Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 09 Issue 02
2025

Research Article
https://doi.org/10.56979/902/2025

Enhancing Handwritten Prescription Recognition with AI-Driven OCR

Haseeb Ullah[1*], Muhammad Tanveer[2], and Atif Jan[3]

[1]Department of Computer Systems Engineering, University of Engineering and Technology, Peshawar, 25000, Pakistan.
[2]Department of Computer Science, Preston University, Islamabad Campus, Islamabad, 44000, Pakistan.
[3]Department of Electrical Engineering, University of Engineering and Technology, Peshawar, 25000, Pakistan.
[*]Corresponding Author: Haseeb Ullah. Email: haseeb.ullah@uetpeshawar.edu.pk

Received: June 20, 2025 **Accepted:** August 03, 2025

Abstract: Accurate interpretation and understanding of medical prescriptions are crucial for healthcare providers to ensure suitable treatment for patients. However, the increasing number of prescriptions and the complexity of pharmaceutical regimens may lead to errors, which could have severe consequences. To overcome this problem, artificial intelligence (AI) can automate tasks such as identifying the correct medication, determining the correct dose, and checking for drug interactions. This makes prescription analysis more accurate and faster. This study presents an AI-driven optical character recognition (OCR) framework that uses TrOCR with Roboflow to convert handwritten prescriptions into a digital format. Our method achieves a Word Error Rate (WER) of 12.5%, a Character Error Rate (CER) of 8.7%, and an Exact Match Accuracy of 81.3%. These results show that the system can accurately transcribe prescriptions and help reduce medication errors, making healthcare workflows safer and more efficient.

Keywords: Deep Learning; TrOCR; Medical Prescription; Handwriting Recognition; Roboflow

1. Introduction

Optical character recognition (OCR) technology has completely transformed how computer systems handle the printed text, allowing businesses of all kinds to get information quickly, save documents, and process data. However, older OCR systems don't do as well with handwritten text, especially when the writing styles vary, the formatting is messy, or it's just hard to read [1]. This problem is particularly noticeable in healthcare. Even though there's been a big push to go digital, handwritten prescriptions are still widespread. Doctors' handwriting is unreadable and often challenging to decipher, which really gets in the way of accurately extracting information. This can lead to medication mistakes, reduced patient safety, and significant time and effort wasted in pharmacies and healthcare facilities [2]. Making prescriptions digital is absolutely vital for reducing errors and improving efficiency. Manual transcribing, besides being slow and requiring significant resources, is also prone to human error, and in a medical setting, those errors can have serious consequences. Misunderstanding patient instructions, dosages, or even drug names can lead to harmful reactions to medicine, extended hospital stays, and, tragically, even death. That's why e-prescribing systems have become such an essential tool for mitigating risks, thanks to their benefits, including increased accuracy, improved patient safety, more efficient workflows, and a decrease in adverse drug events [3]. When healthcare providers convert handwritten prescriptions into a structured, machine-readable format, they can ensure better clarity, reduce dispensing errors, and facilitate seamless interaction with pharmacy management systems and electronic health records (EHR).

To address the handwriting recognition problem in medical prescriptions, this study explores the combined use of Roboflow and (Transformer-based Optical Character Recognition) TrOCR. By employing Transformer models to convert textual information within images without the need for segmentation or other intermediate steps, TrOCR provides a direct approach to text recognition.

Advances in machine translation, text summarization, image interpretation, and text generation from visual data have been enabled by these networks' ability to understand sequences holistically rather than as discrete characters or words [4]. Initially, developing custom OCR solutions is simplified by Roboflow, a computer vision platform that streamlines dataset creation, annotation, maintenance, and model training. This study demonstrates that combining Roboflow's efficient data pipeline with TrOCR's advanced recognition capabilities offers a successful method for solving handwriting recognition issues in medical records.

