Proceeding of the International Conference on Electrical Engineering and Informatics
Volume. 1 Nomor. 2, 2024
e-ISSN : XXXX-XXXX, p-ISSN : XXXX-XXXX, Hal 17-33
Available Online at: https://prosiding.arteii.or.id/index.php/ICEEI

An Enhancement of Optical Character Recognition (OCR) Algorithm Applied in Translating Signages to Filipino

Shieryl E. Tendilla [1], Ivan Rey R. Dumago [2], Francis Arlando L. Atienza [3], Dan Michael A. Cortez [4]
[1,2,3,4] Computer Science Department, College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila, Manila City, Philippines,
Author correspondence: setendilla2021@plm.edu.ph

Abstract. Optical Character Recognition (OCR) systems often struggle to extract text accurately from images captured at various distances, particularly under challenging conditions such as blurriness, noise, or poor lighting. These issues are common in real-world scenarios and limit the effectiveness of existing OCR technologies. This study addresses these challenges by applying Gaussian blur after the grayscale conversion. This method reduces noise for the image's clarity without sacrificing the original algorithm's key features. Results revealed that the enhanced OCR algorithm significantly outperformed existing methods in terms of accuracy and confidence levels. It demonstrated the ability to read signages with higher precision, even in difficult conditions such as intricate designs, poor lighting, and long distances. This advancement enables more reliable text recognition and translation, offering practical applications for public signage translation, cross-cultural communication, and improved accessibility in multilingual environments.

Keywords Optical Character Recognition (OCR), Image Processing, Text Recognition, Computer Vision, Machine Learning.

1. INTRODUCTION

In an increasingly globalized world, understanding signs across languages has become vital for effective communication, especially in multicultural and tourist-heavy environments. Signages, which are visual displays of information such as text, symbols, and images, play a critical role as navigational and informational tools. They are designed to convey essential information, guide individuals through complex environments, and provide critical instructions. However, language barriers often limit their accessibility for individuals who are not familiar with the local language (Battula Anoosh, 2024). These limitations create a gap in achieving reliable and efficient multilingual translations, which is crucial for promoting inclusivity and cross-cultural understanding.

Addressing these gaps presents challenges, particularly for existing Optical Character Recognition (OCR) systems, which struggle with accurately processing text from signages due to poor image quality and noise interference. These challenges highlight the need for advanced image preprocessing techniques, such as Gaussian blur. The Gaussian blur is a convolution technique used to smooth images by reducing sharpness and averaging pixel values in a neighborhood (Chauhan, 2018). Essential for noise reduction and image preprocessing in computer vision algorithms like OCR, it operates as a linear low-pass filter, calculating pixel values using the Gaussian function (Novak et al., 2012). Its

Received: October 12, 2024; Revised: November 18, 2024; Accepted: December 26, 2024; Published: December 30, 2024;

application lays the groundwork for improved OCR performance by addressing issues such as noisy images, a common obstacle in recognizing environmental text.

Building on these principles, this study aims to enhance OCR algorithms, focusing on translating signages into Filipino. OCR technology automates the conversion of text from images into readable formats, extending beyond static documents to include dynamic and environmental text. In real-world scenarios, OCR demonstrates its versatility across diverse fields, such as streamlining document verification in finance, digitizing patient records in healthcare, enabling identity verification in security, and optimizing shipment tracking in logistics and transportation. These applications underscore the potential of OCR to facilitate cross-linguistic understanding.

By integrating advanced preprocessing techniques, robust text extraction methods, and deep learning-based models, this study seeks to improve OCR accuracy and efficiency, particularly in noisy and complex environments. The adaptation of OCR technology to the linguistic nuances of Filipino ensures accurate translations that address local and regional needs.

The enhanced OCR algorithm will not only improve text recognition and translation but also foster more accessible public spaces, assist travelers and tourists, and contribute to the growing field of multilingual communication technologies. By overcoming current OCR limitations, this research bridges language barriers, promotes inclusivity, and paves the way for further innovations in text recognition and translation technologies.

2. LITERATURE REVIEW

The advancement of Optical Character Recognition (OCR) technology has garnered significant attention due to its transformative impact on multilingual communication, particularly in translating signages. Despite its potential, several challenges hinder the effectiveness of OCR systems in real-world scenarios. This literature review critically evaluates and synthesizes previous studies to establish the background, motivation, and gaps that drive this research.

