MediVA AI: A Smart, Secure Doctor Network for Accessible Healthcare Using Deep Learning, RAG, and Blockchain

Thatiparthi Karthik Reddy, Thota Akshith Somnath Mishra, Suresh Chauhan
Department of Computer Science and Engineering
Bharath Institute of Higher Education and Research
Chennai, India

Ms. R. Femimol
Assistant Professor
Department of Computer Science and Engineering
Bharath Institute of Higher Education and Research
Chennai, India

Abstract—*Access to specialized medical care is uneven across the world. Rural populations, non-English speakers, and low-income patients routinely face delays that worsen outcomes. This paper describes MediVA AI, a web-based healthcare platform that pairs a conversational AI assistant with a secure doctor scheduling backend. A patient first interacts with the AI assistant, which handles preliminary symptom analysis and generates actionable guidance. When the assistant cannot resolve a case, it schedules a consultation with a verified doctor and forwards a conversation summary so the doctor arrives prepared. Medical images and confidential reports are encrypted using a reversible deep learning model, then stored on a blockchain via IPFS hashing—so records are tamper-evident and retrievable without depending on a single server. The system uses Retrieval-Augmented Generation (RAG) with a vector database for context-aware medical responses, a SQL database for structured patient and appointment data, and a NoSQL store for unstructured media. Multilingual support, voice input, and report-upload features make the platform usable across demographics. We describe the architecture, the encryption pipeline, the scheduling logic, and the doctor-facing dashboard. Early testing shows the assistant handles common queries accurately and escalates appropriately when confidence is low.*

Keywords—*AI healthcare; RAG; blockchain; medical image encryption; doctor network; voice assistant; deep learning; IPFS; telemedicine*

I. INTRODUCTION

Healthcare delivery is slow, expensive, and unevenly distributed. A patient in a small town seeking a second opinion on a dermatology complaint may wait weeks for a specialist—if one is accessible at all. Language compounds the problem: a patient who speaks only Tamil or Hindi gains little from an English-only consultation interface. Even when patients do see a doctor, fragmented records force physicians to start from scratch.

MediVA AI is built around one practical idea: handle routine cases automatically, escalate the complex ones with full context. The AI assistant covers symptom queries, medication reminders, basic report interpretation, and preliminary triage. When a query falls outside what it can reliably handle, the system books a video consultation with a verified doctor and forwards a summary of everything discussed. The doctor does not re-interview the patient from zero.

Security is treated as a first-order requirement. Medical images and lab reports are encrypted using a reversible neural network before leaving the patient's device, stored via IPFS hashing on a blockchain ledger, and retrieved only through verified keys. Records are auditable, tamper-resistant, and recoverable even if a server fails.

The remainder of this paper is structured as follows: Section II states the problem, Section III surveys relevant prior work, Section IV describes system architecture, Section V covers the encryption and blockchain pipeline, Section VI details the AI assistant module, Section VII presents results, and Section VIII concludes with future directions.

C. Scheduling and Doctor Backend

When escalation is triggered, the system generates a structured conversation summary—chief complaint, duration, associated symptoms, relevant uploads—and inserts it into the scheduling queue. The patient selects a time slot from available doctors filtered by specialty. The doctor's dashboard shows this summary before the call begins. After the call, the doctor can add notes, update the patient's record, and issue a digital prescription stored in the same encrypted pipeline.

D. Database Design

A SQL database stores structured data: patient profiles, appointment records, doctor availability, and prescription histories. A NoSQL document store holds unstructured data: chat logs, voice transcripts, and session metadata. Medical images and sensitive reports are processed through the encryption pipeline; the application database records only their IPFS hash—the actual file never resides in the application database.

V. ENCRYPTION AND BLOCKCHAIN STORAGE

Protecting medical images requires more than TLS encryption in transit. A compromised server should not expose readable patient data. MediVA uses a reversible neural network to encrypt images at upload time.

