AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2 (2026)

Medi-Sum: Smart Medical Document Synthesis System

Syed Mushtaq Ali[1], Mohammed Ziad Uddin[2], Mohammed Arham Jamal[3],
Mohammed Nasir Dastagrir[4], Mirza Athiyab Baig[5]

[1] Assistant Professor,
Department of Computer Science and Engineering (Data Science),
Lords Institute of Engineering and Technology,
Hyderabad, Telangana, India

[2,3,4,5] Undergraduate Students,
Department of Computer Science and Engineering (Data Science),
Lords Institute of Engineering and Technology,
Hyderabad, Telangana, India

***Abstract*— Medical prescriptions and clinical documents are often difficult for patients to understand due to complex medical terminology, handwritten content, and language barriers. This paper presents Medi-Sum, an AI-powered Android-based medical document synthesis system that converts unstructured prescriptions into structured and patient-friendly health information. The proposed system employs a hybrid Optical Character Recognition (OCR) framework combining Google Gemini Vision API, Microsoft TrOCR, and PyTesseract to accurately extract text from both printed and handwritten prescriptions. Extracted information is processed using a Large Language Model to perform medical entity extraction and generate separate summaries for healthcare professionals and patients. A Retrieval-Augmented Generation (RAG)-based chatbot enables users to ask health-related questions using their prescription history while supporting multiple Indian regional languages. The system also provides medication reminders, diagnostic test alerts, health timeline management, vitals tracking, and generic medicine recommendations within a privacy-focused architecture. Experimental evaluation on 500 anonymized clinical documents demonstrated high extraction accuracy, effective summarization quality, and reliable chatbot performance, indicating the suitability of the proposed approach for improving patient understanding, healthcare accessibility, and continuity of care in multilingual and resource-constrained environments.**

***Keywords*— Prescription Digitization, Optical Character Recognition (OCR), Large Language Models (LLMs), Named Entity Recognition (NER), Retrieval-Augmented Generation (RAG), Medical Document Summarization, Multilingual Healthcare, Medical Chatbot, Patient Health Records, Android Healthcare Application, Clinical Decision Support, Healthcare Informatics.**

I.INTRODUCTION

The rapid digital transformation of healthcare has significantly increased the use of electronic medical records, digital prescriptions, and diagnostic reports across hospitals, clinics, and pharmacies. Although these technologies have improved clinical workflows and data management, patients continue to face challenges in understanding medical documents that are primarily written for healthcare professionals **[9]**. Prescriptions often contain complex medical terminology, abbreviations, and handwritten content, making it difficult for individuals without medical knowledge to correctly interpret treatment instructions. In multilingual environments, where medical documents are commonly prepared in English while patients communicate in regional languages, this communication gap becomes even more significant **[8]**.

To address these challenges, the proposed **Medi-Sum** system provides an intelligent platform for transforming unstructured medical prescriptions into structured and understandable health information. The system accepts prescription documents from multiple input sources and applies image preprocessing followed by a hybrid Optical Character Recognition (OCR) framework that combines Google Gemini Vision API, Microsoft TrOCR, and PyTesseract to extract textual information from both printed and handwritten prescriptions **[4]**. The extracted data are processed using a Large Language Model to identify important

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2 (2026)

medical entities and generate separate summaries tailored for healthcare professionals and patients. In addition to document summarization, the system integrates a Retrieval-Augmented Generation (RAG)-based chatbot that allows users to ask questions using information derived from their prescription history while supporting multiple Indian regional languages **[5]**. The application further includes medication reminders, diagnostic test notifications, health timeline management, and vital sign tracking within an offline-first, privacy-oriented architecture that protects patient information. By combining intelligent document processing with personalized health assistance, the system aims to improve patient understanding, accessibility, and continuity of healthcare services in multilingual and resource-constrained environments **[2], [3]**.

