGY

3.1 Research Design

This study employs a controlled comparative experimental design to investigate the effect of image preprocessing on Optical Character Recognition (OCR) performance for English printed receipt images. Two open-source OCR engines, Tesseract OCR and EasyOCR, were evaluated under four image conditions: the original image as the baseline, Gaussian Blur, Contrast Limited Adaptive Histogram Equalization (CLAHE), and Fast Non-Local Means (Fast NL-Means) denoising.

The same images, ground-truth references, text-normalization procedure, and evaluation criteria were used across all experimental conditions. Preprocessing parameters were fixed throughout the experiment and were not individually optimized for particular images or OCR engines. No OCR model was trained or fine-tuned on the dataset.

Each OCR configuration was evaluated using Character Error Rate (CER), Word Error Rate (WER), and OCR inference processing time. The resulting design allows the study to examine both the overall effect of preprocessing and whether the effect of a given preprocessing technique differs between OCR engines.

The experimental structure was:

Receipt Image → Preprocessing Condition → OCR Engine → Text Normalization → CER/WER and Time Evaluation.

3.2 Dataset and OCR Systems

3.2.1 Dataset

The experiments used the Scanned Receipts OCR and Information Extraction (SROIE) dataset from the ICDAR2019 competition. The original SROIE benchmark contains whole scanned receipt images together with annotations for OCR-related tasks [2]. The dataset was selected because it provides receipt images and corresponding textual annotations suitable for evaluating OCR recognition performance.

A subset of 200 images was systematically extracted from the training split by selecting the sequential continuous samples starting from the baseline index. This deterministic selection procedure was applied explicitly to ensure absolute evaluation integrity and eliminate any potential manual selection bias, thereby providing a standardized, reproducible benchmark block that captures the diverse noise and quality variations of the original dataset.

The selected images exhibit variation in visual characteristics, including differences in image resolution, illumination, and degrees of degradation or distortion. This variation provides a practical setting for examining whether preprocessing techniques affect OCR performance differently across receipt images.

The word-level annotations supplied by the dataset were used to reconstruct the ground-truth text for each receipt. The annotated words were concatenated in their original sequential order to form a complete reference string.

No manual image enhancement, cropping based on ground-truth bounding boxes, or image-specific correction was performed before OCR inference.

3.2.2 OCR Systems

Two OCR systems were evaluated: Tesseract OCR 5.x and EasyOCR 1.7.x.

For Tesseract, images were processed through the PyTesseract interface using the baseline default configuration of the native Tesseract OCR engine. Tesseract is a well-established open-source OCR engine, and its original architecture and recognition approach were described by Smith [9].

No custom Page Segmentation Mode (PSM) or layout-specific overrides were applied.

EasyOCR was configured for English recognition using the language argument ['en']. GPU acceleration was explicitly disabled using `gpu = False`, forcing inference to execute on the host CPU. These configurations follow the documented EasyOCR interface and usage [10].

Both OCR systems were applied to the complete receipt image rather than to manually cropped text regions or ground-truth bounding boxes. The pretrained OCR models were used directly without training or fine-tuning on SROIE.

3.3 Image Preprocessing

Three preprocessing techniques were investigated to represent different image-processing strategies: Gaussian Blur, CLAHE, and Fast NL-Means denoising. Each technique was applied independently, and all parameters were fixed across the 200 images.

3.3.1 Gaussian Blur

Gaussian Blur was used as a conventional smoothing operation to reduce high-frequency image variations and local noise [3]. It was configured with a 3×3 kernel and:

$$\sigma = 0$$

With this configuration, the standard deviation is automatically determined by the underlying implementation from the kernel dimensions.

3.3.2 CLAHE

Contrast Limited Adaptive Histogram Equalization (CLAHE) was used to enhance local image contrast while limiting excessive amplification of local intensity variations. CLAHE was originally introduced by Zuiderveld as a contrast-enhancement technique based on adaptive histogram equalization with contrast limitation [4].

The method was configured with a clip limit of 2.0 and a tile grid size of 8×8 .

The fixed configuration was applied consistently to all images without image-specific parameter tuning.

3.3.3 Fast Non-Local Means

Fast Non-Local Means (Fast NL-Means) denoising was employed to evaluate the effect of non-local noise reduction on OCR recognition. The underlying Non-Local Means approach was introduced by Buades et al. and uses similarities between image regions for denoising [5].

The method was configured with a filter strength $h = 10$, a template window of 7×7 pixels, and a search window of 21×21 pixels.

The parameters were kept constant across the complete dataset.