2. Related Works

From early template-matching and rule-based systems like Tesseract to approaches based on statistical techniques and machine learning, optical character recognition (OCR) has advanced significantly. Due to baseline differences, noisy input, and variations in stroke width, these methods consistently performed poorly when applied to handwritten texts, but they performed well for structured, printed documents with consistent fonts [5]. Because of these limitations, these systems were insufficient for essential domains, such as prescription drugs, where errors could result in major clinical problems. Handwriting recognition advanced significantly with the advent of deep learning architectures, particularly Convolutional Recurrent Neural Networks (CRNNs). When compared to rule-based systems, CRNNs improve performance by combining CNNs and RNNs for feature extraction and sequence modelling, respectively [6]. However, their efficacy in real-life medical settings was constrained by their heavy reliance on line segmentation and preprocessing. Furthermore, rather than unstructured handwritten prescriptions, a large portion of the early research in medical OCR concentrated on structured data, such as forms, electronic health records, or simplified prescription fields [7]. As a result, these systems could usually extract drug names or dosages only within limited contexts, leaving the challenge of free-form handwriting largely unresolved. Google Translate. Unlike CRNNs, which face challenges with long-range dependencies and the variability in handwriting flow, TrOCR is designed to grasp global context and handle noisy, cursive scripts common issues in handwritten medical prescriptions. Recent benchmarking studies have shown its superiority over traditional OCR systems and CRNN-based models, particularly in clinical settings where prescription handwriting is highly unpredictable [8].

Furthermore, efforts have been made to address the multilingual challenge of clinical prescriptions, especially in regions with high linguistic diversity. Rani et al. developed a multilingual OCR system capable of reading and translating prescriptions written in regional languages similar to English [9]. This solution adapts to healthcare environments with many languages by combining a translation feature with advanced OCR technology. In healthcare settings, especially in multicultural areas where prescriptions often include multiple languages, the ability to analyze multilingual text is essential. In order to enhance text segmentation and recognition, Zitu et al. [10] investigated Transformers for OCR in handwritten prescriptions, utilising self-attention mechanisms. When it came to highly cursive and cluttered prescriptions, this model performed better than conventional RNN-based OCR systems. However, its real-time use was restricted by its computational complexity. By producing synthetic handwritten prescriptions, Sousa Neto et al. [11] suggested a GAN-based process to increase training datasets and improve OCR accuracy for uncommon or underrepresented words. Although this method increased the diversity of the data, it occasionally added noise that decreased the accuracy of recognition. Patel et al. [12] made a significant contribution by creating a graph-based OCR technique that uses graph neural networks (GNNs) to identify intricate spatial relationships in handwritten medical text. Their model performed well in identifying ambiguous characters, but it had trouble with fragmented words or extremely compact handwriting styles. Maleki Varnosfaderani et al. [13] used meta-learning techniques to further enhance OCR for prescriptions, enabling models to adjust to new handwriting styles with little retraining. Despite learning from a small amount of data, their system performed worse when handwriting from various practitioners was wildly inconsistent. To improve sequential text recognition in OCR, Kataria et al. [14] integrated CNN-based feature extraction with Hidden Markov Models (HMMs). This hybrid approach achieved strong results in denoising and segmenting complex scripts but required significant computational resources. Deep learning- driven advances in OCR technology have greatly improved the accuracy of handwritten prescription recognition. Challenges posed by complex handwriting, scientific terminology, and multilingual text have been addressed through transformer-based methods.

Preparing datasets is another difficulty in medical OCR. The majority of earlier systems used small, carefully selected datasets that were challenging to grow. By providing an integrated platform for dataset annotation, augmentation, and versioning, Roboflow addresses this issue and facilitates the creation of robust OCR pipelines. It helps build models that are resistant to the variability found in real-world prescriptions by introducing controlled variations like noise addition, rotations, or contrast changes. Furthermore, Roboflow Universe makes shared datasets and models accessible, promoting cooperation and reproducibility [15,16].

In summary, these technological developments highlight an ongoing knowledge gap: although deep learning has boosted the accuracy of OCR, few studies have effectively integrated innovative transformer models with modern data engineering pipelines to address the unique requirements of handwritten medical prescriptions. By combining Roboflow's dataset management and augmentation tools with TrOCR's end-to-end recognition capabilities, our approach aims to bridge this gap. In contrast to previous CNN- or CRNN-based systems, this integrated framework seeks to improve clinical reliability by lowering errors in drug name and dosage interpretation, where errors can seriously risk patient safety.

