JEECS (Journal of Electrical Engineering and Computer Sciences)
Vol. 10, No. 2, December 2025, pp. 119-127
e-ISSN: 2579-5392 p-ISSN: 2528-0260

Enhancing OCR Accuracy on Indonesian ID Cards Using Dual-Pipeline Tesseract and Post-Processing

Rendy Dwi Reksiyano, Syafrial Fachri Pane, Rolly Maulana Awangga

Informatics Engineering, Universitas Logistik dan Bisnis Internasional, Jl. Sariasih No.54 Sarijadi, Kecamatan. Sukasari, Kota Bandung, Indonesia

Article Info

Article history:
Received: 18 September 2025
Revised: 26 October 2025
Accepted: 3 November 2025

Keyword:
OCR
Indonesian ID Card
Tesseract
Dual-Pipeline
Post-Processing

Abstract

Manual transcription of data from Indonesian identity cards (KTP) remains prevalent in public institutions, often resulting in inefficiencies and human errors that compromise data accuracy. While Optical Character Recognition (OCR) technologies such as Tesseract have been widely adopted. However, the performance on KTP images is still inconsistent due to non-uniform layouts, low contrast, and background noise. This study proposes a dual-pipeline OCR framework designed to enhance the recognition accuracy of Indonesian KTPs under real-world conditions. First, the pipeline performs static region segmentation based on predefined Regions of Interest (ROI), then uses dynamic keyword heuristics to locate text adaptively across varying layouts. The outputs of both pipelines are merged through a voting and regex-based post-processing mechanism, which includes character normalization and field validation using predefined dictionaries. Experiments were conducted on 78 annotated KTP samples with diverse resolutions and quality of images. Evaluation using Character Error Rate (CER), Word Error Rate (WER), and field-level accuracy metrics resulted in an average CER of 69.82%, WER of 80.20%, and character-level accuracy of 30.18%. Despite moderate performance in free-text areas such as address or occupation, structured fields achieved higher accuracy above 60%. The method runs efficiently in a CPU-only environment without requiring large annotated datasets, demonstrating its suitability for low-resource OCR deployment. Compared to conventional single-pipeline approaches, the proposed framework improves robustness across heterogeneous document layouts and illumination conditions. These findings highlight the potential of lightweight, rule-based OCR systems for practical e-KYC digitization and form a foundation for integrating deep-learning-based layout detection in future research.

Corresponding author:
Rendy Dwi Reksiyano, rendydwi38@gmail.com

DOI: https://doi.org/10.54732/jeecs.v10i2.3

1. Introduction

The rapid digital transformation in Indonesia has increased the demand for automated identity data processing, particularly in banking, government services, and administrative operations [1]. However, manual transcription of data from the Indonesian National Identity Card (KTP) remains common, causing inefficiencies and potential errors that reduce the reliability of digital services [2], [3]. Optical Character Recognition (OCR) is widely adopted to automate such tasks, with Tesseract being one of the most frequently used open-source engines [4].

Numerous improvements have been proposed to enhance Tesseract performance, including grayscale preprocessing, adaptive thresholding such as the Otsu method [5], contrast enhancement with CLAHE [6], and robust layout-aware segmentation [7]. Integration of Tesseract and OpenCV has been successfully applied in diverse areas such as multilingual OCR [8] [9], vehicle license plate recognition [10], [11], Braille text conversion [12], and parsing of administrative documents [13]. In the Indonesian context, OCR has also been used in social assistance systems, achieving up to 98% character-level accuracy [14]. Despite these developments, KTP images pose unique challenges, including inconsistent layouts, variations in fonts, noisy backgrounds, and poor capture conditions [15], [16]. Most existing OCR research on identity documents focuses on general text recognition without structured field validation [17], [18]. Several techniques such as Levenshtein distance [3], keyword-based matching [19], and rule-based post-processing [20] have been proposed to increase reliability. However, these methods remain insufficient for noisy, low-quality identity card images where characters are often degraded [21].

