

OCR + Generative AI Powered Prescription to Explanation System

¹Manvi Breja, ²Ishan Singh, ³Apoorv Srivastava, ⁴Piyush Yadav

^{1, 2, 3, 4}The NorthCap University, India

manvibreja@ncuindia.edu

Information of Article

Article history:

Received: 19 May 2026

Revised: 2 July 2026

Accepted: 29 July 2026

Online: 20 October 2026

Keywords:

OCR

Generative AI

Agentic AI

RAG

ABSTRACT

Healthcare systems produce and manage lots of highly sensitive information, especially in the form of prescriptions and other documents related to patients' private data. However, due to the nature of handwritten medical prescriptions and medical jargon, patients often find it difficult to follow medication instructions. Optical Character Recognition (OCR) is still in its nascent phase of digitizing handwritten prescription documents into text. Furthermore, the traditional OCR systems are noisy and difficult to understand when it comes to recognizing the handwriting of doctors to understand medical prescriptions properly. This paper utilizes Generative AI to test the ability of LLMs in conducting OCR on handwritten text and analyze its performance, and merge OCR with Generative AI to analyze and extract medicines from prescriptions in order to give a detailed and specialized description of the medicine to the user.

The paper further analyzes some OCR methods for text extraction from medical prescriptions and compares their performance. It has also been demonstrated that the usability of prescription data and the possibility of misinterpretation of medication is decreased when OCR is combined with Generative and Agentic AI. Thus the combination of artificial intelligence and OCR will help improve the understanding of patients and the accessibility of digital healthcare.

1. Introduction

OCR is one of the technologies in the digital healthcare industry that helps improve the need for automation for the interpretation and processing of medical documents. Medical prescriptions are one of the key documents and modes of communication between patients and healthcare providers. However, due to the ambiguity in medical terminologies, common use of abbreviated forms, and the difficult legibility of handwritten notes from doctors, medical prescriptions are usually very hard to read and understand. OCR (Optical Character Recognition) is one of the key computer vision techniques that captures text in an image or a scanned document and converts it into editable text. In the healthcare sector, OCR has made the digitization of medical prescriptions, lab reports, and patient records possible, enhancing the ease of storage, retrieval, and analysis of these documents. However, due to variations in handwriting, the presence of background noise in scanned documents, and the inconsistencies in document formatting, OCR-based text extraction from prescriptions is still a challenging problem.

Old OCR technologies such as Tesseract [1] have shown to be effective with printed documents. However, they still have difficulties with handwritten medical documents. Some of the more recent studies have shown the potential of using deep learning models including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks [3], and even Transformer models, to enhance the text recognition process. Such models have the ability to learn spatial characteristics of images as well as the sequential relationships between characters, and therefore, they are also more appropriate than the traditional models to handle the recognition of highly complex handwriting.

Some other recent studies [2][5] have also looked at the integration of OCR technologies with Artificial Intelligence (AI), and Natural Language Processing (NLP) in order to enhance the accuracy of text correction, and improve the systems' ability to understand context. For example, some hybrid OCR systems use cloud-based OCR [1] and language models to improve the accuracy of text recognition in combination with enhanced comprehension of the extracted text. Within the medical field, the automated interpretation of prescriptions is of the utmost necessity, because a minor misinterpretation of a medication can result in fatal outcomes. Although some existing applications in healthcare [4] do allow the ability to manually search for a medicine, none of them work that well in the automatic extraction of information from an image of a prescription and output an explanation in a personalized and easy manner.

To address this limitation, this research proposes a **Medical OCR system integrated with Generative and Agentic AI**, which not only extracts text from prescriptions but also analyzes the extracted information to provide personalized explanations based on the patient's health profile. Additionally, the system introduces a **user login and database storage module**, enabling secure storage of patient data and improving the personalization capability of the system, and the RAG (Retrieval-Augmented Generation) architecture, where the model refers to a vector database to retrieve relevant information, which would act as context for answering the user's questions.

There are 2 main research objectives of this paper -

- 1) To test and compare the ability of 3 different OCR technologies (EasyOCR, PyTesseract, LLMs) on handwritten text and see whether it is viable to use LLMs as OCR on handwritten text.
- 2) To create a personalized medicine explanation that identifies the names of medicine from the prescription and explains their uses and implications specific to the user's biodata, existing medications and conditions, and web search results.

2. Literature Review

2.1 Table: Comparison of Existing OCR Research

Research Paper	Challenges	Technique Used	Performance	Dataset
Development of Hybrid Optical Character Recognition System [1]	Difficulty in recognizing handwritten text and noisy images	Google Cloud Vision OCR + GPT based text correction	Improved usability and accuracy through AI correction	Handwritten documents
Improving Classification Accuracy for Unstructured Medical Documents via Multi-Engine OCR [2]	Single OCR engines struggle with heterogeneous medical documents	Multi-engine OCR + Ensemble Deep Learning	94.7% classification accuracy	Clinical medical document dataset
Intelligent OCR through CNN-LSTM Fusion [3]	Traditional OCR systems struggle with domain-specific documents like prescriptions	Hybrid CNN + LSTM architecture with dictionary validation	Achieved 96.8% recognition accuracy	MNIST handwritten dataset
Deep Neural Network-based Prescription Recognition [4]	Irregular handwriting, non-standard abbreviations, and poor scalability of traditional OCR	Hybrid CNN + MLP + LSTM-based deep learning architecture	Improved OCR accuracy and text correction performance	Large-scale medical prescription dataset