3. Research Methodology

The proposed method for digitizing handwritten medical prescriptions combines Roboflow's dataset management and preprocessing features with TrOCR's end-to-end transformer-based text recognition. The approach includes three key components: dataset preparation, model training and setup, and evaluation.

3.1. Dataset Preparation

To facilitate the training of a transformer-based OCR model, TrOCR, we used Roboflow as the primary tool for dataset preparation. For this study, we used the 580 images from the publicly available Doctors' Prescriptions Handwriting dataset hosted on Roboflow Universe, which provides diverse collections of handwritten prescription images. Multiple handwriting styles and different levels of document readability appear in these images and scans, presenting real-life document characteristics like poor image quality and diverse conditions and backgrounds. It offers annotated samples of handwritten prescriptions specifically designed for OCR research. According to a standard ratio, the data set is divided into 80% for training, 10% for validation, and 10% for testing. This division ensures that the model is trained on a substantial volume of data, validated during the training process to prevent overfitting, and tested on unseen data to assess its capacity for generalization.

3.2. Model Training and Experimental Setup

The annotated dataset was exported from Roboflow in a format compatible with the Hugging Face Transformers library. Each image was converted to PIL format and paired with its ground truth transcription. Input text was tokenized using the TrOCR Processor, which combines a ViT Feature Extractor for image encoding and a Roberta Tokenizer for target sequence encoding.

The model was initialized with pre-trained weights from microsoft/trocr-base-handwritten and fine-tuned on the prepared dataset using transfer learning. Training used cross-entropy loss and the Adam optimizer. The training configuration is summarized in Table 1.

Table 1. Training Hyperparameters

Parameter	Value
Batch Size	16
Learning Rate	2e-5
Epochs	8
Optimizer	Adam
Loss Function	Cross-Entropy
Early Stopping	Enabled (patience = 2 epochs)
Augmentations Applied	Rotation, Contrast, Blur, Noise

3.3. Evaluation

To thoroughly test how well our TrOCR-based system converts handwritten medical prescriptions into digital text, we used several common metrics typically found in optical character recognition and text data processing. The use of these metrics has enabled us to obtain a comprehensive picture of the model's accuracy at various levels of granularity. Metrics used for results are Exact Match Accuracy (EMA), Word Error Rate (WER), and Character Error Rate (CER).

3.3.1. Character Error Rate (CER)

Along with errors of types such as substitutions, insertions, or deletions, CER counts only those characters that are incorrectly recognized, and this number is expressed as a percentage of the total number of characters in the ground truth. Generally, the lower the CER, the better the character-level accuracy it represents [17].

$$CER = \frac{Deletions\ +\ Substitutions\ +\ Insertions}{Total\ Characters\ in\ Ground\ Truth}$$

3.3.2 Word Error Rate (WER)

WER is a similar metric, but it measures only the word level—the proportion of words incorrectly recognized. It considers word substitutions, insertions, and deletions. WER is primarily viewed as more helpful in evaluating the overall performance of the text recognition system [18].

$$WER = \frac{Deletions\ +\ Substitutions\ +\ Insertions}{Total\ Words\ in\ Ground\ Truth}$$

3.3.3 Exact Match Accuracy (EMA)

Exact Match Accuracy is calculated as the percentage of predictions that perfectly match the ground truth text out of all test samples. This metric shows the model's ability to produce complete and error-free transcriptions. Since Exact Match Accuracy is a strict measure, it only counts predictions that are exactly identical to the ground truth. For example, in digitizing medical prescriptions, even a minor error can lead to serious consequences. [19].

$$EM\ Accuracy = \frac{Number\ of\ exact\ matches}{Total\ number\ of\ samples} \times 100$$

4. Results

As reported in the abstract, these results demonstrate the effectiveness of the TrOCR model, which was trained on handwritten prescription datasets. The model achieves competitive accuracy, with a CER of 8.7% and WER of 12.5% at both the character and word levels. Notably, the model achieves an 81.3% Exact Match Accuracy, indicating that it can accurately transcribe a substantial number of prescriptions without errors. This is crucial for patient safety and maintaining accurate data in healthcare.