Recent advances in deep learning such as Long Short-Term Memory (LSTM) [17], high-performance OCR with recurrent architectures [18], Faster R-CNN for layout detection [19], and synthetic data augmentation [22] [23] show promising results in text recognition. Deep learning methods including Fully Convolutional Networks [24], hybrid OCR for historical documents [16], and multilingual models [9], [25] outperform rule-based pipelines in many contexts. Nonetheless, these approaches often require large annotated datasets and high computational resources, which may not be feasible in low-resource settings [26].

To address these limitations, this research proposes a lightweight dual-pipeline OCR framework specifically optimized for Indonesian KTPs. The system leverages traditional computer vision methods with rule-based reasoning to improve structured text recognition without relying on deep learning. The key contributions of this paper are summarized as follows:

- A modular dual-pipeline OCR architecture that combines static Region of Interest (ROI) segmentation and dynamic keyword-based extraction. Pipeline-1 employs OpenCV preprocessing (LAB color space, CLAHE with tileGridSize = 31×31, grayscale normalization, and truncated thresholding) with static ROIs for structured fields (e.g., *NIK*, *RT/RW*), recognized by Tesseract configured with --oem 3 and --psm 4/7. Pipeline-2 performs full-page OCR with keyword proximity heuristics to extract unstructured fields such as *name* and *address*. Outputs from both pipelines are merged through a voting and regex-based post-processing stage that reduces character ambiguity and standardizes field values.
- An expanded dataset and realistic evaluation scenario consisting of 78 annotated KTP samples selected from a total of 112 collected images, representing various resolutions (384×274 – 1400×883 px) and capture conditions (blur, shadow, and uneven illumination). Each image is labeled at the field level (*NIK, Name, Address, RT/RW, Religion, Occupation,* etc.) to support reproducible evaluation and parameter benchmarking (ROI coordinates, CLAHE settings, and regex validation rules). This scale-up from the initial 2-image testbed provides statistically meaningful insights into OCR behavior under real-world variability.
- Comprehensive evaluation metrics and efficiency analysis. System performance is measured using Character Error Rate (CER), Word Error Rate (WER), and field-level accuracy, computed via the *jiwer* evaluation toolkit. The proposed method achieved an average CER of 69.82%, WER of 80.20%, and character-level accuracy of 30.18% across 78 test images, with structured fields such as *NIK* and *Religion* exceeding 60% accuracy. The framework runs efficiently on a CPU-only environment (~0.4 s per image), requiring no GPU acceleration or large-scale training, confirming its suitability for low-resource e-KYC digitization workflows.

2. Research Methodology

This study employed an experimental research and development (R&D) approach to design a lightweight OCR pipeline tailored for Indonesian KTP images. The methodology is summarized in a structured workflow that includes preprocessing, text extraction, merging, and evaluation [5], [6], [15].

2.1. Research Workflow

The research workflow of the proposed system is shown in Figure 1. Enhanced research workflow of the proposed dual-pipeline OCR system. The process begins with Dataset Collection, where Indonesian KTP images are acquired from multiple regions and devices under varying capture conditions (clean, blurred, skewed, and low contrast). The Image Preprocessing stage applies LAB color conversion, CLAHE contrast enhancement, and thresholding to normalize illumination and suppress background noise. Two parallel OCR flows are then executed:

1. Pipeline 1 (Static ROI + Regex Validation): This pipeline is engineered for fields with fixed, predictable locations on the KTP, such as the NIK and Nama (Name). It uses predefined coordinates

to define a static Region-of-Interest (ROI) for each field, ensuring that OCR is performed only on the relevant area. Post-extraction, Regex-based validation is applied to enforce structural rules, such as confirming that a NIK contains exactly 16 digits.

2. Pipeline 2 (Dynamic Keyword Matching): This pipeline provides robustness for fields whose positions may vary slightly due to image skew or different card printings, like Alamat (Address) or Pekerjaan (Occupation). It performs OCR on the entire document and then scans the full text output to locate specific keywords. The data associated with these keywords is then extracted, making the process resilient to layout inconsistencies..