Vision Transformer-based Handwritten Recognition Study [5]	CNNs struggle with global dependencies and complex character structures	Vision Transformer (ViT) + VGG16 + ResNet50 comparison	Achieved ~98.26% accuracy outperforming CNN models	CMATERdb handwritten dataset
Review on Handwritten Text Recognition [6]	High variability in handwriting, noise, and computational cost	CNN + RNN + LSTM + Transformer-based architectures (survey)	Significant improvement in OCR accuracy with deep learning approaches	Multiple datasets (IAM, RIMES, etc.)
Hybrid CNN + Transformer OCR Model for Character Recognition [7]	Difficulty in recognizing complex scripts and contextual dependencies	CNN (VGG16, ResNet50) + Transformer Encoder (hybrid model)	Achieved accuracy above 98% and improved generalization	Arabic OCR & Handwritten datasets
Transformer-Based OCR Framework for Multilingual Documents [8]	Difficulty in handling multilingual handwriting and diverse writing styles	Transformer-based OCR + FA-ResNet + URRNN + CTC	Improved accuracy, precision, recall and F1-score compared to traditional OCR models	IAM OCR + Char74k multilingual dataset
Image-Based Extraction of Prescription Information using OCR-Tesseract [9]	Difficulty in extracting handwritten prescription data with varying layouts and image quality	OCR-Tesseract + Image Preprocessing + Regex-based extraction	Achieved 96% OCR accuracy after preprocessing	Real-world prescription image dataset.
AI-Based OCR System for Handwritten Medical Prescription Recognition and Interpretation [10]	Handling handwritten prescriptions with inconsistent writing styles and medical abbreviations	CNN-LSTM + NLP + Tesseract + Streamlit interface	Achieved 92% recognition accuracy with real-time prescription interpretation	Multilingual handwritten prescription dataset

From the comparison, it is evident that recent research is shifting from traditional OCR systems toward transformer-based and hybrid architectures to improve contextual understanding and accuracy, especially for handwritten and domain-specific documents like medical prescriptions.

From the above studies, it is evident that while significant advancements have been made in OCR accuracy using deep learning and transformer-based approaches, most existing systems focus only on text extraction. Very limited research addresses the integration of OCR with generative AI for contextual understanding, personalization, and automated explanation of medical prescriptions. This research aims to bridge this gap by proposing an intelligent system that combines OCR with generative AI and RAG-based reasoning to provide user-specific medical explanations.

The work illustrated in the first article [1] outlines a novel hybrid approach involving integration of the Google cloud vision optical character recognition (OCR) tool along with generative language models for purposes of error correction. The study indicates that when the two OCR tools are used together as part of an overall strategy, there is an improvement in the text subjected to processing. In another study being conducted, a multi-engine OCR framework incorporating ensemble deep learning to classify unstructured medical records was provided as a basis for improving classification of such documents. The results from that particular study indicated a level of accuracy of approximately 94.7% for classifying unstructured medical records shows that combining various engines for OCR can in fact result in greater accuracy when classifying other forms of medical record documents. The article presents an approach for developing a hybrid OCR architecture leveraging different forms of artificial intelligence (AI) and deep learning technologies using convolutional neural networks (CNN) and long-short term memory models (LSTM). This hybrid architecture is capable of recognizing both the spatial characteristics of the data being extracted along with the sequential and contextual meanings conveyed through the characters making up the words utilized within that document. This study indicates that combining these forms of AI and deep learning technologies resulted in a nearly (96.8%) increase in performance for recognizing documents (e.g. documents that contain noise) from different communities or industries and therefore lend themselves favourability when performing recognition related tasks associated with the processing of medical records. A thorough review of all studies demonstrates advances taken place within the field of optical character recognition (OCR) from traditional forms of rule-based OCR technologies to the latest generation of OCR solutions available today utilizing various forms of hybrid artificial intelligence (AI), deep learning and other advanced data science methodologies.

In the second study [2] (Improving Classification of Unstructured Medical Documents by Multi-Engine OCR), a framework comprising multiple OCR engines is suggested to enhance the classification of unstructured medical documents and to solve the problems associated with using only one OCR engine when classifying handwritten or scanned documents. The authors propose to use a confidence-based fusion of numerous OCR engines to derive the correct representation of text and subsequently utilize two categories of deep learning models (CNNs and transformer architectures) for the classification of these documents.

In the third study [3] (Intelligent OCR Using CNN–LSTM Fusion), a combination of two types of deep learning architectures (CNNs and LSTM units) is proposed to enhance the performance of OCR systems on sequentially arranged and complicated types of text. The authors use a CNN to extract the spatial characteristics of an image and an LSTM to model the sequential relationship among the characters of text for better recognition of the text, whether it be structured or unstructured. By unifying both of these architectures into a single pipeline, the authors claim to achieve improved recognition rates compared with OCR systems that do not use these hybrid techniques for both noisy and handwritten/in-context text.

The framework used in the Fourth Study was based on deep learning methods for identifying and converting handwritten medical prescriptions into structured digital records. It proposes a solution for several issues that arise due to the irregular handwriting of providers and patients (i.e. non-standard written abbreviations, inappropriate line spacing, etc.). These issues have limited the effectiveness of traditional optical character recognition (OCR) in acknowledging a provider's hand-drawn medical scripts as valid/integrable format. A convolutional neural network (CNN) is used to extract features from a patient's handwritten script. Layered perceptrons (MLP) and sequence modelling techniques are also employed in an attempt to achieve higher levels of accuracy. Advantages of this proposed

system include an improved ability to scale and generalise in a wide range of handwriting, making it more reliable for converting handwritten medical prescriptions into an electronic health record (EHR) system maintained by hospitals.

The framework discussed in the Fifth Study is an OCR-based architecture that utilizes transformer-based models to provide OCR functionality for multilingual handwritten documents. This architecture was designed to overcome limits of traditional OCR architectures (by allowing a single OCR application to be used to read handwritten text using multiple scripts and writing styles). The framework consists of a combination of Transformer models, Residual Networks, and Recurrent Neural Networks (RNN) and allows the creation of a neural network capable of capturing the spatial and sequential characteristics of the input data. The proposed system utilizes advanced preprocessing methods, such as segmentation and noise reduction, which help improve the clarity of all am prior to being sent through the OCR framework used in this research project. These advanced preprocessing techniques have produced significant impacts on the quality of the data collected from Multilingual Documents, resulting in higher quality input data for the OCR system. The research findings also indicate that transformer-based OCR systems have been shown to improve both the accuracy of character recognition and other related performance measures (e.g., precision and recall) by large margins compared to previous methods and therefore should be preferred for processing highly complex Multilingual Documents as a result.