A. *Impact of Image Quality on OCR Performance*

A recurring challenge in OCR technology is its sensitivity to image quality. Studies by Garris et al. (n.d.) emphasize that low-quality images, often characterized by noise, distortions, or uneven lighting, significantly increase OCR error rates. While adaptive preprocessing techniques, such as dynamic thresholding, have improved accuracy, Garain et al. (2008) argue that the binarization process may still result in the loss of essential textual features. This highlights the need for advanced image enhancement methods tailored to address specific distortions in signages.

B. *OCR Struggles with Complex Layouts*

Complex layouts, such as multi-column text or mixed content, pose additional hurdles for OCR algorithms. Reisswig et al. (2020) note that dense, non-standard layouts require intricate segmentation techniques to preserve text order and readability. While deep learning models like LayoutLM address some of these issues, Zhu et al. (2022) caution that such methods may require large annotated datasets, limiting their scalability. These findings underscore the importance of robust layout analysis tailored for diverse signage formats.

C. *Challenges with Visually Similar Characters*

OCR algorithms often struggle to differentiate visually similar characters, such as "O" and "0" or "I" and "l," particularly in artistic fonts or handwritten text. Smith (2007) highlights the limitations of traditional shape-based recognition methods, which fail under noisy conditions. Yang et al. (2023) suggest that incorporating contextual language models can mitigate these errors, although such approaches demand computational resources that may not be feasible for real-time applications.

D. *Recognition of Artistic and Decorative Fonts*

The diversity of artistic and decorative fonts remains a significant obstacle for OCR systems. Clausner et al. (2019) observe that even state-of-the-art algorithms underperform when faced with fonts outside their training datasets. Similarly, Pirker and Wurzinger (2016) highlight that historical or unconventional fonts often require manual intervention or retraining to achieve acceptable accuracy. This indicates a gap in developing adaptive OCR systems capable of handling font variability.

E. *Speed and Accuracy Trade-offs in Real-Time Applications*

Real-time OCR applications demand a balance between speed and accuracy. Neumann and Matas (2012) assert that computationally intensive methods, while accurate, are unsuitable for time-sensitive tasks like augmented reality translations. Fragoso et al. (2011) emphasize that robust algorithms must address issues such as false detections and character misrecognitions without compromising efficiency. This highlights the need for streamlined, optimized approaches for real-time OCR deployment.

F. *Gaps and Opportunities*

Although recent advancements in OCR technology have addressed some of these challenges, several gaps remain. Most existing solutions are context-specific, focusing on either document OCR or limited real-world applications. Additionally, the lack of integration between advanced preprocessing techniques, contextual language models, and real-time processing capabilities restricts the scalability of OCR systems. Furthermore, the unique linguistic features of Filipino—such as syntax, morphology, and regional variations—have not been adequately explored in OCR research, creating an opportunity for this study to bridge the gap.

By addressing these gaps, this research seeks to develop an enhanced OCR algorithm that combines advanced image preprocessing, contextual recognition models, and efficient real-time capabilities. These improvements aim to provide accurate, scalable, and linguistically adaptive solutions for translating English signages into Filipino, thereby advancing the field of OCR and promoting cross-cultural accessibility.

3. METHODS

This section details the methodology employed to enhance the Optical Character Recognition (OCR) algorithm for translating signages to Filipino through pipelining EasyOCR using a novel approach of utilizing Gaussian blur during image preprocessing. This adjustment is designed to smooth images and reduce impulse noise while preserving the structural integrity of the image, crucial for reliable key feature extraction in character recognition.

A. *Data Acquisition*

The study utilized images of signages as primary data. These images were captured under various real-world conditions, including poor lighting, noise, and varying distances. Digital and smartphone devices were used for data collection to simulate user scenarios. Another dataset selected was ICDAR 2013 due to its standard use in benchmarking OCR systems.