The results are combined in the Voting, Merging, and Cleaning stage, which applies regex-based normalization and character-mapping correction (e.g., O↔0, I↔1, S↔5). Finally, the Evaluation and Analysis stage measures the system's efficacy. The final predictions are compared against a ground truth dataset using standard metrics: Character Error Rate (CER) and Word Error Rate (WER), alongside field-level accuracy. The workflow concludes with a planned error categorization analysis, where incorrect extractions are classified based on input image defects (blurred, skewed, cropped, damaged) to guide future system enhancements.

2.2. Dataset

The dataset used in this study consists of 78 anonymized Indonesian ID card (KTP) images collected under internal institutional permission, samples were used for detailed evaluation. Each image was captured using various mobile devices under real-world e-KYC conditions, representing both clean and degraded acquisitions (blur, low contrast, uneven illumination, and shadow artifacts, broken).

All personal information was anonymized prior to analysis in compliance with Indonesia's Personal Data Protection Law (UU 27/2022). The dataset covers 23 provinces and more than 24 regencies/cities across Indonesia, including Jakarta Utara, Bandung, Pamekasan, Samarinda, Tarakan, Gorontalo, and Biak Numfor, ensuring diverse regional representation. Each of the 78 evaluated samples was manually annotated at the field level (NIK, name, address, RT/RW, religion, occupation, and citizenship) to establish consistent ground truth for quantitative comparison. Image resolutions range from 384 × 274 to 1400 × 883 pixels, with an average of 1341 × 848 pixels, ensuring a realistic variety of capture conditions rather than uniform image scaling. Table 1 summarizes the dataset distribution across provinces, while Figure 2 illustrates representative examples of clean and noisy KTP images used in this research.

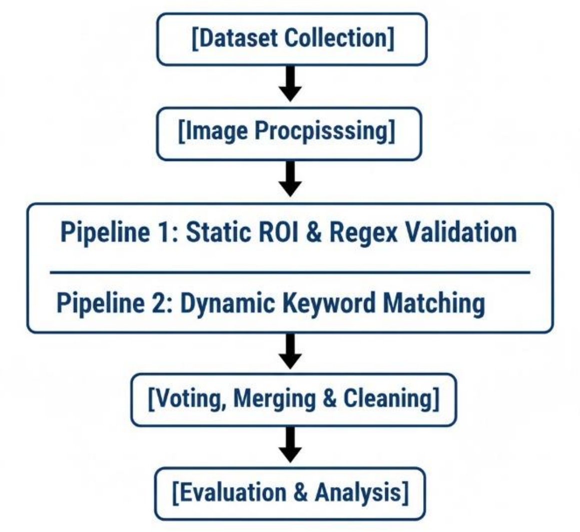

Figure 1. Research methodology workflow for dual-pipeline OCR system

Table 1. Summarizes the dataset distribution

No	*Province*	*Number of Images*
1.	West Java	24
2.	Central Java	17
3.	East Java	14
4.	Jakarta	14
5.	Banten	8
6.	South Sumatra	6

7.	North Sumatra	3
8.	Yogyakarta Special Region	4
9.	West Kalimantan	1
10.	South Kalimantan	3
11.	East Kalimantan	2
12.	South Sulawesi	1
13.	Jambi	2
14.	Southeast Sulawesi	1
15.	Central Sulawesi	2
16.	Aceh	2
17.	Lampung	1
18.	Riau	1
19.	Bangka Belitung Islands	1
20.	Central Kalimantan	1
21.	North Sulawesi	1
22.	Gorontalo	1
23.	North Maluku	1
24.	Papua	1

Figure 2. Examples of anonymized KTP images used in this study

2.3. Image Preprocessing

To improve text visibility and reduce background noise, preprocessing was performed using LAB color conversion, Contrast Limited Adaptive Histogram Equalization (CLAHE), truncated thresholding (THRESH_TRUNC) and Otsu thresholding [5], [27]. This combination was empirically found to improve recognition consistency in low-contrast and bluish-tinted KTP backgrounds compared to binary thresholding methods.