3.4 Experimental Setup and Procedure

Each of the 200 receipt images was evaluated under four image conditions:

1. Original image
2. Gaussian Blur
3. CLAHE
4. Fast NL-Means

Each condition was processed by both Tesseract and EasyOCR. Thus, the total number of OCR evaluations was: $200 \times 4 \times 2 = 1,600$

The original image served as the baseline for each OCR engine, while each preprocessing technique was compared independently with its corresponding baseline.

The experimental procedure consisted of the following steps:

1. Load the complete receipt image.
2. Apply the specified preprocessing technique.
3. Submit the resulting image to the selected OCR engine.
4. Record the raw OCR output.
5. Apply the predefined text-normalization procedure.
6. Compare the normalized OCR output with the corresponding normalized ground truth.
7. Calculate CER and WER.
8. Record OCR inference processing time.

No manual correction or post-recognition editing was applied to the OCR outputs.

The preprocessing techniques were evaluated independently rather than sequentially. Therefore, no combination of Gaussian Blur, CLAHE, and Fast NL-Means was used.

3.5 Text Normalization and Ground-Truth Preparation

To ensure consistent error evaluation, the same normalization procedure was applied to both the reconstructed ground-truth text and the raw OCR outputs before calculating CER and WER.

Three operations were performed.

First, all text was converted to lowercase, making the evaluation case-insensitive.

Second, whitespace was normalized by removing leading and trailing spaces and reducing consecutive whitespace characters to a single space.

Third, the word-level SROIE annotations were concatenated in their original sequential order to reconstruct the complete reference text for each receipt.

Punctuation marks, currency symbols, and numerical content were preserved during normalization. They were not removed because they constitute part of the textual content being evaluated.

Thus, the normalization procedure standardizes case and spacing differences while preserving the substantive textual content of the receipts.

3.6 Evaluation Metrics

Three complementary measures were used to evaluate the OCR systems: Character Error Rate (CER), Word Error Rate (WER), and OCR inference processing time.

3.6.1 Character Error Rate

Character Error Rate measures the character-level edit distance between the OCR output and the ground-truth reference. It is defined as [16]:

$$\text{CER} = (S + D + I) / N$$

where S, D, and I represent character substitutions, deletions, and insertions, respectively, and N is the number of characters in the ground-truth text.

Lower CER values indicate better character-level recognition performance.

For intuitive presentation, character accuracy may be derived as:

$$\text{Character Accuracy} = (1 - \text{CER}) \times 100$$

This derived accuracy is used only for interpretation and does not represent an independent evaluation metric.

3.6.2 Word Error Rate

Word Error Rate evaluates recognition errors at the word level and is defined as [16]:

$$\text{WER} = (S_v + D_v + I_v) / N_v$$

where S_v , D_v , and I_v represent word substitutions, deletions, and insertions, respectively, and N_v is the number of words in the ground-truth reference.

Lower WER values indicate better word-level recognition performance.

Word accuracy can similarly be derived as:

$$\text{Word Accuracy} = (1 - \text{WER}) \times 100$$

CER and WER are both based on normalized Levenshtein distance but operate at different textual levels [16].

3.6.3 Processing Time

OCR inference processing time was recorded as a complementary measure of computational efficiency. For each evaluation, elapsed time associated with the OCR inference operation was measured locally within the CPU execution loop and recorded in seconds.

The reported processing time represents OCR inference time only; the time required to execute the preprocessing operation itself was excluded.

Accordingly, processing-time results characterize the OCR execution under the specified computational environment rather than the total computational cost of the complete preprocessing-plus-OCR pipeline.

3.7 Comparative Analysis

The analysis compared each preprocessing condition with the original-image baseline for the same OCR engine. The primary comparisons were therefore:

- Original vs. Gaussian Blur
- Original vs. CLAHE
- Original vs. Fast NL-Means

CER and WER were treated as the primary recognition-performance measures, with lower values indicating better performance. Processing time was analyzed separately as a computational measure.

In addition to aggregate CER and WER results across the 200 images, image-level changes were examined to provide further insight into how preprocessing effects were distributed across individual receipts.

A supplementary image-level analysis was performed for CLAHE. For each image, the WER obtained after CLAHE was compared with the WER obtained from the corresponding original image. Each image was classified as improved, worsened, or unchanged according to whether its WER decreased, increased, or remained the same after preprocessing.

This analysis complements the aggregate results by indicating whether an observed overall change is broadly distributed across the evaluated images or concentrated in a subset of them.

3.8 Computational Environment and Reproducibility

All experiments were conducted using a standardized hardware and software environment. The host machine was equipped with an Intel Core i7 64-bit processor and operated under Microsoft Windows 11.

Both OCR systems were executed in CPU-only mode, without GPU acceleration. For EasyOCR, this was explicitly enforced using `gpu=False`.