Table 2. Performance Table

Metric	Value
Character Error Rate	8.7%
Word Error Rate	12.5%
Exact Match Accuracy	81.3%

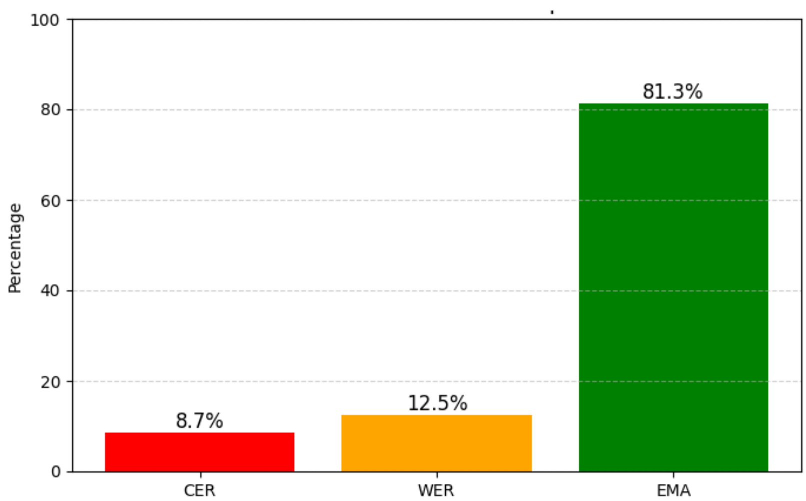

Figure 1. Model performance bar chart.

The roboflow, as shown in the annotated Figure 3, successfully identified and localized the three distinct lines of text. The confidence scores for the detected bounding boxes were consistently high: 93% for 'Petra 40mg. 5 days, 85% for 'Panadol 1+1+1', and 87% for 'Arinac 400mg/60mg 3 day'. These results demonstrate the model's high accuracy in segmenting and localizing relevant text from handwritten notes. the TrOCR was applied to the localized regions. The model achieved a perfect character-level accuracy, transcribing the entire prescription without error. The final concatenated output, Figure 4, was: 'Petra 40mg 5 days Panadol 1+1+1 arinac 400mg/60mg 3 day'. The model successfully recognized and transcribed all characters, numbers, and symbols, including the specific medical notations 'mg' and dosage instructions like '1+1+1' and '400mg/60mg'. This high-fidelity transcription indicates the robustness of the model to variations in handwriting and its ability to handle domain-specific content.