A. Reversible Neural Network Encryption

The model takes a medical image as input and produces an encrypted output paired with a unique encryption key. The network is invertible: given the encrypted image and the correct key, it reconstructs the original with no perceptible

II. PROBLEM STATEMENT

Four interconnected problems motivate this work:

- Specialist access gaps. Rural and underserved patients face long wait times and high travel costs to reach specialists. A voice-enabled AI assistant that can triage and answer common queries addresses this directly.
- Language barriers. Medical platforms are predominantly English-centric. Non-English speakers rely on informal translators or go without. Real-time multilingual support removes this barrier.
- Data fragmentation. Prescriptions, lab results, and imaging files are scattered across providers and formats. MediVA stores all records in one encrypted location with patient-controlled access.
- Slow preliminary triage. The typical path—symptom appears, patient searches online, finds alarming or irrelevant results, eventually books an appointment—is inefficient. A reliable AI assistant that provides structured preliminary guidance shortens this loop considerably.

III. LITERATURE SURVEY

Several threads of prior work inform MediVA's design.

Rajpurkar et al. [7] survey the state of AI in clinical settings and identify a persistent generalization problem: most AI tools perform well on benchmark datasets but struggle when deployed outside controlled conditions. Demographic variation, imaging equipment differences, and clinical workflow heterogeneity all degrade real-world performance.

Lee et al. [4] introduced BioBERT, a biomedical language model pre-trained on PubMed abstracts and clinical notes. Later work combining BioBERT with FAISS vector search demonstrated high retrieval accuracy on symptom-driven queries. The approach lacks multimodal support, however—it cannot reason over an uploaded image alongside a textual complaint.

Nagendran et al. [5] conducted a systematic review of deep learning studies comparing AI performance to clinicians. They found that most studies had design flaws that inflated reported accuracy, and that few looked at patient-facing interfaces at all. Most AI diagnostic tools are designed for radiologists, not for patients navigating the system themselves.

On the security side, IPFS-backed storage with hash verification on a distributed ledger provides tamper evidence and decentralized access control without depending on a single trusted server—an appropriate architecture for sensitive data shared across providers.

Existing commercial telemedicine platforms (Practo, Apollo 247, DocsApp) provide doctor consultations but include no preliminary AI assistant. The handoff context

quality loss. The encryption resists common attacks—noise injection, cropping, and brute force—because the key space is large and the transformation is nonlinear.

This approach differs from standard symmetric encryption (AES) in one important respect: the encrypted output is an image-shaped tensor rather than a binary blob. This means it can be stored in image databases without special handling, and the decryption process is differentiable, which opens the door to future work on learning-based watermarking.

B. IPFS and Blockchain Storage

After encryption, each file is pinned to the InterPlanetary File System (IPFS), which returns a content-addressable hash. That hash is recorded on a blockchain ledger along with the patient's anonymized identifier, timestamp, and access control policy. The actual file never touches the blockchain—only its hash does.

Retrieval works in reverse: the application queries the blockchain for the IPFS hash, fetches the encrypted file from IPFS, and decrypts it using the authorized party's key. An unauthorized party who intercepts the IPFS file obtains an encrypted tensor they cannot read without the key. The blockchain record also functions as an audit trail, logging every access event immutably—relevant for both patient privacy compliance and medico-legal record-keeping.

VI. AI DOCTOR MODULE

A. RAG Pipeline

At ingestion time, medical reference texts are chunked, embedded using a biomedical sentence encoder, and stored in a vector database. At query time, the patient's message is encoded the same way and used to retrieve the top-k most similar chunks. These chunks are prepended to the prompt sent to the language model, which generates a response grounded in retrieved material.

The vector database also maintains a rolling window of the current session's conversation, so the assistant can answer follow-up questions without the patient repeating context. When a patient uploads a lab report or image, the system extracts key values and adds them to the session context as structured text.

B. Escalation Logic

Escalation is triggered by three conditions: the model's confidence score falls below a defined threshold; the query contains a keyword from a high-risk symptom list (chest pain, difficulty breathing, neurological symptoms); or the patient explicitly requests a doctor. In any of these cases, the assistant does not attempt to answer but offers to schedule a consultation and generates the handoff summary.

C. Multimodal Input

Text and voice are the primary input channels. The assistant can also process uploaded images. A skin lesion problem—ensuring the doctor knows what the patient already discussed—is unsolved by any of these services.

IV. SYSTEM ARCHITECTURE

MediVA is organized around three user roles: patients, the AI assistant, and doctors. Patients interact through a web interface that accepts text chat, voice input, and file upload. The AI assistant handles the conversation and mediates escalation. Doctors access a separate dashboard to review scheduled consultations, read pre-generated summaries, and join video calls.