Dataset	Source	Characteristics
CRAFT Dataset (ICDAR 2013 Dataset)	From github by masoudMZB from ICDAR 2013 OCR Dataset	It consist of 462 photos, including 229 for the training set and 233 for the test set.
TSDD - Traffic Sign Detection Dataset	From Kaggle by Larxel	This dataset contains 877 images of 4 distinct classes for the objective of road sign detection.
RVL-CDIP Dataset	Ryerson Vision Lab Complex Document Information Processing Collection	It consists of 400,000 grayscale images of documents across 16 categories, such as letters, forms, emails, and more.

Table 1: Text Extraction Dataset Composition

B. ***Preprocessing***

The preprocessing pipeline was developed to address issues such as noise, uneven illumination, and distortions to enhance text features without compromising key characteristics.

1. Gaussian Blur Algorithm

The Gaussian Blur algorithm is a widely used image processing technique to blur images, reducing noise and detail by applying a Gaussian function. The Gaussian Blur algorithm uses the Gaussian function, which is based on the normal distribution in statistics. In one dimension, the Gaussian function is given by:

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

In two dimensions, it is the product of two such Gaussian functions, one in each dimension:

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

```
Algorithm 1: Gaussian Blur

Input:
Output:
Allocate I'[width(I)][height(I)]  // Initialize output image
kernel = generate_kernel(K, sigma)
edgex ← ⌊K / 2⌋  // Calculate x-axis edge offset
edgey ← ⌊K / 2⌋  // Calculate y-axis edge offset
    for x ← edgex to width(I) - edgex do
      for y ← edgey to height(I) - edgey do
          sum ← 0
          for fx ← 0 to K - 1 do
              for fy ← 0 to K - 1 do
                  sum ← sum + I[x + fx - edgex][y + fy - edgey] *
kernel[fx][fy]
          I'[x][y] ← sum
```

Table 2: Gaussian Blur Algorithm

The gaussian blur (Algorithm 1) operates by the following steps:

1) Defining the Gaussian Kernel
2) Calculating Kernel Values
3) Defining the Sliding Window
4) Collecting Pixel Values Within the Window
5) Calculating the Weighted Average
6) Repeating Steps 4-5 for All Pixels.

2. Enhanced Optical Character Recognition Algorithm

Parameter	Description	Selected Value	Rationale
σ_gaussian	Gaussian blur standard deviation	1.5	Smoothes image without losing details
binarization_threshold	Binarization threshold	127	Separates text from background effectively
sharpening_factor	Image sharpening factor	1.2	Restores lost details due to blurring

Table 3: Optical Character Recognition Algorithm Parameters

The Gaussian blur standard deviation (σ_gaussian), set to 1.5, is used to smooth the image. After these smoothing steps, the image is converted into a binary format using a binarization threshold of 127. This threshold value separates the text from the background effectively, creating a clear distinction that aids the OCR algorithm in identifying characters accurately. However, the blurring process can sometimes result in the loss of fine details. To counter this, an image sharpening factor of 1.2 is applied. With these parameters established, we can now transition to the detailed steps of the Enhanced Optical Character Recognition (Algorithm 2), which systematically integrates a Gaussian blur and median filter to enhance noise resilience for accurate character detection.

```
Algorithm 2: Enhanced Optical Character Recognition

Input: I  // Input image
Output: text  // Recognized text

1: Allocate I'[width(I)][height(I)]  // Initialize output image

2: for x ← 0 to width(I) - 1 do
3:     for y ← 0 to height(I) - 1 do
4:         I'[x][y] ← convert_to_grayscale(I[x][y])

5: kernel = generate_gaussian_kernel(K, σ)
6: edgex ← ⌊K / 2⌋  // Calculate x-axis edge offset
7: edgey ← ⌊K / 2⌋  // Calculate y-axis edge offset
8: for x ← edgex to width(I) - edgex do
9:     for y ← edgey to height(I) - edgey do
10:        sum ← 0
11:        for fx ← 0 to K - 1 do
12:            for fy ← 0 to K - 1 do
13:                sum ← sum + I'[x + fx - edgex][y + fy - edgey] * kernel[fx][fy]
14:        I'[x][y] ← sum

15: threshold_value = 128  // Example threshold value
16: for x ← 0 to width(I) - 1 do
17:     for y ← 0 to height(I) - 1 do
18:         if I'[x][y] > threshold_value then
19:             I'[x][y] ← 255  // White
20:         else
21:             I'[x][y] ← 0   // Black

22: contours = detect_contours(I')
23: results = []
24: for contour in contours do
25:     x, y, w, h = bounding_rect(contour)
26:     if w > 20 and h > 20 then  // Skip small regions
27:         roi = extract_roi(I', x, y, w, h)
28:         text = ocr_engine.recognize(roi)
29:         results.append((x, y, w, h, text))

30: return results
```