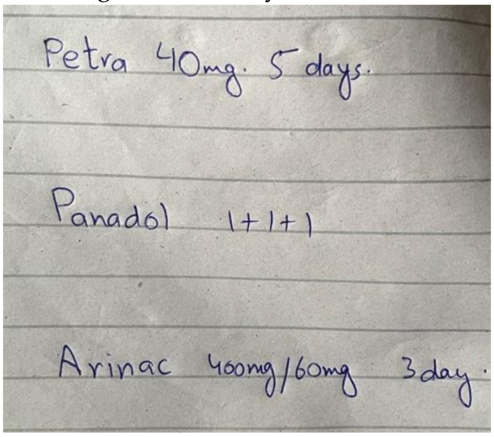

Figure 2. Input image 1

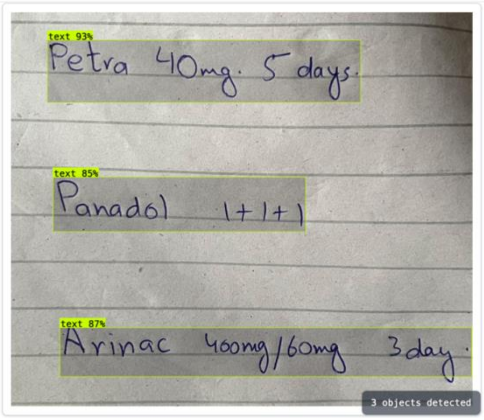

Figure 3. Annotated image 1 by Roboflow

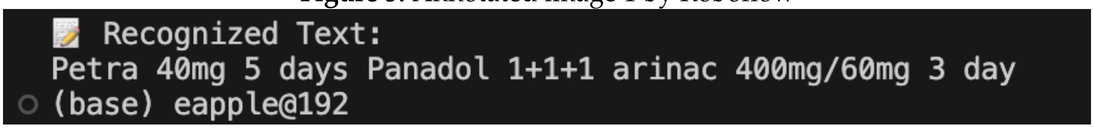

Figure 4. Output image 1 of the model

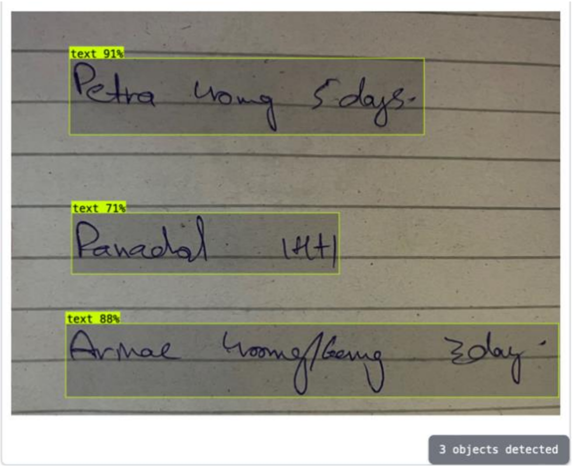

Figure 6. Annotated image 2 by Roboflow

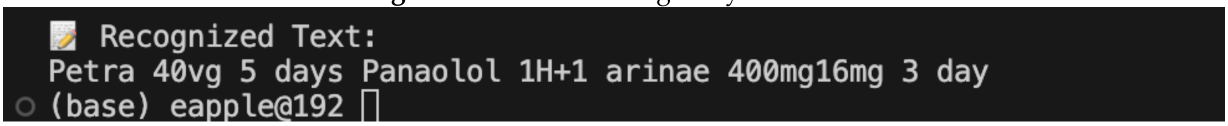

Figure 7. Output image 2 by the model

Similar to the previous evaluation, Roboflow localized three lines of text with confidence scores of 91%, 71%, and 88%, respectively. While the first and third lines were mostly accurate, the second line, with the lowest confidence, produced notable transcription errors. The model misread characters such as 'm' as 'v' ("40mg" → "40vg"), misinterpreted symbols ("1+1+1" → "1H+1"), and dropped a critical slash ("400mg/60mg" → "400mg16mg").

These errors highlight the challenges of fine-grained character recognition in handwritten prescriptions, particularly when handwriting is ambiguous, symbols are closely written, or confidence scores drop. Such mistakes could lead to serious risks in medical contexts. Future work could integrate larger domain-specific datasets, symbol-aware training, and post-processing correction techniques to improve robustness against these variations.

5. Discussion

Our findings clearly demonstrate TrOCR's strong capability in handling handwritten prescriptions, and the precision achieved when combined with Roboflow for dataset preparation. The performance metrics achieved an Exact Match Accuracy (EMA) of 81.3%, a Character Error Rate (CER) of 8.7%, and a Word Error Rate (WER) of 12.5%, representing a notable advancement in OCR for clinical contexts. This performance is especially significant given the diversity of handwriting styles and the complexity of medical vocabulary, where traditional OCR approaches often falter. These results confirm that the model is not only effective at recognizing individual characters and words but also capable of transcribing entire prescription entries with high accuracy, a critical factor for both patient safety and workflow efficiency in healthcare. Nonetheless, the analysis of Figures 6 and 7 highlights important areas for improvement. For example, errors such as interpreting "Panadol" as "Panaolol," replacing "+" with "H," or misreading "60mg/16mg" as "60mg16mg" point to the model's sensitivity to subtle script variations and symbol misclassification. These specific error types suggest targeted avenues for refinement. Symbol-aware training could reduce confusion between "+" and similar handwritten strokes, while post-processing rules could help preserve essential clinical symbols such as slashes. Likewise, misinterpretations of drug names underscore the need for integrating medical lexicons or dictionary-based correction layers, which could catch unlikely substitutions (e.g., "Panaolol") before final transcription. Another key insight is that lower confidence scores often align with the lines containing the most critical transcription errors, as shown in both figures. This pattern suggests that confidence metrics could serve as a practical flag for human review in semi-automated systems, improving safety without sacrificing efficiency. Future systems could actively incorporate these thresholds into a hybrid pipeline that combines automation with clinician oversight.