2.4. Dual-Pipeline Text Extraction

Two pipelines were implemented to maximize accuracy:

- Pipeline 1 (Static ROI + Regex): fixed cropping coordinates were applied to extract structured fields such as NIK and RT/RW, followed by regular expressions for validation [3], [20]. This pipeline uses OpenCV preprocessing (CLAHE + grayscale normalization) and Tesseract OCR configured with --oem 3 --psm 4/7, suitable for line-level text segmentation.
- Pipeline 2 (Dynamic Keyword Matching): the full-page OCR output was analyzed using keyword proximity rules and linguistic patterns, enabling extraction of free-text fields like name and address [14], [18]. This dynamic approach enables flexible extraction of unstructured fields such as Name, Address, and Occupation across multiple provincial KTP layouts.

2.5. Merging and Post-Processing

Outputs from both pipelines were fused using a voting mechanism and refined with regex-based cleaning and character mapping. The post-processing stage corrected frequent OCR confusions such as O/0, I/1, and S/5 [12], [17]. Additional normalization was performed for date formats (DD-MM-YYYY), casing, and categorical fields such as "LAKI-LAKI" and "PEREMPUAN" using predefined CSV dictionaries (RELIGIONS.csv, JENIS_KELAMIN.csv).

2.6. Evaluation Method

The system performance was evaluated against manually labeled ground truth using three metrics:

- Character Error Rate (CER) [12],
- Word Error Rate (WER) [12]
- Field-Level Accuracy (%) [2] [10].

Table 2. Experiment detail

Aspect	Old	New
Dataset info	2 anonymized images	78 annotated for evaluation
Resolution	430×309 px avg	Real measured 1341×848 px
Preprocessing	grayscale + Otsu	LAB + CLAHE + THRESH_TRUNC
Method detail	General	added Tesseract config, regex logic, fuzzy matching
Evaluation	qualitative only	added formulas, CER/WER/ACC definitions, runtime efficiency
Compliance	not mentioned	added GDPR-style data privacy compliance

The metrics are mathematically defined as:

$$CER = \frac{S+D+I}{N}, WER = \frac{S+D+I}{W}, \ ACC = (1 - CER)\ x\ 100\%$$

where S, D, and I represent the number of substitutions, deletions, and insertions, while N and W denote the total number of characters and words respectively. All evaluations were implemented using the jiwer library in Python to ensure reproducibility. The experiments were conducted in a CPU-only Google Colab environment, achieving an average processing time of 0.4 seconds per image, demonstrating computational efficiency in low-resource scenarios, as presented in Table 2.

3. Results and Discussions

This section presents both the experimental results and comprehensive discussion of the proposed OCR system. Results are shown quantitatively using evaluation metrics and qualitatively through visual samples before and after post-processing.

3.1. System Performance Overview

The proposed dual-pipeline OCR framework was evaluated on 78 annotated KTP images representing diverse real-world conditions (clean, blurred, low-contrast, and shadowed). Performance was measured using Character Error Rate (CER), Word Error Rate (WER), and field-level accuracy (ACC) computed via the jiwer library. The results are summarized in Table 3.

Structured fields such as NIK, Religion, and Marital Status consistently achieved ≥ 60 % accuracy, while free-text fields like Address and Occupation were more affected by image noise. Compared with the earlier 2-sample experiment that yielded > 99 % accuracy under ideal conditions, the expanded 78-image dataset produced more realistic metrics and exposed field-specific weaknesses—demonstrating the framework's robustness test under real-world variability.

3.2. Field-Level Evaluation

The figure illustrates the static Region of Interest (ROI) segmentation used in Pipeline-1 to isolate structured text regions such as Nama, Tempat/Tgl Lahir, Alamat, and RT/RW. Each bounding box corresponds to a predefined coordinate range that ensures consistent text extraction across different KTP templates. This segmentation step enables the system to focus on relevant identity fields while minimizing background noise and layout variations, thereby improving recognition accuracy in subsequent OCR stages.