Table 4: Enhanced Optical Character Recognition Algorithm

The enhanced Optical character recognition (Algorithm 2) operates through 5 stages:

1. Image preprocessing:
   - Allocate the output image and convert the input image to grayscale.
   - Generate a Gaussian kernel and apply it to the grayscale image to reduce noise.

2. Binarization:
   - Apply thresholding to convert the blurred grayscale image into a binary image.
3. Text Region Detection:
   - Detect contours in the binary image and filter out small regions.
4. Region of Interest (ROI) Extraction and OCR:
   - Extract ROIs from detected contours and use an EasyOCR engine to recognize text within these regions.
5. Output Generation:
   - Return the recognized text along with the coordinates of the text regions.

3. Integration with EasyOCR

The enhanced preprocessing pipeline was integrated with EasyOCR to assess its impact on recognition performance, particularly for multilingual OCR. The original EasyOCR pipeline without preprocessing served as the baseline for comparison.

4. Computational Environment

The enhanced Optical Character Recognition algorithm was implemented using Python 3.12.2, OpenCV for image processing, and EasyOCR for multilingual OCR recognition on an Intel Core i7 CPU with 16GB RAM. The development was conducted in Visual Studio Code version 1.96.

C. *Testing and Evaluation*

The study conducted a comparative analysis between the original EasyOCR and the enhanced OCR algorithm using the standard metrics: **Character Error Rate (CER)**, **Word Error Rate (WER)**, and **Analysis of Variance (ANOVA)**. The enhanced algorithm's ability to translate text into Filipino was also assessed to ensure linguistic and contextual accuracy.

1. **Character Error Rate (CER) -** The CER measured the accuracy of character recognition in the OCR output. It is defined as the ratio of the number of incorrect characters to the total number of characters in the ground truth text.

   Formula:

$$CER = \frac{I + S + D}{N}$$

Where I represents the number of insertions (extra characters in the OCR output), S represents the number of substitutions (incorrect characters), D represents the number of deletions (missing characters), and N represents the total number of characters in the ground truth text. A lower CER indicates better performance, with a perfect score of 0%.=

2. **Word Error Rate (WER) -** The Word Error Rate (WER) measures the percentage of words that contain one or more inaccurate characters.

Formula:

$$WER = \frac{S + D + I}{N} = 100\%$$

Where S represents the number of word substitutions (words replaced by incorrect ones), D represents the number of word deletions (words missing in the OCR output),I represents the number of word insertions (extra words in the OCR output), and N represents the total number of words in the ground truth text. A lower WER also indicates better performance, with a perfect score being 0%.

3. **Analysis of Variance (ANOVA) -** a statistical test employed to compare the means of three or more groups. In this paper, ANOVA was used to compare the CER and WER of both the original and enhanced algorithms.

To understand the mathematical formula for ANOVA calculation, specifically for one-way ANOVA, it is important to grasp the following terms:

1) Sum of Squares (SS)

- *Between-Groups Sum of Squares (SSB):* Measures variability between the group means.
- *Within-Groups Sum of Squares (SSW):* Measures variability within each group.
- *Total Sum of Squares (SST):* Measures total variability, considering all groups as a single sample.

2) Degrees of Freedom (df)

- *Between-Groups Degrees of Freedom (dfB):* Number of groups minus 1: $df_B = k - 1$, where *k* is the number of groups.
- *Within-Groups Degrees of Freedom (dfW):* Total number of observations minus the number of groups: $df_w = N - k$, where *N* is the total number of observations.

3) Mean Squares (MS)

- *Between-Groups Mean Squares (MSB):* Calculated by dividing the *SSB* by its degrees of freedom.
- *Within-Groups Mean Squares (MSW):* Calculated by dividing the *SSW* by its degrees of freedom.

