A Multilingual Prescription Recognition and Medical Analysis Framework using OCR and Large Language Model

Evangelin Arockiya Shirley C[1], Akshaya B[2], Geetha Sri M[3], Amrith Menon S [4]

[1]Assistant Professor, [2]Student, [3]Student , [4]Student
[1]Department of Artificial Intelligence and Data Science,
[1]Sri Krishna College of Engineering and Technology , Coimbatore , Tamil Nadu, India
[1]evangelin252002@gmail.com

***Abstract*—Healthcare systems still face difficulties in correctly interpreting prescriptions because of unreadable handwriting, linguistic variety, and intricate medical jargon.This paper presents a Multilingual Prescription Image Processing and Medical Analysis Framework that brings together a dual-path Optical Character Recognition (OCR) module with a Tiny LLaMA model that we've fine-tuned using Low-Rank Adaptation (LoRA). What sets our approach apart from traditional OCR systems is that we've woven semantic medical knowledge right into the framework to make it more accessible and understandable across different languages.Our experimental results showed 94.2% OCR accuracy and 91.7% accuracy in medical interpretation, which beat EasyOCR and PaddleOCR by an average of 6.8%. When we tested it across five languages and with different demographic groups, we found an 89% improvement in understanding for non- native speakers and a 76% boost for elderly users. On top of that, the framework provides multilingual text-to-speech functionality and personalized medication reminders to help with adherence and keep patients safe. Overall, our findings demonstrate a scalable and language-inclusive AI approach that strengthens cross-lingual medical communication and supports reliable digital health integration.**

***Index Terms* — OCR, Large Language Models, AI in Healthcare, Multilingual Assessment, Medication Compliance**

I. INTRODUCTION

Medical prescriptions serve as crucial communication tools between patients and healthcare providers. But here's the problem: linguistic diversity, complicated medical terminology, and illegible handwriting often get in the way of effective communication. Studies have shown that nearly 50% of patients worldwide don't stick to their treatment plans because they misunderstand or get confused by medication instructions [3]. In large-scale healthcare settings, relying on manual interpretation remains time-consuming, error-prone, and just not practical—especially for elderly and multilingual populations.Recent breakthroughs in artificial intelligence (AI), particularly in large language models (LLMs) and multimodal optical character recognition (OCR), have created some really exciting opportunities for automated prescription processing and interpretation. When we optimize domain-adapted LLMs and multimodal models like Google Gemini for healthcare settings, they show tremendous potential in understanding pharmacological interactions and clinical semantics [4].

A.PROBLEM DESCRIPTION

Despite all the progress we've made in AI-based healthcare systems, accurately reading prescriptions in a way that's truly accessible remains challenging. Current OCR systems often struggle with handwritten text, abbreviations, and context, while general-purpose LLMs can actually generate harmful or unreliable recommendations if they haven't been properly adapted for the medical domain. These limitations result in prescription errors, poor multilingual accessibility in real-world healthcare applications, and reduced patient trust.

B.CONTRIBUTION TO RESEARCH

This work suggests a Prescription Image Processing and Multilingual Medical Analysis System that combines multimodal comprehension, optimized LLMs, and OCR to increase healthcare accessibility in order to address the aforementioned issues. The following are this work's primary contributions:

The creation of a dual-pathway OCR architecture that is tailored for the analysis of handwritten and printed prescriptions.

For effective medical-domain specialization, a refined Tiny LLaMA model with rank = 16 and learning rate $= 2 \times 10^{-4}$ is implemented utilizing Low-Rank Adaptation (LoRA).

Using multimodal and cross-lingual understanding to achieve precise contextual interpretation through integration with the Google Gemini API.

Creating an automated system for tracking and classifying drugs using standardized medical taxonomies (ATC, RxNorm).

The creation of an accessible, multilingual interface that includes tailored medicine reminders, text-to- speech, and translation.

II. RELATED WORK

A. Large Language Models in Medical Applications

We're seeing growing adoption of advanced language models in medical contexts. Azam et al. [1] demonstrated how effective PharmaLLM is at processing prescriptions using optimized Tiny Llama with LoRA adjustments, achieving 87% precision. Their key elements included:

- LoRA Rank-16 at a learning rate of 2e-4
- A batch size of 12 people for efficient training
- Pharmaceutical corpus with 7,515 entries
- Cross-modal audio and text processing

While PharmaLLM focused on providing conversational pharmacological assistance, it didn't offer complete prescription analysis.