The software configuration consisted of:

- Python 3.14.0
- PyTesseract
- Tesseract OCR 5.x
- EasyOCR 1.7.x

Tesseract was executed using its default configuration without custom PSM or layout overrides, while EasyOCR was configured for English recognition using `['en']`.

All preprocessing parameters were fixed across the 200-image subset, and the same images, ground-truth references, normalization procedure, and evaluation metrics were used for every OCR-condition combination.

The experimental pipeline automatically recorded OCR outputs, recognition metrics, and processing-time measurements, reducing manual intervention and supporting reproducibility of the experimental procedure.

3.9 Methodological Scope and Limitations

The experimental design provides a controlled basis for comparing preprocessing effects, but the findings should be interpreted within the scope of the study.

The evaluation is based on 200 images from the SROIE training split rather than the complete dataset. Therefore, the results should not automatically be generalized to all receipt datasets or OCR applications.

Furthermore, the study considers two OCR engines and three preprocessing techniques. The findings consequently characterize the behavior of the evaluated configurations rather than establishing universal superiority for any preprocessing method.

The natural variation in image resolution, illumination, and degradation is useful for assessing practical preprocessing effects, but it may also contribute to differences in the response of individual images to a given technique.

Finally, processing-time measurements are dependent on the specified hardware, software versions, and runtime conditions. They should therefore be interpreted as benchmark measurements obtained under the reported experimental environment.

Despite these limitations, the controlled use of the same image subset, ground-truth references, normalization procedure, preprocessing parameters, and CPU execution environment provides a consistent basis for comparing the effects of the investigated preprocessing techniques across the two OCR systems.

4. Results and Discussion

The experimental results show that image preprocessing does not produce a uniform improvement in OCR performance. Instead, its effect varies according to both the OCR engine and the preprocessing technique. The evaluation was performed using Character Error Rate (CER) and Word Error Rate (WER), for which lower values indicate better recognition performance. Processing time was also considered to assess the computational cost of OCR inference. Table 1 summarizes the overall results obtained from the 200 SROIE receipt images under the original and three preprocessing conditions.

Table 1. Comparative OCR Performance Across Original and Preprocessed SROIE Receipt Images

OCR Engine	Preprocessing Technique	Mean CER (%)	Mean WER (%)	OCR Processing Time (s)
EasyOCR	Original (Baseline)	27.10	61.17	13.830
EasyOCR	Gaussian Blur	27.01	62.50	13.743
EasyOCR	CLAHE	26.23	58.38	22.971
EasyOCR	Fast NL-Means	27.46	62.11	25.944
Tesseract	Original (Baseline)	26.19	49.46	0.756
Tesseract	Gaussian Blur	26.33	48.54	0.640
Tesseract	CLAHE	26.79	50.59	0.903
Tesseract	Fast NL-Means	26.35	50.09	1.148

4.1 Overall OCR Performance

As shown in Table 1, the baseline performance of the two OCR engines differed before preprocessing was applied. EasyOCR obtained a mean CER of 27.10% and a mean WER of 61.17%, whereas Tesseract achieved lower error rates of 26.19% and 49.46%, respectively. Thus, under the experimental conditions of this study, Tesseract provided better baseline recognition performance than EasyOCR. The difference was particularly evident at the word level, indicating that Tesseract produced fewer word-level recognition errors on the evaluated receipt images.

The CER and WER comparisons are presented graphically in Figures 1 and 2. The figures show that the relative performance of the two OCR engines remained different across the evaluated preprocessing conditions, while individual preprocessing techniques affected the engines differently.