A second study [6], conducted by the authors of this research project, examined how Vision Transformers (ViT) were able to achieve improved performance over traditional deep learning solutions for handwritten character recognition and provide a direct comparison between the use of ViT and the use of traditional convolutional neural network (CNN) techniques (e.g., VGG16 and ResNet50). The authors claim that whereas CNN-based models typically recognize only local features, transformer-based models are capable of identifying global relationships among the characters within the input text due to the use of a self-attention mechanism. The findings suggest that ViT processes data with greater accuracy and more universally than previous CNN-based methods, particularly for handwriting that is significantly more complex or variable. This study supports the conclusions drawn from this research project that have shown a clear advantage of using transformer-based models compared to traditional deep learning OCR models in the development of current OCR systems.

Another article [7], which is a review of OCR technology evolution from traditional rule-based systems through contemporary deep learning technology, summarizes research concerning handwritten text recognition. It describes the use of CNNs for feature extraction, RNN/LSTM for sequential data modelling, and transformer architecture for end-to-end systems. The article explains how, despite remarkable improvements in OCR technology through deep learning, there are still challenges associated with noise, variability in handwriting, and excessive computational resources required to accurately identify handwritten text. The article concludes by suggesting the need for a more robust, efficient modelling technique designed for real-world handwritten data recognition.

A research paper [8] describes a hybrid OCR system composed of CNNs and transformer architectures for improved recognition of characters. Pre-trained CNNs (VGG16 and ResNet50) are the foundation for feature extraction while the transformer encoder captures the global dependencies associated with transformed characters. The hybrid architecture provides a better capacity to recognize complex characters as well as context-dependent characters because of its hybrid architecture. The hybrid architecture increases recognition accuracy and generalization when recognizing handwritten or complex instances of text.

Another research paper [9] describes a system for extracting information from prescriptions using an imaging-based approach along with an existing OCR library called Tesseract and associated advanced image preprocessing. The preprocessing techniques include grayscale and adaptive thresholding and noise reduction/cleanup will enhance the readability of the data processed by the OCR library and improve the extraction accuracy of the data extracted from the prescriptions by the OCR engine. Regular expressions are then utilized to identify and extract important data elements from the prescription document (i.e. patient demographics, medications, dosing/usage instructions). The experimental trials conducted on both handwritten and print prescriptions show an approximate OCR accuracy of about 98%. The study shows the value of preprocessing and structured extraction techniques used in the digitization process regarding healthcare documents.

A different type of research [10] details a new AI-based OCR approach for recognizing and interpreting handwritten medical prescriptions using a hybrid CNN-LSTM architecture with NLP methods. Binarization, noise reduction, skew correction and line segmentation are one of many preprocessing techniques that help to improve accuracy when recognizing handwritten prescriptions. The system not only performs OCR but also uses NER techniques to extract medical entities from the prescriptions, such as drug names, dosages and instructions for administering the medications. The study additionally provides a Streamlit-based web interface to facilitate real-time interpretation of prescriptions and integration into a web-accessible database for healthcare applications. Results indicate an approximate recognition accuracy of 92%, demonstrating the capacity of deep learning, OCR and NLP when combined to develop an intelligent healthcare system.

Overall, the collective body of work discussed demonstrates that OCR systems have advanced considerably through the use of CNNs, LSTMs, transformer-based architectures, as well as through preprocessing techniques for the recognition of handwritten medical prescriptions. A number of studies have also integrated NLP and entity recognition techniques to extract structured medical data; however, most current research focuses on extracting text and converting it to a digital form without considering how context impacts patient-specific interpretation of prescriptions. There are a few studies that look at the use of Optical Character Recognition (OCR) in conjunction with both Generative AI and Retrieval Augmented Generation (RAG) when producing individualized medication-related education and providing interactive health assistance to patients [9] [10]. This study aims to create a means by which to integrate OCR, Generative AI, RAG, and agentic reasoning into one integrated intelligent system for interpreting prescriptions.

3. Research Objectives

The main purpose of this research study is to evaluate and compare the performance of various OCRs when trying to identify handwritten text from images of prescription drug labels.

RO1. Evaluate and Compare OCRs:

To evaluate and compare different OCRs performance in the extraction of handwritten textual information from medical prescription label images.

RO2. Create a System Architecture:

To develop an architecture that combines several forms of AI with CRUD functionality for the interpretation of prescription use and the generation of prescription instructions

4. Proposed Work

The proposed project will include a data store, which will include a data access security layer and a patient information management system. All users are required to create an account and authenticate themselves through an application to access their information. The data store will contain the properties listed above including, but not limited to:

1. User credentials (username and password)
2. Patient profile details (age, allergies, medical history)
3. Uploaded prescription images
4. Extracted prescription text
5. AI-generated explanations with RAG

This stored information allows the system to maintain patient history and provide more personalized medical guidance.