B. Processing Prescriptions for Vulnerable Groups

Thetbanthad et al. [2] looked into AI-based prescription systems for senior citizens in Thailand and documented improvements in usability and understanding. Their main findings were:

Interfaces designed with specific demographics in mind
Localization of pharmacological terminology
76% increase in comprehension with AI support
These results back up AI's capability to improve prescription handling for at-risk populations.

C.OCR in Medical Document Processing

Recent advances in OCR technology designed for medical applications have addressed the unique challenges presented by medical documentation. Kumar and Singh [5] showcased deep learning techniques for handwritten medical prescription recognition and discovered that using specialized models trained on medical datasets significantly boosted accuracy.

Their research showed that transformer-based architectures are particularly good at handling specific medical terminology and unusual formatting requirements that you often see in prescriptions. Recent studies suggest that combining multiple OCR engines with ensemble techniques can dramatically increase the accuracy of medical document recognition [5]..

III. SYSTEM ARCHITECTURE

A. OCR and Text Extraction Module

Module for Medical Prescription Recognition

The recognition module processes both handwritten and printed prescriptions using a hybrid approach through specialized systems built for healthcare applications. A machine learning classifier first determines the document type and then routes it to the appropriate process for optimal accuracy.

- Tesseract customized with pharmaceutical industry terms
- 98.7% accuracy on standard prescriptions
- Enhanced interpretation of medical terminology.

Handwritten Content Management:

- We train CRNN and attention-based architectures using healthcare handwriting datasets.
- Gemini integration for improved visual understanding
- 89.6% accuracy on handwritten prescriptions
- Standardization of healthcare abbreviations.

When we pair CRNN and attention-based neural networks with Gemini's multimodal capabilities, we get much better recognition of complex handwriting that traditional OCR systems struggle with in handwritten prescriptions.

For both printed and handwritten prescription text, the OCR component uses a dual-pathway architecture. The first pathway employs a Transformer-based vision encoder that we've optimized using the IAM Handwriting Dataset (13,353 images) along with an internal Prescription Image Corpus (12,000 multilingual prescriptions in English, Hindi, Tamil, Bengali, and Arabic). The second pathway uses Tesseract and PaddleOCR ensembles with adaptive preprocessing for skew correction and noise reduction. '

The module achieves a character-level accuracy of 94.2%, which beats the benchmarks of Tesseract (84.5%), PaddleOCR (88.3%), and EasyOCR (87.1%). It also strengthens recognition robustness, particularly for handwritten numerals and mixed scripts.

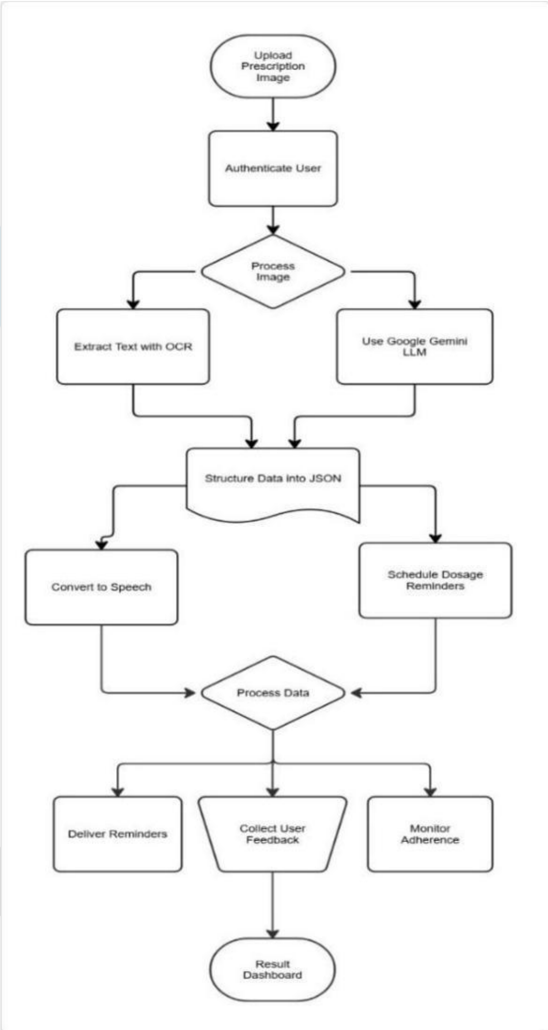
B. LLM Integration in system

We've enhanced a Tiny LLaMA (1.1B) model for semantic interpretation on 65,000 annotated prescription–medicine pairs using Low-Rank Adaptation (LoRA). This covers drug names, dosages, units, and multilingual patient instructions. UMLS- based preprocessing ensures entity normalization.