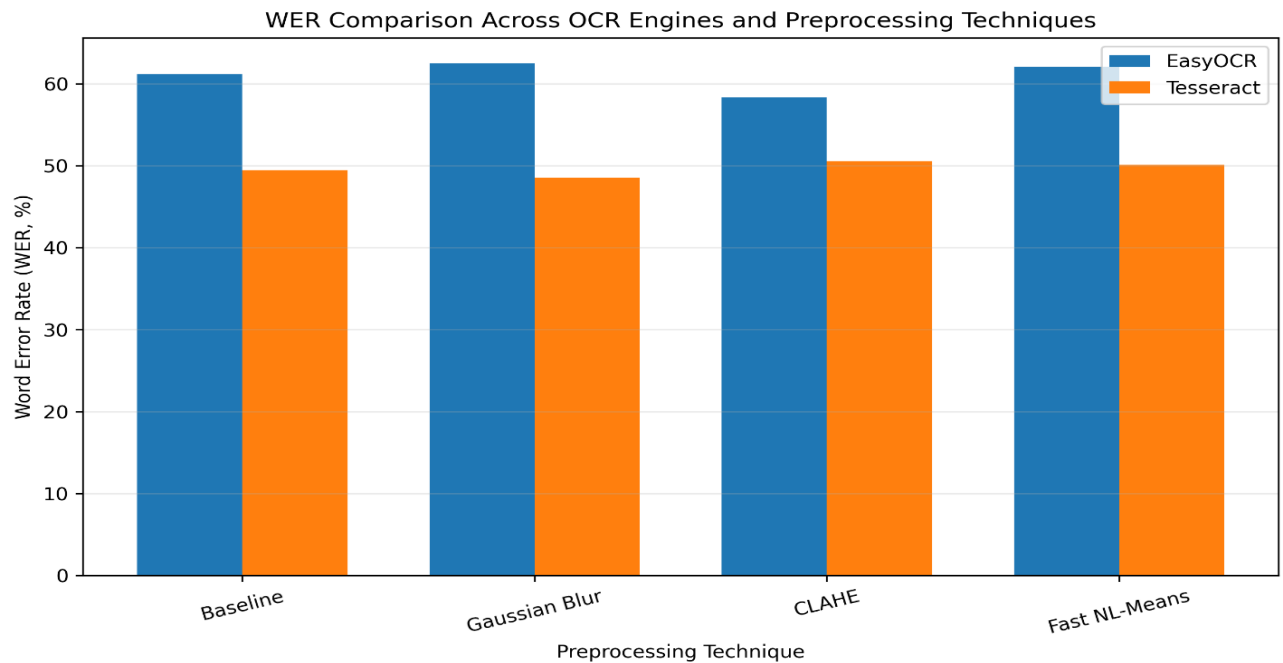


Figure 1. WER comparison across OCR engines and preprocessing techniques.

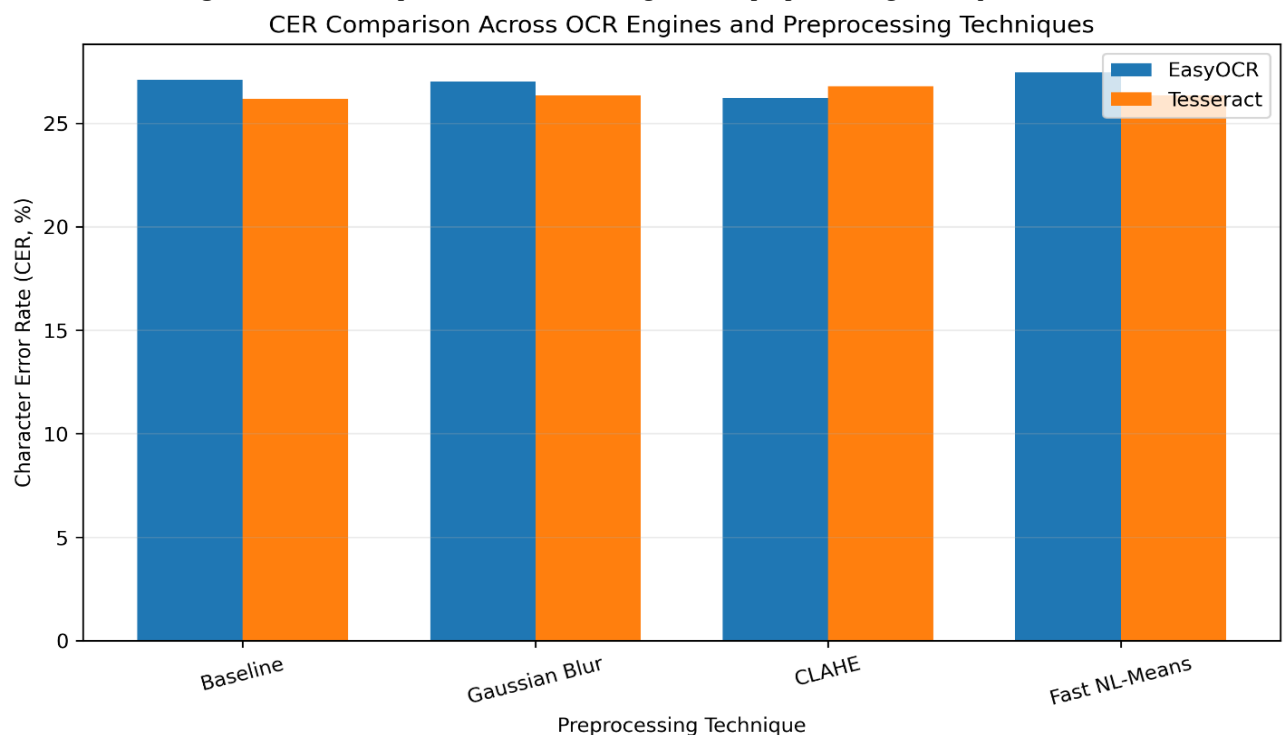


Figure 2. CER comparison across OCR engines and preprocessing techniques.