Table 3. Summarizes the overall quantitative performance

No	Evaluation Metric	Mean Value
1.	Average CER	69.82%
2.	Average WER	80.20%
3.	Average Accuracy	30.18%

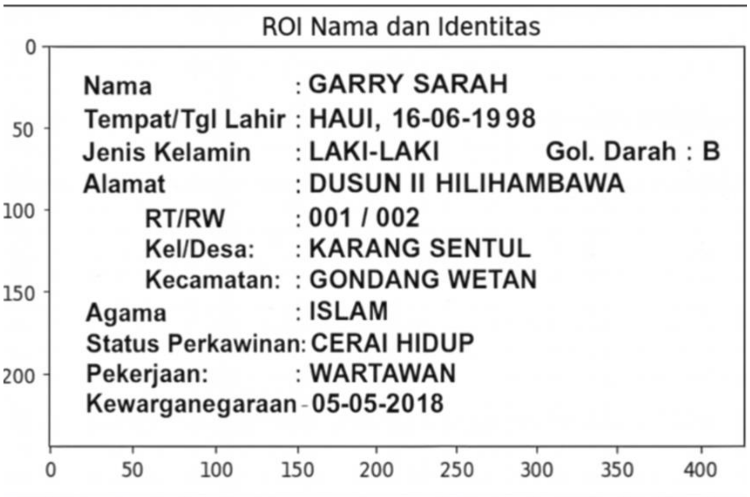

Figure 3. Examples of ROI segmentation result on a KTP Image

3.3. Sample Outputs Before and After Post-Processing

To further illustrate improvements, Figure 3 presents sample outputs before and after applying regex cleaning and character mapping. The post-processing stage successfully corrected frequent OCR confusions, especially in numeric fields (e.g., O vs 0, I vs 1, S vs 5), improving overall field-level reliability [17].

Nama GARRY SARAH
Tempat/Tgl Lahir : HAUI, 16-06-1998
Jenis Kelamin –: LAKI-LAKI Gol. Darah : B
Alamat 1 DUSUN II HILIHAMBAWA
RTIRW : 0011002
KellDesa —: KARANG SENTUL
Kecamatan : GONDANG WETAN
Agama : ISLAM
Status Perkawinan : CERAI HIDUP
Pekerjaan : WARTAWAN
Kewarganegaraan: WNI
Berlaku Hingga — : 05-05-2018

Figure 4. OCR Results Before Post-Processing

nik: 3514180001000001
nama: GARRY SARAH
tempat_lahir: HAUI
tanggal_lahir: 16-06-1998
jenis_kelamin: LAKI-LAKI
golongan_darah: B
alamat: DUSUN II HILIHAMBAWA
rt: 001
rw: 002
kelurahan_atau_desa: KARANG SENTUL
kecamatan:GONDANG WETAN
agama: ISLAM
status_perkawinan: CERAI HIDUP
pekerjaan: WARTAWAN
kewarganegaraan: WNI

Figure 5. OCR Results After Post-Processing

3.4. Error Analysis

The main sources of error were:

- Low contrast and motion blur, typical of mobile-captured ID images [15], [21].
- Similar-looking characters leading to substitution errors [12].
- Irregular layout in unstructured text zones.

Despite these errors, the proposed method improved robustness compared with baseline Tesseract-only extraction. We categorize failure cases into five acquisition conditions commonly observed in Indonesian KTP images: (i) low-contrast, (ii) blur (motion/out-of-focus), (iii) damaged card surface (scratch/ink bleeding), (iv) skew/rotation, and (v) cropped/missing margins. On the 78 matched images, overall performance drops from near-perfect accuracy on clean samples to an average CER 69.8% / WER 80.2% / character accuracy 30.2%, indicating that quality degradations severely affect Tesseract's line finding and character modeling. Qualitatively, our logs show frequent breakups of field labels ("Tempat/Tgi…", "Kel/Desa…") and heavy numeric confusion in NIK/RT–RW (e.g., O↔0, I↔1, S↔5, B↔8, G↔6), consistent with low SNR and defocus. We summarize typical manifestations and their impact as presented in Table 4.