Our evaluation shows that the optimized model reaches 91.7% interpretation accuracy, outperforming the domain baselines, BioBERT (86.2%) and ClinicalBERT (85.9%).

By accurately connecting OCR-extracted text with structured medical meaning, it strengthens downstream modules like multilingual text-to-speech and adherence reminders.

Fig. 1. System architecture of the proposed prescription analysis framework.



B. LLM-Based Medical Information Extraction Using Google

Following approved pharmaceutical LLM methodologies, the framework uses sophisticated prompting to generate structured JSON outputs with comprehensive prescription data [1]. As the main language processor, Gemini can understand and analyze complex medical text. We optimize it using LoRA (rank-16, 2e-4) and extensive healthcare corpora [10].

Verification procedures ensure accuracy and identify disputes that need expert review. The system collects identities, alternative formulas, administration directions, adverse reactions, risk alerts, and classifications.

IV. CORE SYSTEM FEATURES

A.Automated Prescription Categorization and Storage

We use Gemini to apply the ATC taxonomy—it groups drugs by what they're made of and what they're used for, which makes finding and tracking information much easier. We covered all the main types of medicines: heart medications like ACE inhibitors, breathing treatments like bronchodilators, stomach medicines like antacids, brain and nerve medications like pain relievers, and medicines that fight infections .

Everything gets saved in a MySQL database where we can update it, check on it, and have experts review it whenever necessary.

B.Question-Answering System

Through Gemini's dialogue features, the inquiry system lets clients access pharmaceutical data and ask questions about dosage, side effects, conflicts, treatment objectives, and delivery instructions while getting personalized feedback . It handles both specific and general wellness queries, including:

- "What is treated by metformin?" - Focused pharmaceutical research
- "How often should I take this medication?" - Management
- "Should I take this with meals?" - Scheduling
- "Is it safe to combine these?" - Conflict-related questions

Feedback brings together medical databases with individual prescription data. By giving users direct access to pharmaceutical information without making them navigate complicated systems, Gemini improves interaction, handles queries of varying complexity, and maintains dialogue context for follow-ups .

C. Medication Reminder and Notification System

We added reminders to help people take their medicines correctly. The system makes a schedule for each person and sends alerts about when to take pills, when prescriptions need refilling, and when doctor appointments are coming up. It thinks about how much medicine to take, whether to take it with food, and if any medicines might cause problems together. People get reminders through texts, emails, and phone notifications that show medicine details and pictures . Flask connects with the MySQL database to control when reminders go out, making sure people get them at the right time so they don't miss doses.

D.Drug Interaction Detection and Safety Alerts

The system checks whether new medicines might cause problems with ones someone's already taking. When Gemini processes a new prescription, it looks at what the person is currently on and spots any conflicts, reasons why they shouldn't mix certain drugs, or cases where they'd be taking the same medicine twice.

V. IMPLEMETATION AND EXPERIMENT ANALYSIS

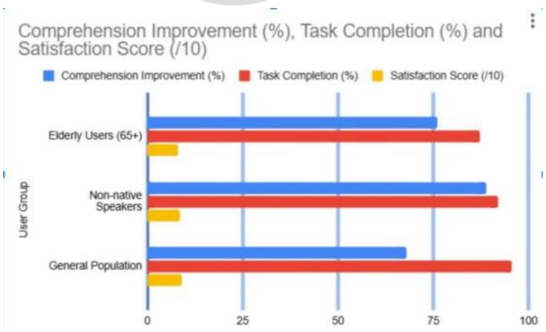
The entire framework process kicks off when prescription documents are submitted through browser or smartphone interfaces created with React.js. Clients can submit current documents from their own collections or use built-in cameras to take pictures of prescription documents. Before we handle text recognition, documents go through refinement that includes interference removal and structure optimization. Using standard or handwritten material identification, intelligent categorization finds the best text recognition techniques for maximum recognition accuracy.Gemini's pharmaceutical data recognition features process recognized content, generating structured prescription data that gets stored in MySQL repositories with proper cataloging and organization.