4.2 Effect of Gaussian Blur

Gaussian Blur produced different effects on the two OCR engines. For EasyOCR, the CER decreased slightly from 27.10% to 27.01%, while the WER increased from 61.17% to 62.50%. Therefore, the small improvement at the character level did not translate into improved word-level recognition.

For Tesseract, Gaussian Blur produced the lowest WER among all the tested Tesseract conditions, decreasing from 49.46% in the baseline to 48.54%. However, the CER increased slightly from 26.19% to 26.33%. This indicates that the effect of Gaussian Blur was metric-dependent as well as engine-dependent. Although it provided a small improvement in word-level recognition for Tesseract, it did not improve character-level recognition. The contrasting results suggest that Gaussian smoothing affected the two OCR systems differently. This difference highlights that the effect of preprocessing may vary depending on the OCR engine and the evaluation metric.

4.3 Effect of CLAHE

Among the evaluated preprocessing techniques, CLAHE produced the clearest improvement for EasyOCR. As shown in Table 1 and Figure 1, EasyOCR WER decreased from 61.17% in the baseline condition to 58.38% after CLAHE. This represents an absolute reduction of 2.79 percentage points, corresponding to approximately a 7.2% relative reduction in WER. CER also decreased from 27.10% to 26.23%.

The image-level analysis in Figure 3 provides additional information about the distribution of this improvement across the 200 images. For EasyOCR, WER improved in 129 images, deteriorated in 50 images, and remained unchanged in 21 images. Thus, although CLAHE improved the aggregate performance of EasyOCR, the improvement was not observed uniformly across all receipt images. This finding suggests that the effectiveness of local contrast enhancement may depend on the visual characteristics of individual receipt images.

The effect of CLAHE on Tesseract was different. Tesseract CER increased from 26.19% to 26.79%, while WER increased from 49.46% to 50.59%. The image-level analysis also shows that WER improved in 87 images but deteriorated in 101 images, with 12 images remaining unchanged. Therefore, both the aggregate results and the image-level results indicate that CLAHE did not provide an overall recognition benefit for Tesseract under the conditions of this study.