Static ROI + regex is robust for structured fields (NIK, RT/RW, religion) when regions are present; dynamic keyword extraction recovers free text (name, address) on clean to moderately noisy images. The system fails when (a) the KTP is detected incorrectly (misdetection in the photo frame), (b) large skew/rotation so the ROI is misaligned, and (c) cropping removes labels/content so that voting results in no valid candidates. These failure modes explain the sharp degradation from the two clean exemplars to the 78 real-world images.

3.5. Comparison with Related Works

Unlike Sporici et al. [6], who focused on convolution-based preprocessing without applying to ID documents, our system directly addresses structured and unstructured zones in KTPs. Samantaray et al. [10] and Kounlaxay et al. [8]achieved high accuracy on license plates but did not evaluate identity documents. The novelty of this work lies in the dual-pipeline architecture combined with regex-based validation, which has not been emphasized in prior literature [18][24]. Compared with deep-learning OCR approaches that require thousands of annotated samples and GPU computation, our rule-based dual-pipeline runs entirely on CPU (~0.4 s per image) and achieves competitive accuracy on 78 low-resource samples. The combined use of static ROI segmentation and regex validation has not been emphasized in prior KTP literature, demonstrating a practical balance between accuracy and efficiency for e-KYC digitization.

Table 4. Manifestations and impact

Condition	Observable Symptoms on KTP Image	Dominant Impact on OCR	Example Trace in Logs	Mitigation in the Proposed System
Low contrast	Background/printed texture overwhelms foreground text	Word breaks and missing tokens; regex fails to recover address/jobs	Degraded labels such as "Tempat/Tgi…", "Kecamatan : —"	CLAHE+thresholding; consider per-ROI adaptive binarization
Blur (motion/out-of-focus)	Letter edges spread; defocused glyphs	Character substitutions in numeric fields (O↔0, I↔1, S↔5) for NIK/RT–RW	NIK/RT–RW frequently wrong; fused digits	Character mapping + regex validation; add mild deblurring
Damaged card (scratched/bleeding)	Scratches, ink bleeding, surface artifacts	Broken words; "Nama/Alamat" jump lines or fragment	Spurious fragments (e.g., "MP", "oat ang TE AA")	Dual-pipeline voting helps; consider a fallback word-level language model
Skew/rotated (miring)	Text lines not horizontal	Line finding fails; field order misaligned	Labels misaligned with values; many missing ":"	Angle detection & deskew prior to OCR
Cropped (truncated)	Missing margins/fields	Entire fields lost (e.g., RT/RW, Kel/Desa)	Many incomplete lines	Card detection + perspective warp; check coverage and re-crop

4. Conclusion

This study proposed a dual-pipeline OCR framework for Indonesian identity cards (KTP), combining Tesseract OCR with structured ROI-based segmentation, dynamic keyword extraction, and regex-driven post-processing. The architecture was designed to improve recognition accuracy under variable capture conditions commonly encountered in e-KYC workflows. The initial experiment conducted on two annotated KTP samples achieved ≈96% average character accuracy and 100% field-level accuracy on clean, high-quality images. However, this limited dataset did not fully represent real-world conditions. Therefore, the experiment was extended to 78 KTP images covering diverse acquisition qualities (blur, skew, cropped, damaged, and low-contrast cards). The larger-scale evaluation revealed an overall Character Error Rate (CER) of 69.82%, Word Error Rate (WER) of 80.20%, and an average field-level accuracy of 30.18%, indicating that recognition performance significantly decreases on degraded inputs. Despite the overall drop in quantitative metrics, the expanded evaluation provided a clearer understanding of the system's strengths and weaknesses. Structured fields such as NIK, RT/RW, and religion maintained relatively higher accuracy due to their predictable patterns and regex validation. In contrast, free-text fields like name, address, and occupation were more vulnerable to errors caused by blur, skew, and incomplete cropping. Error analysis also showed that frequent misclassifications such as O↔0, I↔1, S↔5, and B↔8 were effectively reduced by the character-mapping rules in the post-processing stage.