Flow Diagram

The system workflow follows these steps:

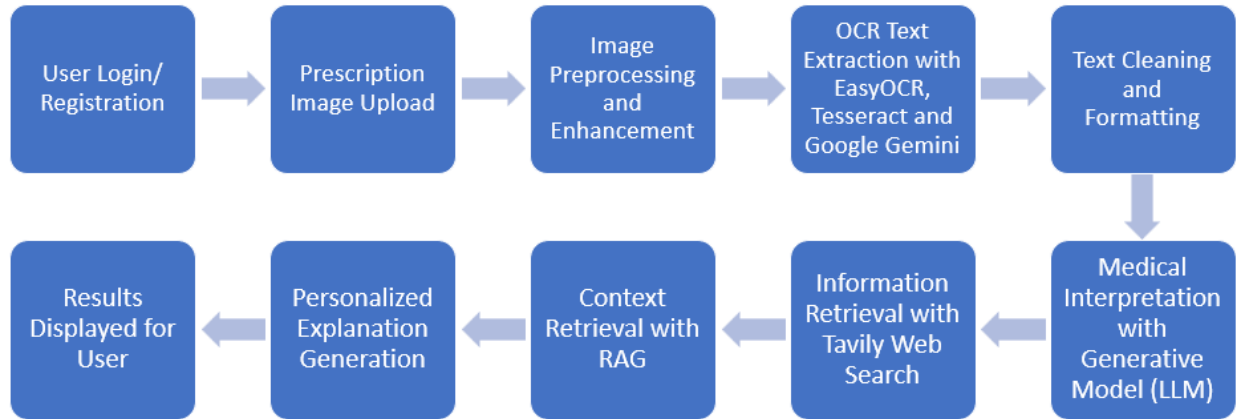


Fig.1. Flow Diagram

1. User Login/Registration
2. Prescription Image Upload
3. Image Preprocessing and Enhancement
4. OCR Text Extraction with EasyOCR, Tesseract and Google Gemini
5. Text Cleaning and Formatting
6. Medical Interpretation with Generative Model (LLM)
7. Information Retrieval with Tavily Web Search
8. Context Retrieval with RAG
9. Personalized Explanation Generation
10. Results Displayed for User

6. Implementation

The system is implemented using Python as the core programming language. Streamlit is used for the UI used for comparing the performance of the 3 OCR engines, while Next.js, React, TypeScript and Tailwind CSS are used to build the web interface that allows users to login, upload prescription images and interact with the medical prescription explanation system. Image processing operations such as preprocessing and format conversion are performed using Pillow. Text extraction is carried out using 3 techniques – EasyOCR, PyTesseract OCR, and Google Gemini LLM which convert prescription images into machine-readable text. The ability to search the web for results is done with the help of Tavily Web API, and the RAG Architecture is implemented with ChromaDB.

7. Observations and Result

- 1) Regarding the first objective, the experimental comparison of different OCR techniques indicates a significant difference between the performance of generic OCR libraries and LLMs. While common OCR engines like Easy OCR and Tesseract may perform well on digital/highlighted text, their performance significantly deteriorates when applied on images containing handwritten text. This is in contrast to Although Google Gemini did not achieve perfect accuracy to achieve 100% accuracy on handwritten text, still outperformed the other OCR engines by a good margin.

Traditional OCR engines such as Tesseract and EasyOCR provide reliable results for printed text but may struggle with irregular handwriting.

7.1 Sample Prescriptions and Their Outputs

a)

Date of Consultation 18/12/19 +919717002200

Current medication: Past illness / surgery	Family history:	Allergies
--	-----------------	-----------

Chief complaints:

History & Examination:

Provisional diagnosis: Ame Froptic

Tests advised	Treatment /Advice
- Ame Aid soap	① 30 sec ① 15 min
- 1/2 Aincup wif 20 sec	① 15 min
- Salyrap spray	4A night ① 1 hr
- Madexia gel	4A night ① 1 hr
- fixderme foot area	① 1 hr

Follow up: _____


 Signature of Practitioner



Tings
 Making the World City



10 AM to 1.00PM (Monday to Saturday)

Fig.2. First Handwritten Prescription Image

<div>  Extracted OCR Text </div>		
Tesseract OCR:	EasyOCR:	LLM OCR:
<p>+919717002200</p> <p>Date of Consultation Ff) 2/19</p> <p>Current medication: Past illness / Family history: Allergies] surgery </p> <p>Chief complaints:</p> <p>History & Examination:</p> <p>Provisional diagnosis: Atnr~ i= prea Tests advised Ate My «os —Sehjres & 'YA Page OF - Npanleen~ ¢ V7 Sf— zyesy (> ya 2 fk sgeee ¢ jot cecal) b.</p> <p>Treatment /Advice</p> <p>7. Sete</p> <p>Follow up: i 5</p> <p>Signature of Practitioner</p> <p>the 'ity 10AMto 1.00PM(Monday to Saturday)</p>	<p>+919717002200 Date of Consultation 8 [12/19 Current medication: Past illness Family history: Allergies surgery Chief complaints: History & Examination: Provisional diagnosis: HLLS < EutLs Tests advised Treatment /Advice Ai c ~Ad 82 ep 7 /++ 24€ / 14" Z4/ '-7541 3a Njcxelck^ 44"7v6/ Bk Acceszv Zv / 22-0 h Follow up: Signature of Practitioner the 10.AMto onpM(Mandav to Saturdav) dal</p>	<p>+919717002200</p> <p>Date of Consultation 18/12/19</p> <p>Current medication: Past illness / Family history: Allergies surgery</p> <p>Chief complaints:</p> <p>History & Examination:</p> <p>Provisional diagnosis: Acne Tropicalis</p> <p>Tests advised Treatment / Advice</p> <ul style="list-style-type: none"> • Acne-Aid soap ① Bali 2-2 N (15) Tab • Viseup 250 Q night ① • Salizap spray Q/A morning ① • Nadoxia gel cream ① • Firderma face cream <p>Follow up:</p> <p>[Signature of Practitioner] Signature of Practitioner</p> <p>Timings Medanta the Medicity 10.AM to 1.00PM (Monday to Saturday)</p>

Fig.3. Comparison of OCR Engines on First Handwritten Image

b)