The F value in ANOVA is calculated as the ratio between the variability between groups and the variability within groups:

$$F = \frac{MSB}{MSW}$$

To calculate the F value, we follow these steps:

a) Calculate the Total Sum of Squares (SST):

- SST is the sum of the variation of each value from the overall mean.

$$SST = \sum_{i=1}^{N} \quad (X_i - \underline{x})^2$$

Where X_i are the observed values and $\bar{x}$ is the overall mean.

b) Calculate the Between-Groups Sum of Squares (SSB):

- SSB measures the variation of the group means from the overall mean.

$$SSB = \sum_{j=1}^{k} \quad n_j(X_j - \underline{x})^2$$

Where n_j is the number of observations in group *j*, $\bar{x}_j$ is the group *j* mean, and $\bar{x}$ is the overall mean.

c) Calculate the Within-Groups Sum of Squares (SSW):

- SSW measures the variation of individual values from their group mean.

$$SSW = \sum_{j=1}^{k} \quad \sum_{i=1}^{n_j} \quad (X_{ij} - \underline{x}_j)^2$$

Where Xi_j are the observations in group *j*, and $\bar{x}_j$ is the group *j* mean.

d) Calculate the Mean Squares:

- Between-Groups Mean Squares (MSB)

$$MSB = \frac{SSB}{df_w} = \frac{SSW}{N-k}$$

- Within-Groups Mean Squares (MSW)

$$MSW = \frac{SSW}{df_w} = \frac{SSW}{N-k}$$

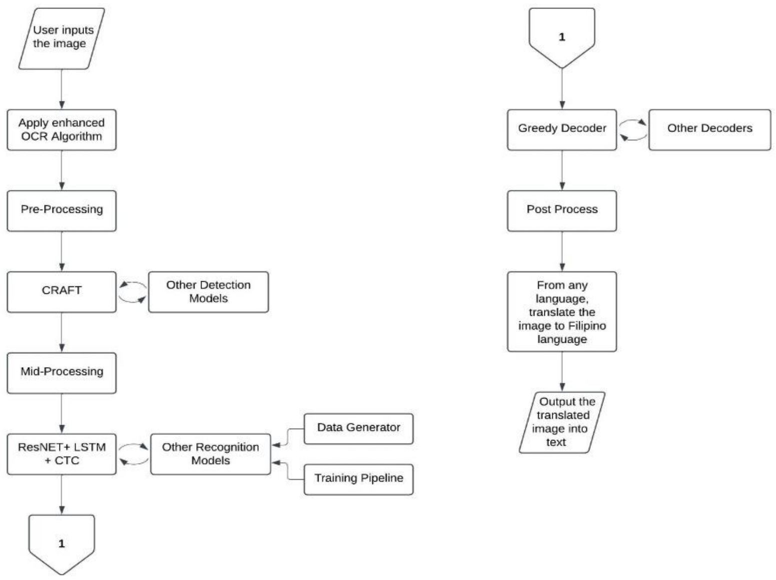
e) *Calculate the F Value:*

- Finally, the F value is the ratio between *MSB* and *MSW*:

$$F = \frac{MSB}{MSW}$$

A high F value indicates that the variability between group means (between-group variability) is greater than the variability within groups, suggesting a significant difference between the groups. A low F value suggests that the variability within groups is greater or comparable to the variability between groups, indicating that the group means are not significantly different. In summary, the F statistic in ANOVA measures how different the group means are compared to the within-group variability. The formula involves comparing the variability between the groups (how much the group means differ from the overall mean) with the variability within the groups (how much the individual values differ from the group mean).

4. Signage Translator System



The Signage Translator to Filipino is an advanced system that converts textual content from signage images into the Filipino language, ensuring accessibility for local audiences. The process begins when a user inputs an image containing text. This input undergoes an enhanced Optical Character Recognition (OCR) algorithm, which detects and extracts text from the image. Pre-processing steps, such as noise reduction and image normalization, optimize the image for accurate analysis.

The CRAFT (Character Region Awareness for Text) algorithm identifies text regions in collaboration with other detection models to ensure precision. Once text regions are detected, mid-processing steps prepare the text for recognition using a combination of ResNet for feature extraction, LSTM for sequence understanding, and CTC (Connectionist Temporal Classification) for sequence alignment. This recognition process is continuously refined using a training pipeline and a data generator to improve accuracy.