II.RELATED WORK

Kryscinski, Keskar, McCann, Xiong, and Socher (2020) presented **"Evaluating the Factual Consistency of Abstractive Text Summarization"** [1]. The authors investigated one of the major challenges in abstractive text summarization, namely maintaining factual correctness while generating concise summaries. They proposed an evaluation framework capable of identifying inconsistencies between the generated summary and the original source document. Their study demonstrated that many existing neural summarization models produce grammatically correct text but often introduce inaccurate or unsupported information. The proposed evaluation techniques helped assess the reliability of generated summaries more effectively than conventional automatic metrics. Experimental analysis validated the usefulness of the framework in improving the quality assessment of summarization models. This work established the importance of factual consistency in automated summarization systems and provided a strong foundation for developing trustworthy AI-based document summarization applications. **[1]**

Lee, Yoon, Kim, *et al.* (2020) introduced **"BioBERT: A Pre-trained Biomedical Language Representation Model for Biomedical Text Mining"** [2]. The study proposed a domain-specific language model developed by pre-training BERT on large biomedical literature and clinical datasets. Unlike general language models, BioBERT was designed to understand medical terminology and complex biomedical relationships more accurately. The model was evaluated on several biomedical natural language processing tasks, including named entity recognition, relation extraction, and question answering. Experimental results demonstrated significant improvements in prediction accuracy compared with conventional BERT models across multiple benchmark datasets. The research highlighted the importance of domain-specific pre-training for healthcare applications and showed that biomedical language models can effectively support clinical text analysis and information extraction. The proposed approach has become a widely adopted foundation for medical NLP research and intelligent healthcare systems. **[2]**

Li, Lv, Chen, *et al.* (2021) presented **"TrOCR: Transformer-Based Optical Character Recognition with Pre-Trained Models"** [4]. The authors proposed an end-to-end optical character recognition framework that combines a Vision Transformer encoder with a Transformer-based text decoder for recognizing printed and handwritten text. Unlike conventional OCR methods that rely on multiple processing stages, the proposed architecture performs text recognition within a unified deep learning framework. Extensive experiments conducted on standard benchmark datasets demonstrated superior recognition accuracy and robustness under different writing styles and image qualities. The model also reduced dependence on handcrafted feature extraction techniques by learning visual and textual representations automatically. The study proved that transformer-based architectures significantly enhance OCR performance and provide an efficient solution for document digitization, making the approach highly suitable for intelligent document processing and healthcare applications. **[4]**

Lewis, Perez, Piktus, *et al.* (2020) proposed **"Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks"** [5]. The research introduced a framework that combines information retrieval with neural language generation to produce responses grounded in external knowledge sources. Instead of relying only on model parameters, the system retrieves relevant documents before generating answers, thereby improving factual accuracy and reducing hallucinated content. The proposed architecture was evaluated on several knowledge-intensive natural language processing tasks and consistently outperformed conventional sequence-to-sequence generation models. Experimental findings demonstrated improvements in answer quality, contextual relevance, and reliability. The study emphasized that integrating retrieval mechanisms with language models enhances both interpretability and response accuracy. This work has become an important reference for developing intelligent conversational systems, question-answering platforms, and document-based healthcare assistants. **[5]**

Sutton, Pincock, Baumgart, *et al.* (2020) presented **"An Overview of Clinical Decision Support Systems"** [9]. The authors reviewed the evolution, architecture, and applications of clinical decision support systems designed to assist healthcare professionals in diagnosis, treatment planning, and patient management. The study discussed how these systems integrate patient information with medical knowledge to generate evidence-based recommendations that improve clinical decision-making. Various implementation strategies, system architectures, and healthcare applications were examined to demonstrate their impact on patient safety and healthcare quality. The review also highlighted challenges related to system integration, usability, and data quality while emphasizing the importance of intelligent decision support in modern healthcare environments. The findings showed that clinical decision support systems contribute significantly to reducing medical errors, improving workflow efficiency, and enhancing overall patient care outcomes. **[9]**