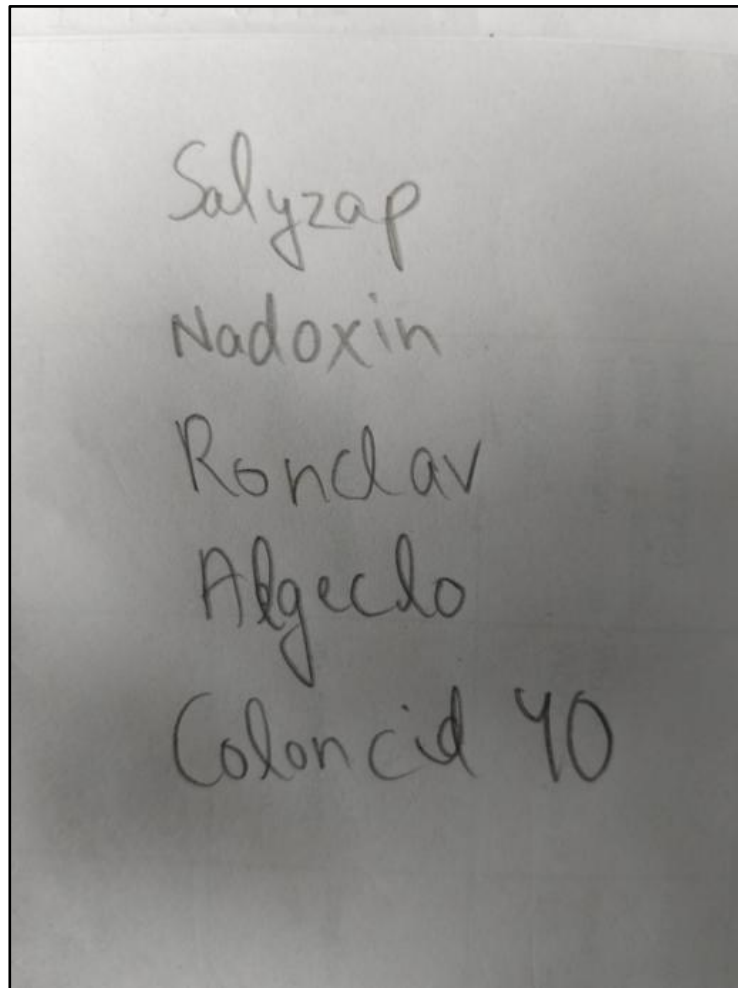


Fig.4. Second Handwritten Image

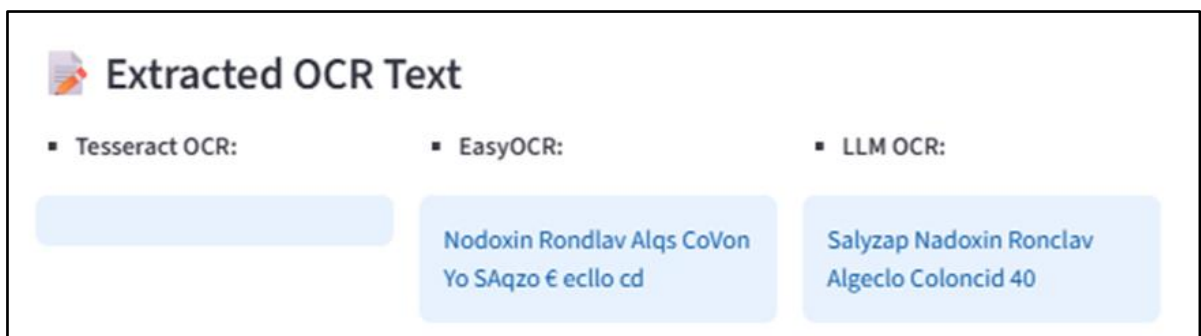


Fig.5. Comparison of OCR Engines on Second Handwritten Image

- 2) After seeing the results of the performance of 3 OCR engines, it was conclusive to determine that there was not going to be a viable way to conduct OCR on handwritten prescriptions with good accuracy. Therefore, it was decided that for the medicine explanation system, the OCR would be conducted on digital text with basic OCR engines. Upon performing the same, medicine names were extracted successfully from the prescriptions and a concise and informative explanation system was successfully integrated with it.

7.2 Sample Prescriptions and Their Outputs

Dr. Akshara M.S. Reg. No: MMC 2018		SMS hospital B/503, Business Center, MG Road, Pune - 411000. Ph: 5465647658, Timing: 09:00 AM - 01:00 PM, 06:00 PM - 08:00 PM Closed: Sunday
---	---	--

ID: 11 - OPD6 PATIENT (M) / 13 Y Mob. No.: 9423380390 Address: PUNE Weight (Kg): 80, Height (Cm): 200 (B.M.I. = 20.00), BP: 120/80 mmHg	Date: 30-Aug-2023
--	--------------------------

Chief Complaints	Clinical Findings
* FEVER WITH CHILLS (4 DAYS) * HEADACHE (2 DAYS)	* THESE ARE TEST FINDINGS FOR A TEST PATIENT * ENTERING SAMPLE DIAGNOSIS AND SAMPLE PRESCRIPTION