From a technical view, TrOCR's transformer-based design remains a key strength. Its Vision Transformer (ViT) encoder and autoregressive text decoder enable seamless recognition while effectively managing messy or connected characters. The attention mechanisms further allow understanding of context, which explains the strong results even with difficult handwriting. However, the errors seen in Figures 6 and 7 highlight that good architecture alone is not enough without robust training data. While our dataset of 580 annotated prescriptions provided a solid base, it may not fully reflect regional handwriting differences or non-English terms. Expanding to multi-center, multilingual datasets is therefore crucial to enhance generalizability. Roboflow played an equally important role by supporting efficient annotation and augmentation. Its data augmentation tools, including simulated rotations, scaling, and brightness adjustments, made the model more resistant to common scanning artifacts. Still, errors with very illegible writing and rare abbreviations show the need for augmentation strategies that better imitate these specific challenges. Future dataset development should aim to diversify handwriting styles and include synthetic samples of uncommon symbols, abbreviations, and degraded scan quality.

Taken together, our discussion underscores that high transcription accuracy alone does not guarantee clinical reliability. By linking observed error types in Figures 6 and 7 to targeted improvements such as symbol-aware training, dictionary integration, confidence-based review triggers, and broader dataset diversity, we establish a clear roadmap toward safer and more dependable OCR deployment in healthcare. Although challenges remain, the demonstrated correctness of TrOCR indicates a meaningful breakthrough for prescription digitization and positions this approach as a viable foundation for future clinical systems.

6. Conclusion and Future Work

This study successfully explored the integration of TrOCR and Roboflow for digitizing handwritten medical prescriptions, which are among the most challenging tasks for traditional Optical Character Recognition (OCR) systems due to the inherently illegible and diverse nature of human handwriting. Our research reveals that combining TrOCR's most advanced Transformer-based architecture for end-to-end text recognition with Roboflow's comprehensive platform for dataset management, annotation, and augmentation enables us to achieve the highest degree of accuracy in transcribing these critical documents. The results of experimentation, showing an 8.7% (CER), a 12.5% (WER), and an extremely high (EMA) of 81.3%, indicate the completeness of this integrated approach. Together, these metrics indicate TrOCR's potential, which, when combined with Roboflow-prepared datasets, can accurately comprehend the unpredictability of handwritten material and faithfully transcribe prescription information. The nature of TrOCR as an end-to-end one, which utilizes Vision Transformers for image encoding and autoregressive decoders for text generation, is robust in surpassing the barriers of traditional OCR methods. The benefits of this tool for healthcare data management and automation are very significant. Digitized, accurate prescription of drugs can potentially increase patient safety significantly by avoiding medication mistakes, decrease the workload and consequently the cost of pharmacies by eliminating manual data entry, and at the same time, enable EHR and telemedicine platforms to communicate with each other without interoperability issues. This step towards a more efficient, secure, and patient-centric healthcare ecosystem is the contribution of this technology.

Although the collaboration between TrOCR and Roboflow represents a significant step toward digitizing handwritten medical prescriptions, several issues still need to be addressed. One of the considerable problems in recognizing handwriting is the variation in writing styles, which often leads to low recognition accuracy in highly cursive and irregular scripts. Here, the authors suggest that the use of advanced preprocessing techniques or the setup of personalized and adaptive models might be beneficial. Besides, protecting data privacy is a matter of the highest importance, especially when considering the requirements of HIPAA and GDPR; thus, it is necessary to examine the methods of machine learning that will guarantee the privacy of participants, for instance, through the differential privacy concept and secure multi-party computation. Moreover, the burst of changes in the handwriting style or the introduction of new medical terms that may cause the model to deteriorate is also a hazard; thus, provisions must be made for continuous supervision and rebooting of the automated training programs. Furthermore, the shortage of a comprehensive, diverse, and well-annotated prescription dataset is a factor hindering progress in this subject matter; hence, it is necessary to collaborate on public and anonymized resources. However, it remains a technical challenge to integrate it into existing healthcare systems; therefore, future work should focus on creating standard APIs and a modular architecture that can be easily and quickly deployed. Additionally, the extension of OCR systems with a natural language understanding component may facilitate a structured extraction of clinical data, including drug names, dosages, and frequencies, thus leading to the emergence of more intelligent applications in healthcare automation.