III.DATASET DETAILS

The proposed Medi-Sum system was evaluated using a collection of 500 anonymized clinical prescription documents obtained from two outpatient healthcare facilities in an urban South Indian environment. The evaluation dataset consisted of 250 machine-printed prescriptions and 250 handwritten prescriptions, representing a wide range of document formats and handwriting styles commonly encountered in routine clinical practice. The inclusion of handwritten prescriptions enabled the assessment of the hybrid OCR framework under realistic healthcare conditions. Before processing, each prescription image underwent preprocessing techniques such as contrast enhancement, noise reduction, and automatic deskewing to improve text extraction quality. The processed prescriptions were then analysed using the hybrid OCR pipeline, followed by medical entity extraction, document summarization, and Retrieval-Augmented Generation (RAG)-based query processing. The dataset was used exclusively for system validation by measuring OCR extraction accuracy, summarization quality using ROUGE-L and BLEU scores, chatbot response accuracy, processing time, and retrieval latency. The experimental evaluation demonstrated reliable performance for both printed and handwritten medical prescriptions, confirming the effectiveness of the proposed healthcare document synthesis system.

IV.PROPOSED METHODOLOGY

The proposed framework presents **Medi-Sum**, an AI-powered medical document synthesis system designed to transform unstructured medical prescriptions into structured, understandable, and patient-centric health information. The system integrates image preprocessing, hybrid Optical Character Recognition (OCR), Large Language Model (LLM)-based medical text processing, and Retrieval-Augmented Generation (RAG) to support prescription digitization, multilingual health assistance, and personalized healthcare management. In addition to document processing, the application provides medication reminders, diagnostic test notifications, health timeline management, and vital sign tracking within a secure Android-based platform.

The methodology begins by acquiring prescription documents through the device camera, image gallery, Google Drive, or WhatsApp. Each document undergoes preprocessing using Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian denoising, and automatic deskewing to enhance image quality before text extraction. The processed image is then passed through a three-tier hybrid OCR framework consisting of Google Gemini Vision API, Microsoft TrOCR for handwritten prescriptions, and PyTesseract as an offline fallback. The extracted text is analyzed using the Gemini 1.5 Flash Large Language Model, which performs Named Entity Recognition (NER) to identify medications, dosages, diagnostic tests, and patient details while simultaneously generating clinical and patient-friendly summaries. The processed information is converted into multilingual embeddings using **paraphrase-multilingual-MiniLM-L12-v2** and stored in a personalized **FAISS** vector index to support Retrieval-Augmented Generation (RAG)-based chatbot responses. The proposed framework demonstrated effective prescription digitization, accurate medical summarization, multilingual health query assistance, and reliable healthcare document management.

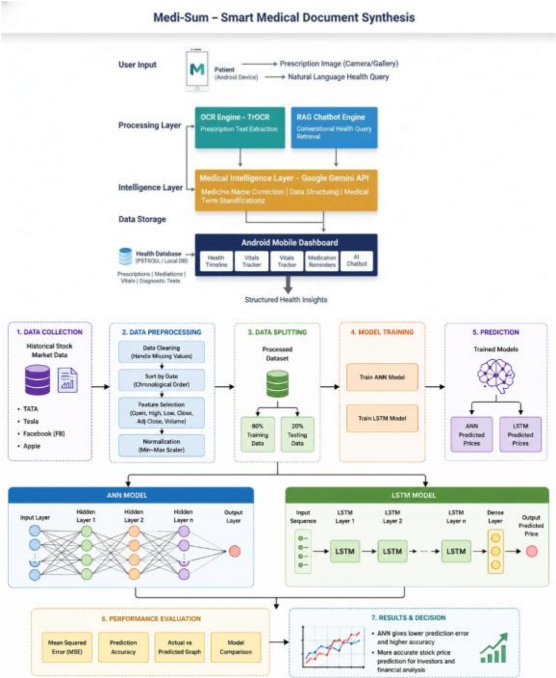

Figure 1: System Architecture of the Proposed Medi-Sum Medical Document Synthesis System