The extracted text is then passed through a greedy decoder, with additional decoders available to enhance output optimization. Post-processing ensures the recognized text is clean and error-free before it is translated into Filipino. The final step outputs the translated text, making the system capable of accurately interpreting and localizing signage content for multilingual and inclusive environments.

D. *Ethical Considerations*

No sensitive or personal data were used in the study. The dataset consisted solely of public signages captured for research purposes, ensuring compliance with ethical standards.

4. RESULTS

The performance of the existing and enhanced Optical Character Recognition (OCR) was evaluated using the standard OCR system metrics: Character Error Rate (CER), Word Error Rate (WER), and Analysis of Variance (ANOVA), using the test datasets.

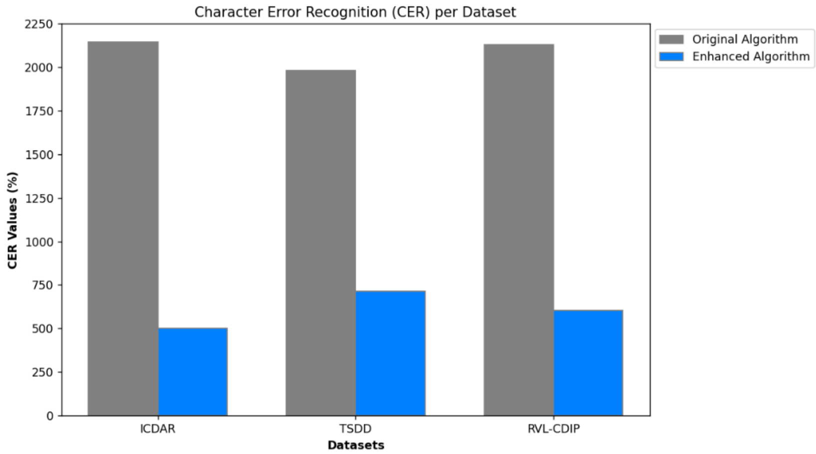


Figure 1: Character Error Recognition (CER) per Dataset

Figure 1 illustrates a substantial reduction in the Character Error Rate (CER) metric achieved by the enhanced Optical Character Recognition algorithm compared to the original implementation. Across the three test datasets, the original OCR algorithm exhibited CER values of 21.4285, 19.8277, and 21.3090. In contrast, the enhanced algorithm achieved significantly lower CER values of 7.0309, 7.1489, and 6.02837 for the ICDAE, TSDD, and RVL-CDP datasets, respectively.

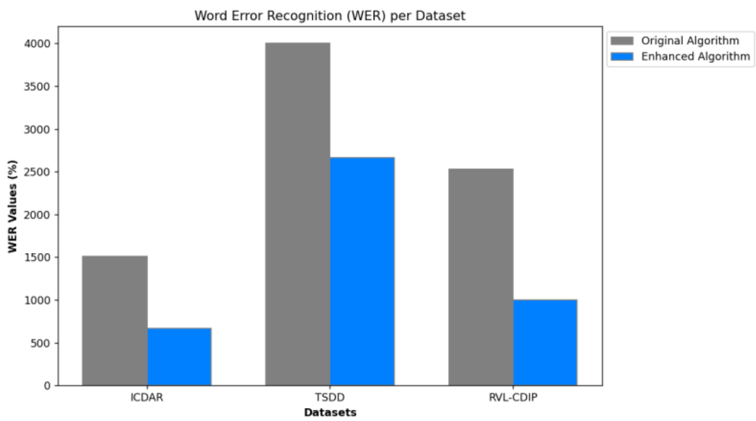


Figure 2: Word Error Recognition (WER) per Dataset

Figure 2 shows the Word Error Rate (WER) analysis, which indicates that the enhanced algorithm resulted in WER values of 6.656, 26.486, and 10.02837, compared to the original algorithm's WER values of 15.06599, 40.0, and 25.30904. This demonstrates a reduction in word-level errors when using the enhanced algorithm.