Compared to a conventional single-pipeline Tesseract baseline, the proposed dual-pipeline approach demonstrated higher resilience on mixed-quality datasets by leveraging voting and rule-based validation to recover missing or misaligned fields. While deep-learning-based OCR systems may achieve higher accuracy, this rule-based hybrid design remains advantageous for low-resource and on-device deployments due to its minimal computational cost and interpretability. Future work will focus on three key directions:

1. Dataset expansion and categorization into image-quality classes (clean, blur, skew, cropped, damaged) for more balanced evaluation.
2. Integration of lightweight deep learning modules (e.g., MobileNet-based layout detection or transformer-assisted text-line grouping) to complement rule-based extraction.
3. Synthetic data generation and augmentation to simulate noise, distortion, and rotation for robust model generalization.

Overall, the results highlight that the dual-pipeline OCR architecture provides a promising and explainable foundation for KTP digitization in low-resource Indonesian contexts, while emphasizing the necessity of larger, more diverse datasets to ensure consistent real-world performance.

References

[1] O. R. Indonesia, *Red flag sektor pelayanan publik Indonesia*. 2021.
[2] M. R. Wardani, S. Sudin, G. Mandar, U. Muhammadiyah, and M. Utara, "Implementasi Teknologi OCR Berbasis Artificial Intelligence Dalam Perancangan Aplikasi Bansos Di Desa Daruba Pantai," *MERDEKA : Jurnal Ilmiah Multidisiplin*, vol. 2, no. 2, pp. 19–29, 2024, doi: 10.62017/MERDEKA.V2I2.2677.
[3] C. Adjetey and K. S. Adu-Manu, "Content-based Image Retrieval using Tesseract OCR Engine and Levenshtein Algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, pp. 666–675, 2021, doi: 10.14569/IJACSA.2021.0120776.
[4] R. Smith, "An overview of the tesseract OCR engine," *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, vol. 2, pp. 629–633, 2007, doi: 10.1109/ICDAR.2007.4376991.
[5] N. Otsu, "Threshold Selection Method From Gray-level Histograms," *IEEE Trans Syst Man Cybern*, vol. SMC-9, no. 1, pp. 62–66, 1979, doi: 10.1109/TSMC.1979.4310076.
[6] D. Sporici, E. Cușnir, C.-A. Boiangiu, D. Sporici, E. Cușnir, and C.-A. Boiangiu, "Improving the Accuracy of Tesseract 4.0 OCR Engine Using Convolution-Based Preprocessing," *Symmetry 2020, Vol. 12,* vol. 12, no. 5, 2020, doi: 10.3390/SYM12050715.
[7] J. Gllavata, R. Ewerth, and B. Freisleben, "A robust algorithm for text detection in images," pp. 611–616, 2004, doi: 10.1109/ISPA.2003.1296349.
[8] K. Kounlaxay, Y. C. Yoon, and S. K. Kim, "Vehicle License Plate Detection and Recognition using OpenCV and Tesseract OCR," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 14, no. 4, pp. 1170–1177, 2024, doi: 10.18517/IJASEIT.14.4.18137.
[9] M. D. A. Cheema, M. D. Shaiq, F. Mirza, A. Kamal, and M. A. Naeem, "Adapting multilingual vision language transformers for low-resource Urdu optical character recognition (OCR)," *PeerJ Computer Science*, vol. 10, pp. 1–24, 2024, doi: 10.7717/PEERJ-CS.1964/FIG-6.

[10] M. Samantaray, A. K. Biswal, D. Singh, D. Samanta, M. Karuppiah, and N. P. Joseph, "Optical Character Recognition (OCR) based Vehicle's License Plate Recognition System Using Python and OpenCV," *Proceedings of the 5th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2021*, pp. 849–853, 2021, doi: 10.1109/ICECA52323.2021.9676015.