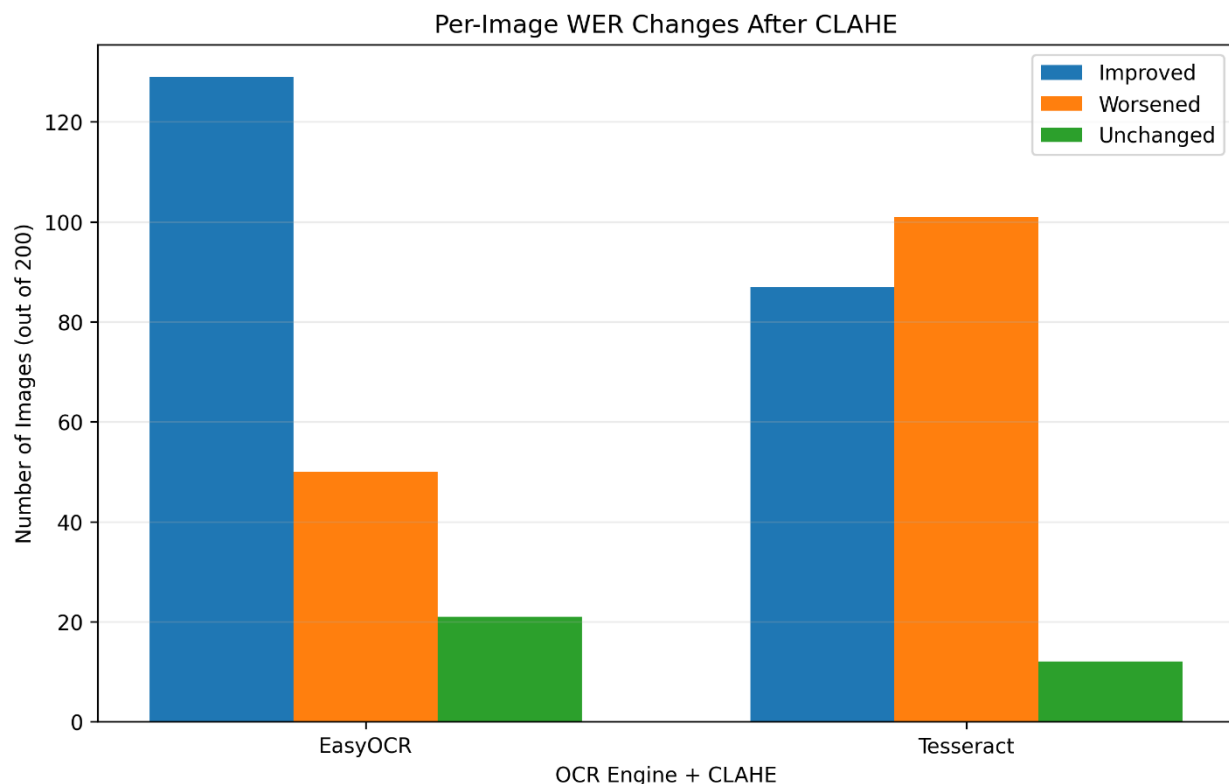


Figure 3. Per-image WER changes after applying CLAHE to the 200 receipt images.

The contrast between EasyOCR and Tesseract is particularly relevant because the same CLAHE operation was applied to the same set of receipt images. This provides direct empirical evidence that the effect of a preprocessing technique can vary between OCR engines. Consequently, the benefit observed for one OCR system should not automatically be assumed to apply to another system.

4.4 Effect of Fast NL-Means Denoising

Fast NL-Means did not provide an overall improvement for either OCR engine. For EasyOCR, CER increased from 27.10% to 27.46%, while WER increased from 61.17% to 62.11%. Tesseract also showed increases in both measures, with CER increasing from 26.19% to 26.35% and WER increasing from 49.46% to 50.09%.

These results show that the denoising operation used in this study did not improve OCR recognition for the evaluated receipt images. For EasyOCR, both CER and WER increased after applying Fast NL-Means, while Tesseract also showed increases in both measures. Therefore, Fast NL-Means, with the parameters used in this study, did not provide an overall recognition benefit for either OCR engine.

4.5 Comparison Between OCR Engines

A comparison of the overall results reveals different optimal conditions for the two OCR engines. EasyOCR achieved its lowest CER and WER with CLAHE, obtaining values of 26.23% and 58.38%, respectively. In contrast, Tesseract achieved its lowest WER with Gaussian Blur at 48.54%, while its lowest CER was obtained under the original baseline condition at 26.19%.

Therefore, no single preprocessing technique was optimal across both OCR engines and both evaluation metrics. This is one of the main findings of the study. The results demonstrate that the effectiveness of preprocessing should be evaluated in relation to the OCR engine and the characteristics of the evaluated document images rather than assuming that a particular image enhancement technique will universally improve recognition.

It is also important to distinguish between character-level and word-level performance. For example, Gaussian Blur slightly improved EasyOCR CER but worsened its WER, while Gaussian Blur improved Tesseract WER but slightly worsened its CER. These differences demonstrate that evaluating OCR performance using a single metric may not fully capture the effect of preprocessing.

4.6 Processing Time

Processing time provides an additional perspective on the practical differences between the two OCR systems. As shown in Figure 4, Tesseract required substantially less OCR execution time than EasyOCR under the tested CPU-only environment. The measured times were generally below approximately 1.2 seconds for Tesseract, whereas EasyOCR required substantially longer execution times.