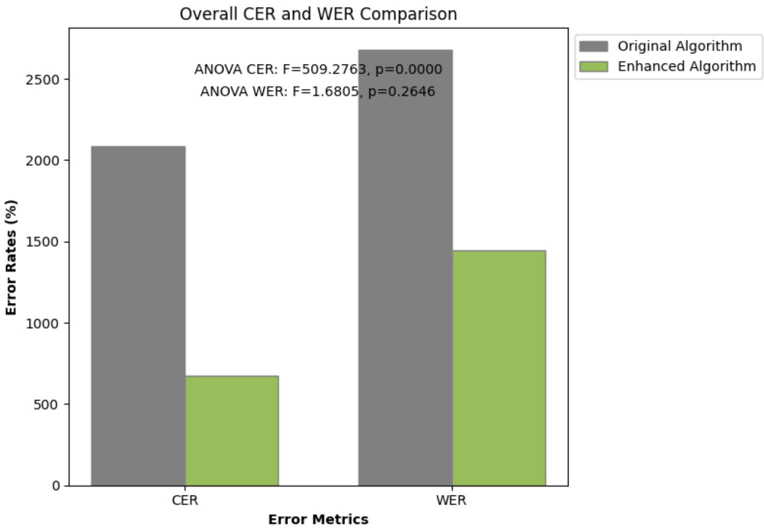


Figure 3: Analysis of Variance (ANOVA) based on the CER and WER average.

Figure 3 indicates a significant difference between the CER values of the original OCR algorithm and the enhanced OCR algorithm. The F-statistic was found to be 509.2763, with a p-value of 0.0000. This extremely low p-value suggests that the difference in CER between the two algorithms is statistically significant at any conventional significance level. Therefore, the null hypothesis that there is no difference in CER between the original and enhanced algorithms can be rejected. Conversely, the results for WER showed a different outcome. The F-statistic was 1.6805, with a p-value of 0.2646. Since the p-value is greater than the conventional significance level, the null hypothesis that there is no significant difference in WER between the original and enhanced OCR algorithms cannot be rejected.

5. DISCUSSION

The study aimed to evaluate the performance of an enhanced Optical Character Recognition (OCR) algorithm compared to its original implementation using standard OCR metrics, such as Character Error Rate (CER) and Word Error Rate (WER). The results reaffirm the importance of this study, demonstrating the enhanced algorithm's ability to significantly reduce errors, thereby contributing to the broader field of OCR technology improvement.

A summary of the findings reveals a substantial improvement in CER across all three test datasets (ICDAE, TSDD, and RVL-CDP). The enhanced algorithm achieved average CER values of 7.0309, 7.1489, and 6.02837, representing a marked reduction compared to the original algorithm's CER values of 21.4285, 19.8277, and 21.3090. Similarly, the WER analysis showed the enhanced algorithm’s average WER values as 6.666, 26.666, and 10.02837, significantly lower than the original algorithm’s values of 15.06599, 40.0, and 25.30904. These findings confirm the hypothesis that the enhanced algorithm would outperform the original in error reduction.

Relating these results to existing literature, the improvements align with previous studies emphasizing the role of advanced OCR frameworks, such as the integration of neural networks like ResNet and LSTM. The observed reduction in errors highlights the utility of robust preprocessing, detection, and recognition models, as implemented in the enhanced algorithm.

From a practical perspective, these results demonstrate that the enhanced OCR algorithm is better suited for real-world applications requiring high accuracy in text detection and translation, such as multilingual signage translators. By reducing errors, this system has the potential to enhance accessibility and usability across diverse domains, including education, navigation, and public services.

6. CONCLUSION

The findings of this study highlight the enhanced OCR algorithm's superior performance in reducing both character-level and word-level errors compared to the original implementation. These improvements underscore the potential for deploying advanced OCR technologies in applications requiring high accuracy, particularly in signage translation systems.

7. LIMITATION

The study focuses on enhancing the Optical Character Recognition (OCR) algorithm for translating signages from various languages into Filipino. While the algorithm is designed to accommodate images from any language, its current scope emphasizes translations specifically into Filipino, excluding broader multilingual translation capabilities.