A. Patient Frontend

The patient interface has three main components. The chat and voice module accepts free-text input or spoken queries, transcribed in real time. The file upload module handles prescriptions, lab reports, and medical images—each encrypted client-side before transmission. The report context module ensures previously uploaded files remain available to the AI across sessions, so the assistant can reference a blood panel from two weeks prior when answering a new question.

B. AI Assistant

The assistant uses Retrieval-Augmented Generation (RAG). A vector database holds embeddings of medical reference material, anonymized case summaries, and drug information. When a patient sends a query, the system retrieves the most relevant chunks and provides them as context to the language model. This approach grounds responses in factual material rather than relying on the model's parametric memory alone.

The assistant also evaluates its own confidence. Queries that involve symptoms outside its training distribution, require physical examination, or involve dosing decisions are flagged for escalation rather than answered with fabricated confidence.

photo, for example, passes through a CNN-based classifier that produces a preliminary assessment (normal, possibly concerning, or recommend specialist review). The classifier output is included in the assistant's context, not shown directly to the patient as a diagnosis.

VII. RESULTS AND DISCUSSION

The system was evaluated across three dimensions: assistant accuracy, encryption performance, and scheduling reliability.

For the AI assistant, 200 queries were drawn from common outpatient scenarios—fever, skin rashes, musculoskeletal complaints, and medication questions. The assistant provided guidance consistent with a clinician reviewer's assessment on 78% of queries. On the remaining 22%, the assistant either escalated correctly (18%) or provided a partially correct answer requiring qualification (4%). No cases produced clearly harmful advice.

The encryption pipeline adds approximately 340 ms of latency per image on consumer hardware—acceptable for a healthcare context where images are not streamed in real time. Decryption adds a comparable delay. Peak signal-to-noise ratio (PSNR) for reconstructed images averaged 42.3 dB across the test set, which is indistinguishable from the original at diagnostic quality.

The scheduling module was tested in a simulated environment with 50 patients and 10 doctors over a one-week window. All escalated cases received a confirmed appointment within the same session. Nine out of ten doctors reported that the pre-call summary reduced the time needed to understand the patient's situation.

VIII. CONCLUSION AND FUTURE SCOPE

MediVA AI addresses a real gap: patients who cannot easily access specialized care, and doctors who receive patients without useful prior context. The platform handles routine queries automatically, escalates sensibly, and moves context with the patient through each handoff. The blockchain-backed encryption pipeline keeps sensitive records private and auditable.

The AI assistant performs well on common outpatient queries but has not been tested on complex multi-system diseases or rare conditions. The encryption model was validated on a limited image set; production deployment would require broader validation across imaging modalities.

Near-term priorities include expanding the RAG knowledge base with specialty-specific literature, adding wearable device integration so vital signs feed automatically into the assistant's context, and running a formal clinical trial with a partner hospital to measure first-contact resolution rates. Longer term, the same encrypted storage and scheduling architecture could support dental, mental health, or physiotherapy workflows with modest modifications to the AI module.

REFERENCES

[1] G. Hinton, "Deep Learning—A Technology With the Potential to Transform Health Care," JAMA, vol. 320, no. 11, pp. 1101–1102, 2018.

[2] S. Jha, "Adapting to Artificial Intelligence: Radiologists and Pathologists as Information Specialists," JAMA, vol. 316, no. 22, pp. 2353–2354, 2016.

[3] A. Rajkomar et al., "Machine Learning in Medicine," New England Journal of Medicine, vol. 380, pp. 1347–1358, 2019.

[4] J. Lee et al., "BioBERT: A pre-trained biomedical language representation model for biomedical text mining," Bioinformatics, vol. 36, no. 4, pp. 1234–1240, 2020.

[5] M. Nagendran et al., "Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies," BMJ, vol. 368, 2020.

[6] E. Neri et al., "What the radiologist should know about artificial intelligence—An ESR white paper," Insights into Imaging, vol. 10, no. 1, 2019.

[7] P. Rajpurkar et al., "AI in healthcare: The hope, the hype, the promise, the peril," PLOS Medicine, vol. 19, no. 4, 2022.

[8] M. Huisman et al., "An international survey on AI in radiology in 1041 radiologists and radiology residents," European Radiology, vol. 31, pp. 866–879, 2021.

[9] H. Suresh and J. Guttag, "A framework for understanding sources of harm throughout the machine learning life cycle," ACM FAccT, 2021.

[10] A. Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems, vol. 30, 2017.