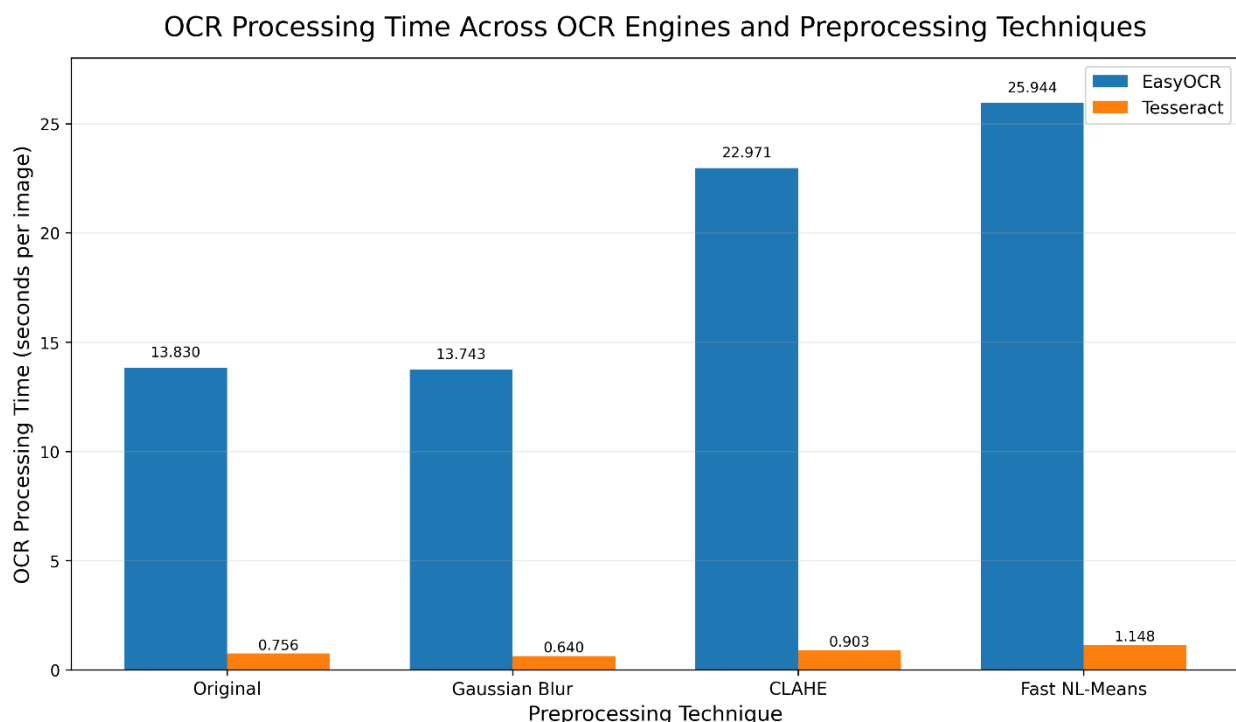


Figure 4. OCR processing time across OCR engines and preprocessing techniques.

The OCR processing times varied across the evaluated preprocessing conditions. For EasyOCR, the processing time ranged from 13.743 to 25.944 seconds per image, while Tesseract required between 0.640 and 1.148 seconds per image. The CLAHE condition required 22.971 seconds per image for EasyOCR and 0.903 seconds per image

for Tesseract. These measurements should be interpreted within the context of the experimental environment rather than as universal execution times, as processing time can be affected by hardware, software configuration, system load, and implementation details. Nevertheless, under the same experimental environment, the results indicate a substantial difference in OCR execution time between the two systems.

4.7 Overall Findings

Overall, the experimental results reveal three distinct patterns. First, CLAHE provided the clearest improvement for EasyOCR, reducing both CER and WER compared with the baseline. Second, Gaussian Blur produced the best word-level result for Tesseract, although its CER was slightly higher than the baseline. Third, Fast NL-Means did not improve the overall recognition performance of either OCR engine.

Taken together, these results suggest that image preprocessing does not provide a universal improvement in OCR performance. Instead, its effectiveness depends on the interaction between the preprocessing operation, the OCR engine, and the characteristics of the input images. The results therefore suggest that preprocessing should be selected empirically for the specific OCR system and document type rather than being treated as a mandatory enhancement step.

The results also highlight the value of using multiple evaluation measures. CER and WER captured different aspects of recognition performance, while processing time provided an additional measure of computational efficiency. Considering these measures together provides a broader assessment of the practical effects of preprocessing than relying on a single recognition metric.

5. Conclusion

This study investigated the effect of image preprocessing on OCR performance for English printed receipt images using two open-source OCR engines, Tesseract and EasyOCR. A controlled experimental evaluation was conducted on 200 SROIE receipt images under four image conditions: the original images, Gaussian Blur, CLAHE, and Fast NL-Means denoising. OCR performance was evaluated using Character Error Rate (CER), Word Error Rate (WER), and OCR processing time.

The results show that image preprocessing did not consistently improve OCR performance across the two systems. Its effect varied with both the OCR engine and the preprocessing technique. CLAHE produced the best overall recognition results for EasyOCR, reducing both CER and WER compared with the baseline. For Tesseract, Gaussian Blur produced the lowest WER, while the original images provided the lowest CER. In contrast, Fast NL-Means did not improve the overall recognition performance of either OCR engine under the tested conditions.

Character-level and word-level evaluation did not always lead to the same conclusions about the effect of preprocessing. In some cases, a preprocessing technique reduced CER while increasing WER, indicating that a single metric may not fully capture changes in recognition quality. Processing-time results further showed a substantial difference between the two OCR systems under the experimental environment, with Tesseract requiring considerably less OCR execution time than EasyOCR.

Within the conditions examined in this study, preprocessing was not consistently beneficial across the evaluated OCR systems. The choice of preprocessing technique should therefore take into account the OCR engine, document type, and characteristics of the input images. The results indicate that preprocessing decisions may need to be adapted to the OCR system rather than applying the same enhancement method to all systems.

6. Future Work

Future experiments could extend the dataset by using a larger number of receipt images and additional document types, allowing the observed preprocessing effects to be examined across a wider range of documents. The comparison could also be expanded to include additional OCR engines and preprocessing techniques, which would provide further evidence on how different OCR systems respond to image preprocessing.