References

1. Rice, Stephen V., Junichi Kanai, and Thomas A. Nartker. A report on the accuracy of OCR devices. University of Nevada, Information Science Research Institute, 1992.
2. Riaz, M.K., Riaz, M., Latif, A.: Medication errors and strategies for their prevention. Pakistan journal of pharmaceutical sciences 30(3) (2017)
3. 15 Ammenwerth, E., Schnell-Inderst, P., Machan, C., and Siebert, U., "The Effect of Electronic Prescribing on Medication Errors and Adverse Drug Events: A Systematic Review," *Journal of the American Medical Informatics Association*, vol. 15, no. 5, pp. 585–600, Sep.–Oct. 2008. doi: 10.1197/jamia.M2667
4. S. Islam, H. Elmekki, A. Elsebai, J. Bentahar, N. Drawel, G. Rjoub, and W. Pedrycz, "A Comprehensive Survey on Applications of Transformers for Deep Learning Tasks," *arXiv*, vol. arXiv:2306.07303, Jun. 2023.
5. H. Wang, C. Pan, X. Guo, C. Ji, and K. Deng, "From object detection to text detection and recognition: A brief evolution history of optical character recognition," *WIREs Computational Statistics*, vol. 13, no. 5, e1547, Jan. 2021.
6. Liebl, B., Burghardt, M.: On the accuracy of crnns for line-based ocr: A multi-parameter evaluation. arXiv preprint arXiv:2008.02777 (2020).
7. C. Tao, M. Filannino, and Ö. Uzuner, "Prescription Extraction Using CRFs and Word Embeddings," *Journal of Biomedical Informatics*, vol. 72, pp. 60–66, Aug. 2017.
8. M. Li, T. Lv, J. Chen, L. Cui, Y. Lu, D. Florencio, C. Zhang, Z. Li, and F. Wei, "TrOCR: Transformer-Based Optical Character Recognition with Pre-trained Models," in *Proc. AAAI Conf. Artificial Intelligence*, vol. 37, no. 11, pp. 13094–13102, Feb. 2023.
9. Rani, S.J., Saranya, N., et al.: Medical prescription analysis using machine learning. 2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS), 376–379 (2025). IEEE.
10. Zitu, M.M., Owen, D., Manne, A., Wei, P., Li, L.: Large language models for adverse drug events: A clinical perspective. Journal of Clinical Medicine 14(15),5490 (2025).
11. Sousa Neto, A.F., Bezerra, B.L.D., Moura, G.C.D., Toselli, A.H.: Data augmentation for offline handwritten text recognition: a systematic literature review. SN Computer Science 5(2), 258 (2024)
12. Patel, M., Thakkar, S.P.: Handwritten character recognition in English: a survey. International Journal of Advanced Research in Computer and Communication Engineering 4(2), 345–350 (2015).
13. Maleki Varnosfaderani, S., Forouzanfar, M.: The role of ai in hospitals and clinics: transforming healthcare in the 21st century. Bioengineering 11(4), 337 (2024).
14. Kataria, B., Jethva, H.B.: CNN-Bidirectional LSTM-based optical character recognition of Sanskrit manuscripts: a comprehensive systematic literature review. Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.(IJSRCSEIT) 5(2), 2456–3307 (2019).
15. Roboflow Inc. (2023). Roboflow API Documentation. Retrieved from https://roboflow.com/
16. Roboflow, "Roboflow Universe: the Computer Vision Community," Roboflow Blog, Aug. 11, 2021. Retrieved from https://roboflow.com/
17. Pradana, Musthofa Galih, et al. "Levenshtein Distance Algorithm in Javanese Character Translation Machine Based on Optical Character Recognition." JOIV: International Journal on Informatics Visualization 9.4 (2025): 1411-1418.
18. Lund, W.B., Walker, D.D., Ringger, E.K. 2011 International Conference on Document Analysis and Recognition, 764–768 (2011). IEEE.
19. Mozer, R., Miratrix, L., Kaufman, A.R., Anastasopoulos, L.J.: Matching with text data: An experimental evaluation of methods for matching documents and of measuring match quality. Political Analysis 28(4), 445–468 (2020).