Diagnosis:
* MALARIA

R

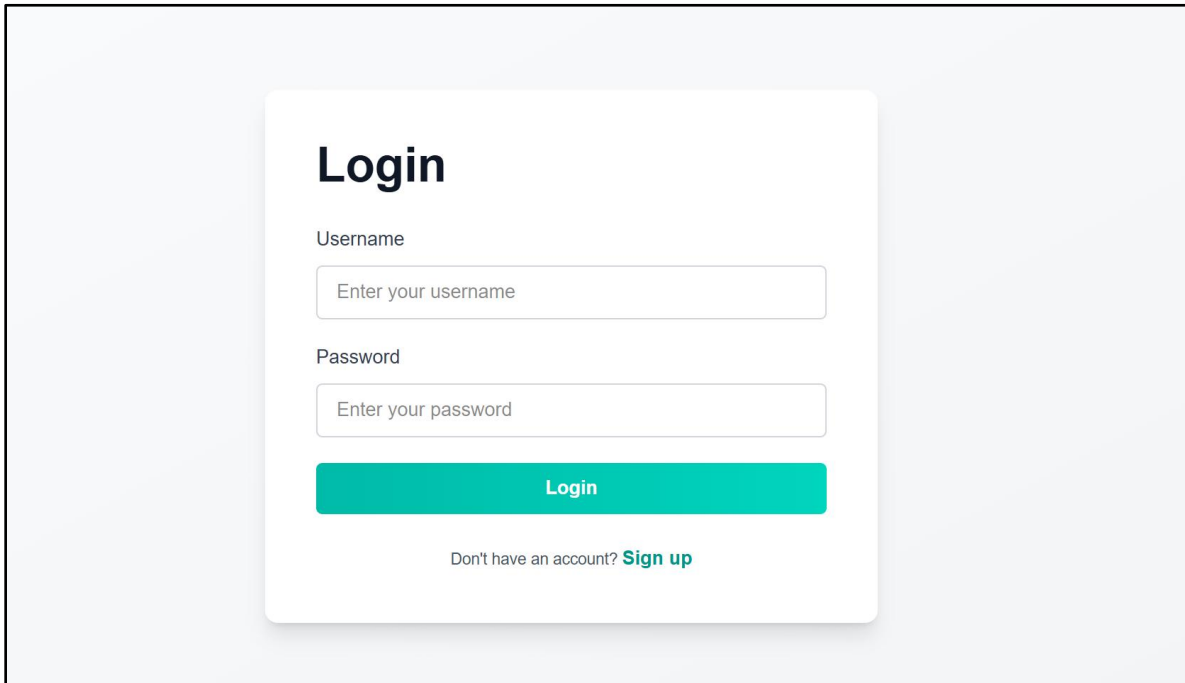
Medicine Name	Dosage	Duration
1) TAB. ABCIXIMAB	1 Morning	8 Days (Tot:8 Tab)
2) TAB. VOMILAST DOXYLAMINE 10MG + PYRIDOXINE 10 MG + FOLIC ACID 2.5 MG	1 Morning, 1 Night (After Food)	8 Days (Tot:16 Tab)
3) CAP. ZOCLAR 500 CLARITHROMYCIN IP 500MG	1 Morning	3 Days (Tot:3 Cap)
4) TAB. GESTAKIND 10/SR ISOXSUPRINE 10 MG	1 Night	4 Days (Tot:4 Tab)

Advice:
* TAKE BED REST
* DO NOT EAT OUTSIDE FOOD
* EAT EASY TO DIGEST FOOD LIKE BOILED RICE WITH DAAL

Follow Up: 04-09-2023

Susbtitute with equivalent Generics as required.

Fig.6. Sample Digital Text Prescription

A login page design featuring a white card on a light gray background. The card has a title 'Login' in bold black font. Below the title are two input fields: 'Username' with a placeholder 'Enter your username' and 'Password' with a placeholder 'Enter your password'. A teal 'Login' button is positioned below the password field. At the bottom of the card, there is a link 'Don't have an account? Sign up' in a smaller font, where 'Sign up' is highlighted in teal.

Login

Username

Enter your username

Password

Enter your password

Login

Don't have an account? [Sign up](#)

Fig.7. Login Page

Create Account

Username

Choose a username

Password

Create a password

Medical Information

Age

Age

Gender

Select Gender

Height (cm)

Height in cm

Weight (kg)

Weight in kg

Blood Group

Select Blood Group

Existing Medical Conditions

e.g., Diabetes, Hypertension

Allergies

e.g., Penicillin, Peanuts

Current Medications

e.g., Aspirin 100mg daily

Medical History

e.g., Previous surgeries, hospitalizations

Sign Up

Already have an account? [Login](#)