Further experiments could investigate combinations or sequential applications of preprocessing techniques rather than evaluating each operation independently. Experiments using different languages, particularly Arabic receipt and financial documents, could also examine whether the observed effects are influenced by language and script characteristics.

Future studies could also investigate the relationship between image characteristics and OCR performance to determine whether specific types of degradation, illumination variation, or document quality are associated with greater benefits from particular preprocessing techniques. Such analysis could provide further insight into the conditions under which preprocessing is more effective for different OCR systems.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] D. Doermann and K. Tombre, Eds., *Handbook of Document Image Processing and Recognition*. London, U.K.: Springer, 2014, doi: 10.1007/978-0-85729-859-1.
- [2] Z. Huang, K. Chen, J. He, X. Bai, D. Karatzas, S. Lu, and C. V. Jawahar, "ICDAR2019 competition on scanned receipt OCR and information extraction," in *Proc. 15th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 2019, pp. 1516–1520, doi: 10.1109/ICDAR.2019.00244.
- [3] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-030-34372-9.
- [4] K. J. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, P. S. Heckbert, Ed. San Diego, CA, USA: Academic Press, 1994, pp. 474–485, doi: 10.1016/B978-0-12-336156-1.50061-6.
- [5] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2005, pp. 60–65, doi: 10.1109/CVPR.2005.38.
- [6] D. Sporici, E. Cuşnir, and C.-A. Boiangiu, "Improving the accuracy of Tesseract 4.0 OCR engine using convolution-based preprocessing," *Symmetry*, vol. 12, no. 5, Art. no. 715, 2020, doi: 10.3390/sym12050715.
- [7] G. Ramadhan, D. I. Mulyana, and S. Adrianto, "Optimization of Tesseract OCR for automatic text extraction on Indonesian ID cards (KTP) through image quality enhancement using preprocessing techniques," *International Journal of Software Engineering and Computer Science*, vol. 5, no. 3, pp. 938–948, 2025, doi: 10.35870/ijsecs.v5i3.5183.
- [8] K. A. Puteri, F. Alifia, P. A. Lailatulrahmi, G. P. Mindara, and E. P. Giri, "Perbandingan akurasi Tesseract dan EasyOCR sebelum dan sesudah prapemrosesan pada citra nota," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 1, 2026, doi: 10.23960/jitet.v14i1.8603.
- [9] R. Smith, "An overview of the Tesseract OCR engine," in *Proc. Ninth International Conference on Document Analysis and Recognition (ICDAR)*, 2007, pp. 629–633, doi: 10.1109/ICDAR.2007.4376991.
- [10] JaidedAI, "EasyOCR: Ready-to-use OCR with 80+ supported languages," GitHub repository, 2024. [Online]. Available: <https://github.com/JaidedAI/EasyOCR>
- [11] A. Rexhepi, E. Hasi, A. Haxholli, and E. Bytyçi, "Invoice and receipt optical character recognition: Review on current methods and future trends," in *Proc. RTA-CSIT*, 2025, pp. 107–116. [Online]. Available: <https://ceur-ws.org/Vol-4044/paper12.pdf>
- [12] R. I. Indrakusuma, A. S. Ahmadiyah, and N. F. Ariyani, "Pengenalan dan klasifikasi tulisan pada nota pembelian material (studi kasus proyek konstruksi)," *Jurnal Teknik ITS*, vol. 10, no. 2, pp. A478–A483, 2021, doi: 10.12962/j23373539.v10i2.77109.
- [13] K. Maliński and K. Okarma, "Analysis of image preprocessing and binarization methods for OCR-based detection and classification of electronic integrated circuit labeling," *Electronics*, vol. 12, no. 11, Art. no. 2449, 2023, doi: 10.3390/electronics12112449.
- [14] S. Sutriawan, A. Kurniawan, and R. Rosyid, "Improving the quality of optical character recognition (OCR) based on neural network with the image enhancement process," *Scientific Journal of Computer Science and Informatics*, vol. 1, no. 2, pp. 58–67, 2024, doi: 10.34304/scientific.v1i2.334.
- [15] S. Kumar, R. Jaiswal, and S. Kumar, "Improve OCR accuracy with advanced image preprocessing using machine learning with Python," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 7, pp. 1026–1030, 2020, doi: 10.35940/ijitee.G5745.059720.
- [16] M. Mayr, K. Neumeier, J. Krenz, S. Bürcky, F. Kordon, M. Seuret, J. Zöllner, F. Wu, A. Maier, and V. Christlein, "Data-efficient handwritten text recognition of diplomatic historical text," *Multimedia Tools and Applications*, vol. 84, pp. 39107–39122, 2025, doi: 10.1007/s11042-024-20545-9.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of **AJAPAS** and/or the editor(s). **AJAPAS** and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.