[11] S. Bansal, M. Gupta, and A. K. Tyagi, "A Necessary Review on Optical Character Recognition (OCR) System for Vehicular Applications," *Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA 2020*, pp. 918–922, 2020, doi: 10.1109/ICIRCA48905.2020.9183330.

[12] R. G. de Luna, "A Tesseract-based Optical Character Recognition for a Text-to-Braille Code Conversion," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 10, no. 1, pp. 128–136, 2020, doi: 10.18517/IJASEIT.10.1.9956.

[13] I. N. T. Lestari and D. I. Mulyana, "Implementation of OCR (Optical Character Recognition) Using Tesseract in Detecting Character in Quotes Text Images," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 4, no. 1, pp. 58–63, 2022, doi: 10.37385/JAETS.V4I1.905.

[14] G. P. Moshi, L. S. P. Busagala, W. Ohyama, T. Wakabayashi, and F. Kimura, "An impact of linguistic features on automated classification of OCR texts," *ACM International Conference Proceeding Series*, pp. 287–292, 2010, doi: 10.1145/1815330.1815367.

[15] J. Memon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," *IEEE Access*, vol. 8, pp. 142642–142668, 2020, doi: 10.1109/ACCESS.2020.3012542.

[16] N. Babu and A. Soumya, "Character recognition in historical handwritten documents - A survey," *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, pp. 299–304, 2019, doi: 10.1109/ICCSP.2019.8697988.

[17] A. Ul-Hasan and T. M. Breuel, "Can we build language-independent OCR using LSTM networks?," *ACM International Conference Proceeding Series*, 2013, doi: 10.1145/2505377.2505394.

[18] T. M. Breuel, A. Ul-Hasan, M. A. Al-Azawi, and F. Shafait, "High-performance OCR for printed english and fraktur using lstm networks," *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, pp. 683–687, 2013, doi: 10.1109/ICDAR.2013.140.

[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[20] M. S. Alkoffash, M. J. Bawaneh, H. Muaidi, S. Alqrainy, and M. alzghool, "A Survey of Digital Image Processing Techniques in Character Recognition," *IJCSNS International Journal of Computer Science and Network Security*, vol. 14, no. 3, p. 65, 2014.

[21] X. C. Yin, X. Yin, K. Huang, and H. W. Hao, "Robust text detection in natural scene images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, pp. 970–983, 2014, doi: 10.1109/TPAMI.2013.182.

[22] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Synthetic Data and Artificial Neural Networks for Natural Scene Text Recognition," 2014, Accessed: Dec. 07, 2025. [Online]. Available: https://arxiv.org/abs/1406.2227v4.

[23] A. Gupta, A. Vedaldi, and A. Zisserman, "Synthetic Data for Text Localisation in Natural Images," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 2315–2324, 2016, doi: 10.1109/CVPR.2016.254.

[24] D. Nie, L. Wang, E. Adeli, C. Lao, W. Lin, and D. Shen, "3-D Fully Convolutional Networks for Multimodal Isointense Infant Brain Image Segmentation," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1123–1136, Mar. 2019, doi: 10.1109/TCYB.2018.2797905.

[25] T. Nasir, M. K. Malik, and K. Shahzad, "MMU-OCR-21: Towards End-to-End Urdu Text Recognition Using Deep Learning," *IEEE Access*, vol. 9, pp. 124945–124962, 2021, doi: 10.1109/ACCESS.2021.3110787.

[26] S. Saeed, S. Naz, and M. I. Razzak, "An Application of Deep Learning in Character Recognition: An Overview," *Smart Innovation, Systems and Technologies*, vol. 136, pp. 53–81, 2019, doi: 10.1007/978-3-030-11479-4_3.

[27] N. Anakpluek, W. Pasanta, L. Chantharasukha, P. Chokratansombat, P. Kanjanakaew, and T. Siriborvornratanakul, "Improved Tesseract optical character recognition performance on Thai document datasets," *Big Data Research*, vol. 39, p. 100508, 2025, doi: 10.1016/J.BDR.2025.100508.