Fig.8. Sign Up Page for New User

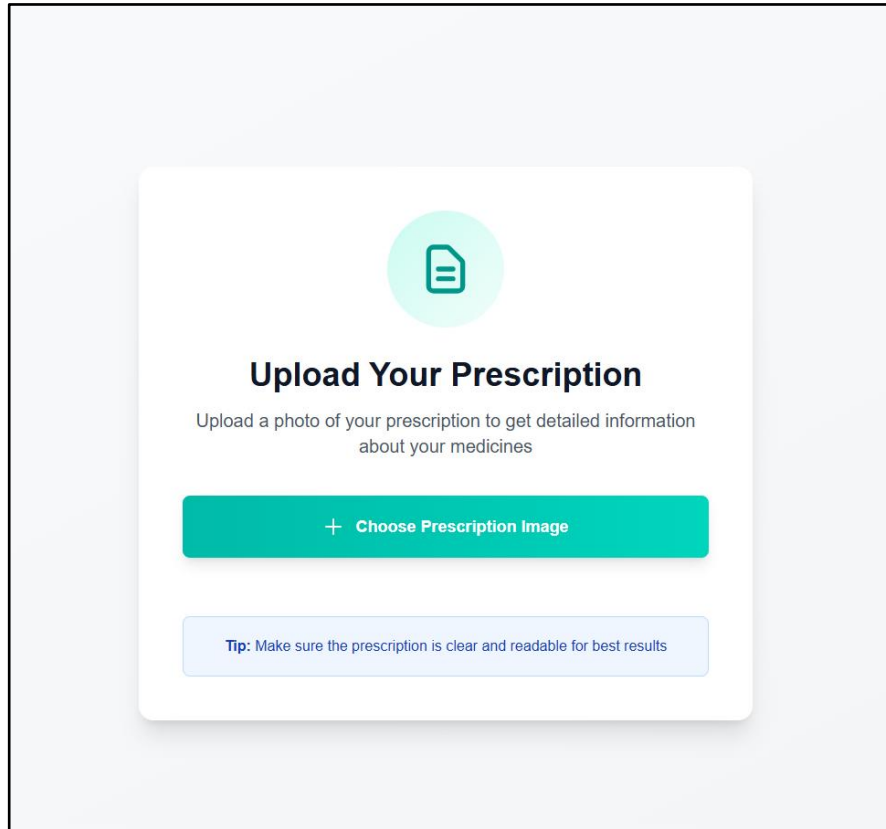


Fig.9. Image Upload Section

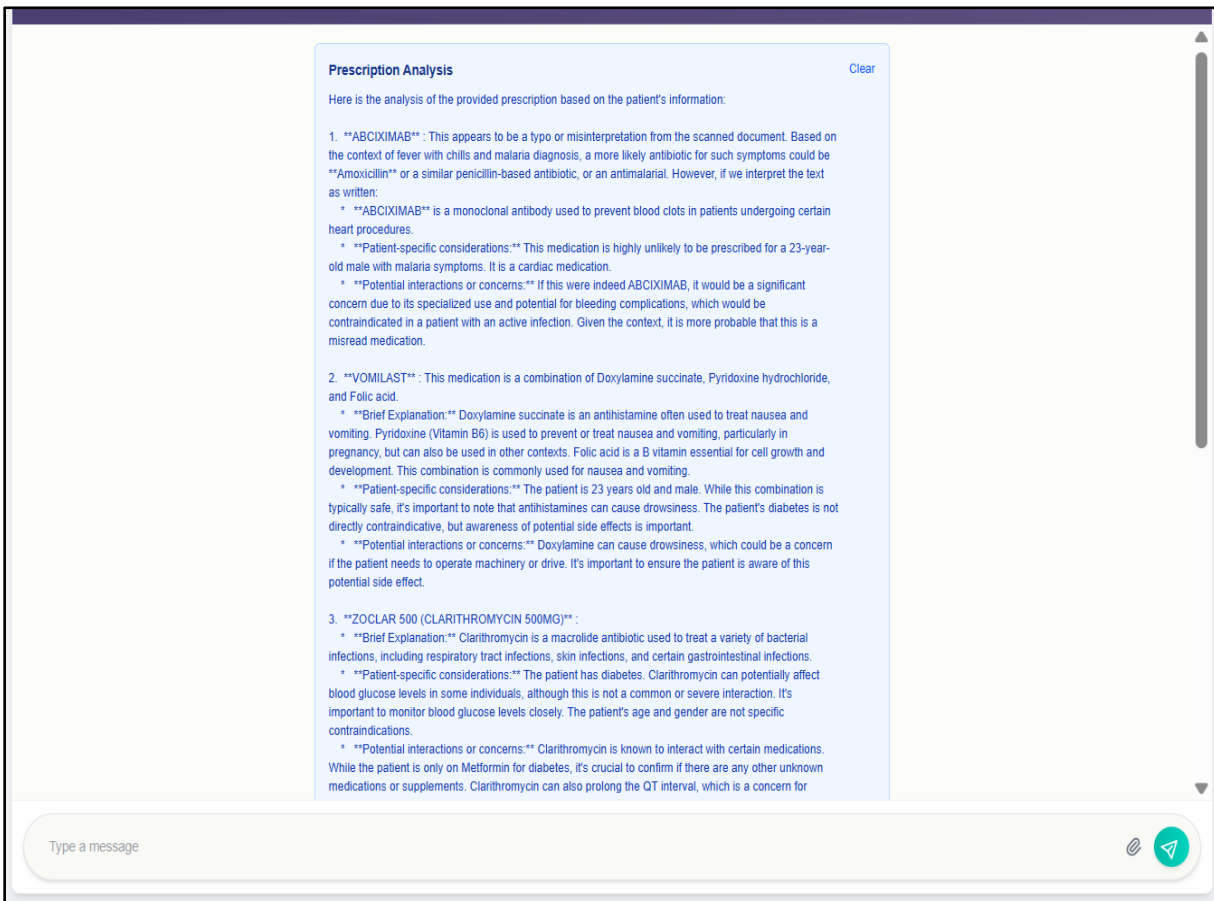


Fig.10. Chatbot Initialization with Medicine Identification and Explanation

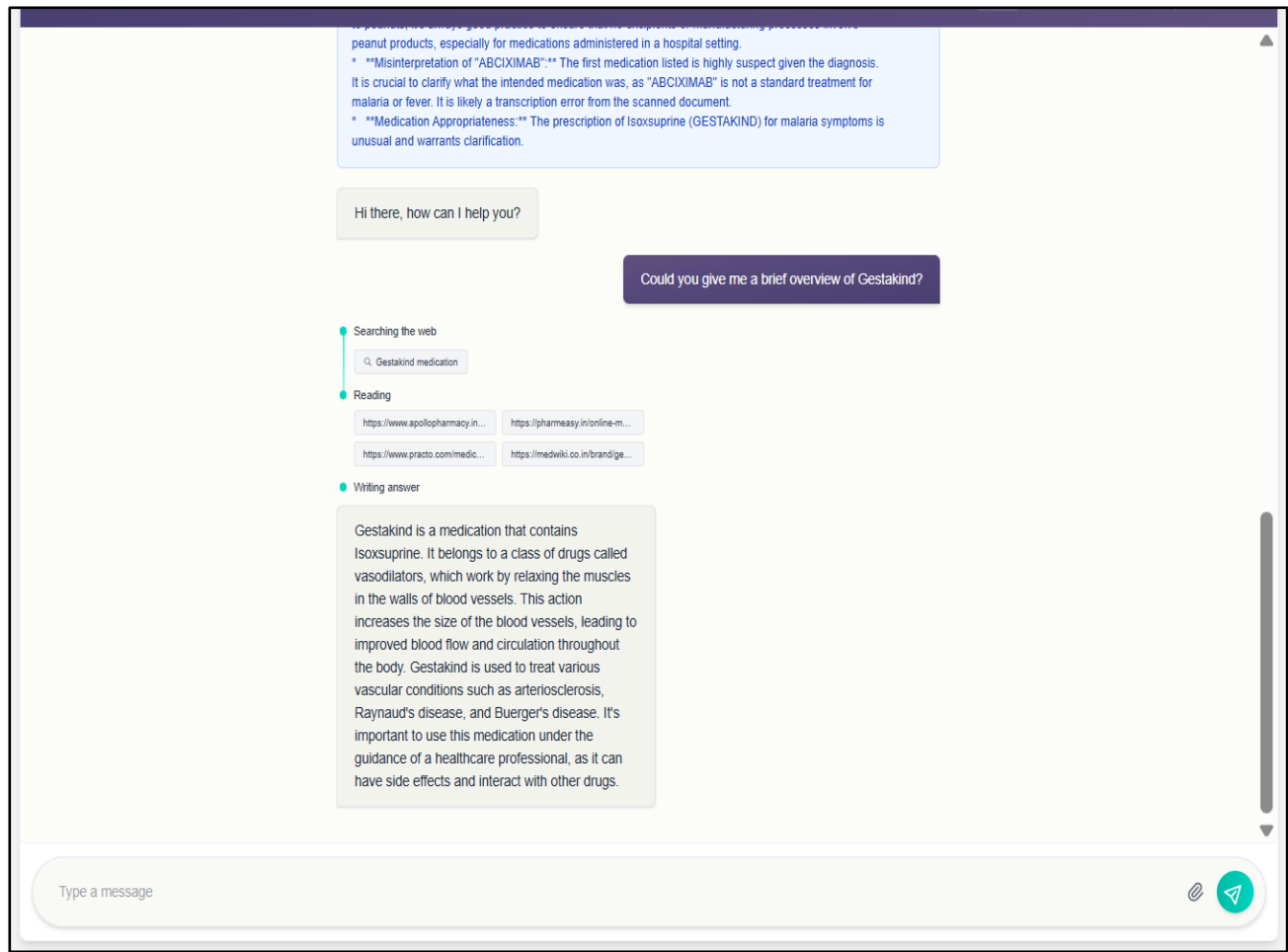


Fig.11. Chatbot Execution

Performance of Different OCR Engines on Handwritten Text	
OCR Engine	Accuracy (Handwritten)
EASYOCR	12.7%
PYTESSERACT	32.6%
GOOGLE GEMINI LLM	83.1%

8. Comparison Table

Research Paper	Technology Used	Technology Used
Improving Classification Accuracy for Unstructured Medical Documents via Multi-Engine OCR	Multi-engine OCR + Ensemble Deep Learning	95.8%
Deep Neural Network-based Prescription Recognition	Hybrid CNN + MLP + LSTM based deep learning architecture	95.5%
Image-Based Extraction of Prescription Information using OCR-Tesseract	OpenCV + Tesseract	96%
AI-Based OCR System for Handwritten Medical Prescription Recognition and Interpretation	CNN + LSTM	92%
Our Approach	Open-Source LLM	83.1%

9. Conclusion

Based on the observations made, it can be concluded that the LLM-based OCR approach significantly outperformed traditional OCR engines such as EasyOCR and PyTesseract in recognizing handwritten medical prescriptions. Despite the numerous obstacles presented by different writing styles, use of abbreviations, and variability in the quality of images, recognition of handwritten prescriptions is an area of research that is both challenging and has numerous opportunities for development. Our experimental results illustrate that there are ways to positively affect OCR systems by adding Generative AI to the existing technologies. In addition to using OCR tools to extract text from a physical document, we also integrated RAG (Retrieval-Augmented Generation) and Agentic AI capabilities to generate personalized medicine explanations and increase access to digital healthcare applications.

The development of Multimodal LLMs as well as continued improvements in transformer-based architecture for OCR may be the key ingredients to advancing handwriting prescription recognition technology. Overall, we see that existing OCR technologies are still faced with formidable barriers with respect to accurate recognition of handwritten prescriptions; however, it appears that LLM-based solutions are achieving good progress towards improving recognition performance.