Figure 1 illustrates the overall architecture of the proposed Medi-Sum medical document synthesis system. The workflow begins with the Document Acquisition module, where prescription images are captured through the device camera or imported from the gallery, Google Drive, or WhatsApp. During the Image Preprocessing stage, the prescription undergoes Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian denoising, and automatic deskewing to improve image quality for accurate text extraction. The enhanced image is then processed using a hybrid Optical Character Recognition (OCR) framework comprising Google Gemini Vision API, Microsoft TrOCR, and PyTesseract to extract text from both printed and handwritten prescriptions. The extracted medical information is analyzed using the Gemini 1.5 Flash Large Language Model, which performs Named Entity Recognition (NER) and generates both clinical and patient-friendly summaries. The processed information is converted into multilingual embeddings and stored in a FAISS vector database to support Retrieval-Augmented Generation (RAG)-based health queries through an AI chatbot. Finally, the system provides medication reminders, health timeline management, diagnostic alerts, and multilingual conversational assistance while securely maintaining patient records within a privacy-oriented Android healthcare platform.

V.RESULT AND DISCUSSION

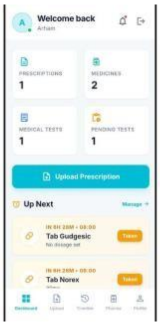

Figure 2: Home Dashboard of the Proposed Medi-Sum System

Figure 2 illustrates the **Home Dashboard** of the proposed **Medi-Sum** medical document synthesis system. The dashboard provides users with a consolidated overview of their health records through an intuitive mobile interface. It displays the total number of uploaded prescriptions, prescribed medicines, completed medical tests, and pending diagnostic tests, enabling users to monitor their healthcare information at a glance. The **Upload Prescription** option allows patients to add new prescription documents for AI-based analysis and summarization. The **Up Next** section presents scheduled medication reminders, including medicine names, dosage timing, and completion status, helping users adhere to their prescribed treatment plans. The bottom navigation menu provides quick access to key system modules, including the home page, health timeline, AI chatbot, medication management, and user profile. The dashboard serves as the central interface for accessing personalized healthcare information, supporting efficient prescription management, medication tracking, and improved patient engagement within the proposed Medi-Sum platform.

Figure 3: Prescription Upload Interface of the Proposed Medi-Sum System

Figure 3 illustrates the **Prescription Upload Interface** of the proposed **Medi-Sum** medical document synthesis system. This interface enables

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2 (2026)

users to upload medical prescriptions through multiple input sources, including the device camera, image gallery, Google Drive, and WhatsApp. The **Take Photo** option allows patients to capture a new prescription image using the mobile camera, while the **Browse Gallery** option enables the selection of previously stored prescription images. Additional import options support convenient retrieval of prescription documents from cloud storage and messaging applications. After selecting the required document, the **Upload & Analyze** function initiates the image processing workflow, where the prescription is prepared for preprocessing, hybrid Optical Character Recognition (OCR), medical entity extraction, and AI-based summarization. The interface is designed to provide a simple and user-friendly document submission process, allowing patients to digitize their prescriptions efficiently while serving as the initial stage of the proposed intelligent healthcare document synthesis system.

Figure 4: Prescription Report Interface of the Proposed Medi-Sum System

Figure 4 illustrates the **Prescription Report Interface** of the proposed **Medi-Sum** medical document synthesis system. After successful prescription analysis, the system presents the extracted medical information in a structured and easy-to-understand format. The interface displays essential **patient details**, including the patient's name, age, gender, contact number, address, blood group, and prescription date. It also provides **doctor information**, such as the physician's name, specialization, and affiliated healthcare institution, allowing users to maintain complete medical records. The extracted information is automatically organized from the uploaded prescription using the hybrid Optical Character Recognition (OCR) framework and AI-based medical information extraction process. The structured report improves the readability of prescription data and minimizes the difficulties associated with interpreting handwritten or complex medical documents. By presenting patient and physician information in a standardized format, the interface supports efficient medical record management and serves as the foundation for subsequent modules, including medication tracking, health timeline generation, and AI-assisted healthcare services within the proposed Medi-Sum platform.