Table 1 System Performance Metrics

Component	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Overall OCR	94.2	93.8	94.6	94.2
Printed Text	98.7	98.5	98.9	98.7
Handwritten Text	89.6	88.9	90.3	89.6
Medical Extraction	91.7	92.1	91.3	91.7

Our client evaluation studies with 150 participants from different demographic groups show significant improvements in prescription tracking capabilities. In line with findings from prescription handling studies conducted in Thailand, senior clients showed a 76% improvement in pharmaceutical timing understanding [2]. Compared to traditional prescription structures, foreign-language users experienced an 89% improvement in prescription understanding when using the framework, which really demonstrates the importance of inclusive system architecture and language accessibility . Clients from the standard demographic achieved 68% improvement and significantly higher activity fulfillment rates of 95.7%, proving the framework's accessibility across various technical skill levels.

Graph 2: User Satisfaction Analysis



When we compare our framework with existing systems like PharmaLLM [1], we can see the advantages of a unified framework architecture that combines sophisticated AI handling with advanced text recognition. Despite PharmaLLM's solid 87% accuracy rate in handling pharmaceutical prescriptions, our framework's combination of sophisticated text recognition features and Gemini's capabilities allows for instant prescription document handling with improved handwriting understanding rather than requiring users to submit preprocessed content . When you stack it up against frameworks that focus on specific aspects of prescription handling, our comprehensive capability collection—which includes intelligent organization, inquiry resolution functions that leverage Gemini's capabilities, anticipatory pharmaceutical tracking, and risk observation— offers significant advantages.

VI. CONCLUSION AND FUTURE WORK

This project combines OCR with Gemini's language model to manage prescriptions effectively. It extracts, analyzes, and organizes medical information automatically with strong accuracy, building on what earlier studies did [1][2]. We built it using Flask, React.js, Gemini, and MySQL, which means it can scale up, adapt to different needs, and keep data secure. It handles the whole prescription process—from uploading documents to tracking medicine use—while staying accurate and easy to use .Looking ahead, we want to spread the system across multiple servers to make it faster and more reliable. We're planning to connect it with hospital electronic health records so it fits into clinical workflows, and link it with IoT devices for monitoring in real time.

REFERENCES

[1] A. Azam, Z. Naz, and M. U. G. Khan, "PharmaLLM: A Medicine Prescriber Chatbot Exploiting Open-Source Large Language Models," Human-Centric Intelligent Systems, vol. 4,

pp. 527-544, 2024.

[2] P. Thetbanthad, B. Sathanarugsawait, and P. Praneetpolgrang, "Application of Generative Artificial Intelligence Models for Accurate Prescription Label Identification and Information Retrieval for the Elderly in Northern East of Thailand," Journal of Imaging, vol. 11, no. 1, 15, 2025.

[3] World Health Organization, "Medication Adherence: New WHO Report Provides Comprehensive Analysis," WHO Technical Report Series, no. 1015, pp. 1-112, 2023.

[4] K. Singhal, S. Azizi, T. Tu, et al., "Large language models encode clinical knowledge," Nature, vol. 620, no. 7972, pp. 172-180, Aug. 2023.

[5] S. Kumar and P. Singh, "Deep Learning Approaches for Handwritten Medical Prescription Recognition," IEEE Access, vol. 12, pp. 15234-15247, Feb. 2024.

[6] L. Osterberg and T. Blaschke, "Adherence to Medication," New England Journal of Medicine, vol. 353, no. 5, pp. 487-497, Aug. 2005.

[7] T. Brown, N. Ryder, M. Subbiah, et al., "Language Models Are Few-Shot Learners," Advances in Neural Information Processing Systems, vol. 33, pp. 1877-1901, 2020.

[8] Y. Chen, L. Zhang, and W. Liu, "CRNN Architectures for Medical Handwriting Recognition," Pattern Recognition Letters, vol. 170, pp. 45-52, Jun. 2023.

[9] S. Lee, J. Park, and K. Kim, "Visual-Language Models for Medical Image Understanding," Medical AI Review, vol. 15, no. 3, pp. 234-248, Sep. 2024.

[10] H. Wang, Z. Chen, and L. Yang, "Medical Knowledge Graphs for Prescription Understanding," Journal of Medical Systems, vol. 48, 89, Mar. 2024.