Additionally, the research centers on addressing challenges in the existing algorithm without considering technical issues or external factors unrelated to its functionality. These limitations may influence the algorithm's applicability in broader multilingual contexts or situations involving technical variations, such as poor image quality or complex text layouts.

However, the study's focus on Filipino translations enables a thorough examination of linguistic nuances and specific challenges. By refining the algorithm within this defined scope, the study provides a robust foundation for future research to expand its capabilities to additional languages and address external factors, ensuring a more comprehensive and versatile OCR solution.

8. REFERENCES

Clausner, C., Antonacopoulos, A., & Pletschacher, S. (2019). Efficient and effective OCR engine training. International Journal on Document Analysis and Recognition, 23(1), 73–88. https://doi.org/10.1007/s10032-019-00347-8

DocuClipper. (2024, December 3). What is the OCR accuracy and how it can be improved. Retrieved from https://www.docuclipper.com/blog/ocr-accuracy/

Fateh, A., Fateh, M., & Abolghasemi, V. (2023). Enhancing optical character recognition: Efficient techniques for document layout analysis and text line detection. Engineering Reports, 6(9). https://doi.org/10.1002/eng2.12832

Fragoso, V., Gauglitz, S., Zamora, S., Kleban, J., & Turk, M. (2011). TranslatAR: A mobile augmented reality translator. IEEE Winter Conference on Applications of Computer Vision (WACV). https://doi.org/10.1109/wacv.2011.5711545

Garain, U., Jain, A., Maity, A., & Chanda, B. (2008). Machine reading of camera-held low quality text images: An ICA-based image enhancement approach for improving OCR accuracy. Proceedings of the International Conference on Pattern Recognition. https://doi.org/10.1109/icpr.2008.4761840

Garris, M., Janet, S., & Klein, W. (n.d.). Impact of image quality in machine print optical character recognition. NIST. Retrieved from https://www.nist.gov/publications/impact-imagequality-machine-print-optical-character-recognition

Ifttt-User. (2023, July 17). Gaussian blurring — A gentle introduction. Towards AI. Retrieved from https://towardsai.net/p/l/gaussian-blurring-a-gentle-introduction

Neumann, L., & Matas, J. (2012). Real-time scene text localization and recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://ieeexplore.ieee.org/document/6248097

Patil, S., Varadarajan, V., Mahadevkar, S., Athawade, R., Maheshwari, L., Kumbhare, S., Garg, Y., Dharrao, D., Kamat, P., & Kotecha, K. (2022). Enhancing optical character recognition on images with mixed text using semantic segmentation. Journal of Sensor and Actuator Networks, 11(4), 63. https://doi.org/10.3390/jsan11040063

Pirker, J., & Wurzinger, G. (2016). Optical character recognition of old fonts: A case study. Graz University of Technology. Retrieved from https://graz.elsevierpure.com/en/publications/optical-character-recognition-of-oldfonts-a-case-study

Rao, R. (2024, November 15). Analysis and benchmarking of OCR accuracy for data extraction models. Retrieved from https://www.docsumo.com/blogs/ocr/accuracy

Reisswig, C., Katti, A., Spinaci, M., & Hohne, J. (2020). Chargrid-OCR: End-to-end trainable optical character recognition for printed documents using instance segmentation. arXiv. https://arxiv.org/abs/1912.13318

Smith, R. (2007). An overview of the Tesseract OCR engine. IEEE Conference on Document Analysis and Recognition. https://ieeexplore.ieee.org/document/4376991

Tavares, R. A. (2024). Comparison of image preprocessing techniques for vehicle license plate recognition using OCR: Performance and accuracy evaluation. arXiv. https://arxiv.org/pdf/2410.13622

Waltz, K., & Gonzales, A. (n.d.). Demystifying Gaussian blur. Retrieved from https://www.adobe.com/ph_en/creativecloud/photography/discover/gaussian-blur.html

Yang, X., Arora, A., Yu Jheng, S., & Dell, M. (2023). Quantifying character similarity with vision transformers. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP). https://aclanthology.org/2023.emnlp-main.863.pdf

Zhu, W., Sokhandan, N., Yang, G., Martin, S., & Sathyanarayana, S. (2022). DocBed: A multi-stage OCR solution for documents with complex layouts. arXiv. https://arxiv.org/abs/2202.0141