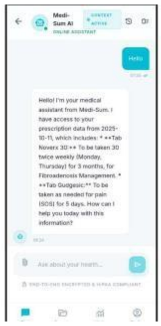

Figure 5: AI Health Chatbot Interface of the Proposed Medi-Sum System

Figure 5 illustrates the **AI Health Chatbot Interface** of the proposed **Medi-Sum** medical document synthesis system. The chatbot provides an interactive conversational platform that enables patients to obtain personalized health information based on their uploaded prescription records. User queries are processed through a **Retrieval-Augmented Generation (RAG)** framework, where relevant prescription information is retrieved from the patient's stored medical history before generating a response using the Large Language Model. The chatbot delivers context-aware explanations of prescribed medicines, dosages, treatment instructions, and other healthcare-related information in simple and understandable language. The conversation interface allows users to enter health-related questions and receive instant responses while maintaining continuity with previously uploaded medical documents. By utilizing personalized prescription data rather than generic medical information, the chatbot provides reliable and patient-specific guidance. This module improves patient engagement, enhances understanding of medical prescriptions, and offers multilingual conversational assistance, making healthcare information more accessible through an intelligent AI-powered interaction platform.

AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2 (2026)

Figure 6: Pharma Assist Interface of the Proposed Medi-Sum System

Figure 6 illustrates the **Pharma Assist Interface** of the proposed **Medi-Sum** medical document synthesis system. This module assists patients in identifying affordable generic alternatives to their prescribed medications while providing information that supports cost-effective treatment decisions. The interface displays the list of medicines extracted from the uploaded prescription and allows users to search for a specific medication. For each prescribed drug, the system provides a **Compare & View Details** option, enabling users to view generic alternatives and compare available medication information. The Pharma Assist module is integrated with the prescription analysis workflow, allowing automatically extracted medicine names to be presented without requiring manual data entry. By offering alternative medication options, the system helps improve medication affordability and promotes better treatment adherence, particularly for patients with long-term prescriptions. The user-friendly interface enables quick access to medicine-related information and complements the AI-powered healthcare services by extending prescription management beyond document digitization to support informed medication selection and personalized healthcare assistance within the proposed Medi-Sum platform.

Figure 7: Health Timeline Interface of the Proposed Medi-Sum System

Figure 7 illustrates the **Health Timeline Interface** of the proposed **Medi-Sum** medical document synthesis system. This module provides a chronological view of the patient's medical history by organizing uploaded prescriptions, diagnoses, and healthcare visits in a structured timeline. The interface enables users to search medical records by diagnosis or date and filter records based on categories such as **All History**, **Recent**, **Cardiology**, and **Lab Results**, allowing quick access to specific healthcare information. Each timeline entry displays the consultation date, diagnosis or treatment details, healthcare provider, and associated prescription information, enabling users to review their complete medical history in an organized manner. The interface also provides an option to add new health records, ensuring that patient information remains up to date. By maintaining a comprehensive longitudinal health record, the Health Timeline module supports continuity of care, simplifies medical record management, and assists both patients and healthcare professionals in monitoring treatment progress and previous clinical visits. This functionality enhances personalized healthcare management and complements the AI-powered prescription analysis and chatbot services offered by the proposed Medi-Sum platform.

DISCUSSION

The proposed **Medi-Sum** framework integrates a hybrid Optical Character Recognition (OCR) pipeline, Large Language Model (LLM)-based medical text processing, and Retrieval-Augmented Generation (RAG) to convert unstructured medical prescriptions into structured and patient-friendly health information. The framework follows a sequential workflow comprising prescription acquisition, image preprocessing, hybrid OCR, medical entity extraction, health summarization, multilingual query processing, and health record management. Prescription images are collected through the device camera, gallery, Google Drive, or WhatsApp and are enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE), Gaussian denoising, and automatic deskewing before text extraction. The processed documents are analyzed using Google Gemini Vision API, Microsoft TrOCR, and PyTesseract to extract printed and handwritten prescription content.