10. References

- [1] ONU, S., et al. (2025). Development of hybrid optical character recognition system. *Journal of Xi'an Shiyou University*, 21(12), 358–373 <https://doi.org/10.31891/2307-5732-2022-309-3-50-54>
- [2] ZHANG, Q. (2026). Improving classification accuracy for unstructured medical documents via multi-engine OCR and deep learning collaboration. *Journal of Advanced Computing Systems*, 6(1), 1–14. <https://doi.org/10.69987/JACS.2026.60201>
- [3] KUMAR, A. N., & APARNA, S. (2026). Intelligent optical character recognition through CNN–LSTM fusion with dictionary validation and error correction. *Journal of Theoretical and Applied Information Technology*, 104(3), 318–321. <https://doi.org/10.5281/zenodo.18666690>
- [4] MITRA, M., & CHAUBEY, A. (2025). A deep learning-based framework for handwritten medical prescription recognition using OCR and NLP <https://doi.org/10.1109/cicn67655.2025.11367989>
- [5] PARVEZ, A. H. M. S., et al. (2025). Enhancing Bangla handwritten character recognition using vision transformers, VGG-16, and ResNet-50: A performance analysis. *Frontiers in Big Data*, 8. <https://doi.org/10.3389/fdata.2025.1682984>
- [6] GARG, V., et al. (2025). A review on handwritten text recognition. *International Journal for Research Trends and Innovation*, 10(5).
- [7] AL-MAAMARI, M. R., et al. (2025). Integrating CNN and transformer architectures for superior Arabic printed and handwriting character classification. *Scientific Reports*, 15. <https://doi.org/10.1038/s41598-025-12045-z>
- [8] DADAPEER, S., & SURESH, R. (2026). Transformer-based OCR framework for multilingual handwritten documents. <https://doi.org/10.48084/etasr.15667>
- [9] PONNURU, M., PONMALAR, P. S., A, L., S. B., T., & G. C., G. (2024). Image-based extraction of prescription information using OCR-Tesseract. *Procedia Computer Science*, 235, 1077–1086 <https://doi.org/10.1016/j.procs.2024.04.102>
- [10] SHARJEEL, S., & ARIF, M. U. (2025). AI-based OCR system for handwritten medical prescription recognition and interpretation. *Indian Journal of Computer Science and Technology*, 4(2), 365–370 <https://doi.org/10.59256/indjcst.20250402050>