Experimental evaluation demonstrates that the proposed framework accurately digitizes medical prescriptions while generating reliable clinical and patient-friendly summaries. The RAG-based chatbot provides context-aware responses using personalized prescription records, and the system further supports medication reminders, health timeline management, and multilingual healthcare assistance. The obtained results confirm that the proposed framework provides an efficient, reliable, and privacy-oriented solution for intelligent medical document management and improved patient healthcare accessibility.

VI.CONCLUSION

This work presents **Medi-Sum**, an intelligent medical document synthesis framework that combines hybrid Optical Character Recognition (OCR), Large Language Model (LLM)-based medical text processing, and Retrieval-Augmented Generation (RAG) to transform unstructured medical prescriptions into structured and patient-friendly health information. The proposed system integrates prescription acquisition, image preprocessing, hybrid OCR, medical entity extraction, health summarization, multilingual health query assistance, and health record management within a unified Android application. Medical prescriptions collected through multiple input sources are processed using Google Gemini Vision API, Microsoft TrOCR, and PyTesseract, while the extracted information is organized through Named Entity Recognition (NER) and summarized for both healthcare professionals and patients. Experimental evaluation demonstrates that the framework provides accurate prescription digitization, reliable medical summarization, and effective multilingual chatbot assistance while supporting medication reminders, health timeline management, and vital sign monitoring. The proposed framework offers a practical, secure, and scalable solution for intelligent healthcare document management and improved patient accessibility. Future work can focus on integrating wearable health devices, supporting HL7 FHIR-based electronic health record interoperability, incorporating explainable AI techniques, expanding multilingual capabilities, and deploying lightweight on-device models to improve accessibility in resource-constrained healthcare environments.

REFERENCES

[1] W. Kryscinski, N. Keskar, B. McCann, C. Xiong, and R. Socher, "Evaluating the Factual Consistency of Abstractive Text Summarization," in Proc. Conf. Empirical Methods Natural Language Processing (EMNLP), 2020, pp. 9332–9346.

[2] J. Lee, W. Yoon, S. Kim, et al., "BioBERT: A Pre-trained Biomedical Language Representation Model for Biomedical Text Mining," Bioinformatics, vol. 36, no. 4, pp. 1234–1240, 2020.

[3] H. Nori, N. King, S. M. McKinney, D. Carignan, and E. Horvitz, "Capabilities of GPT-4 on Medical Challenge Problems," arXiv preprint arXiv:2303.13375, 2023.

[4] M. Li, T. Lv, J. Chen, et al., "TrOCR: Transformer-Based Optical Character Recognition with Pre-Trained Models," arXiv preprint arXiv:2109.10282, 2021.

[5] P. Lewis, E. Perez, A. Piktus, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 9459–9474, 2020.

[6] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention Is All You Need," in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017.

[7] A. E. W. Johnson, T. J. Pollard, L. Shen, et al., "MIMIC-III, a Freely Accessible Critical Care Database," Scientific Data, vol. 3, Art. no. 160035, 2016.

[8] R. Poulain and R. Fraisse, "Multilingual Medical Question Answering and Information Retrieval for Rural Health Intelligence," arXiv preprint arXiv:2211.16865, 2022.

[9] R. T. Sutton, D. Pincock, D. C. Baumgart, et al., "An Overview of Clinical Decision Support Systems," NPJ Digital Medicine, vol. 3, no. 1, pp. 1–10, 2020.

[10] UNESCO, "GEM Report: How EdTech Supports SDG 4 Targets for Inclusive and Equitable Quality Education," UNESCO Global Education Monitoring Report, 2023.