

MedSage: A Multimodal and Multilingual AI Platform for Symptom Analysis and Healthcare Assistance

Author: **Dodda Surya Tejaswi**¹ (MCA student), Ms. **V. Prasanna Lakshmi**² (Assistant Professor) Master Of Computer Applications, GIET ENGINEERING COLLEGE (A), Rajahmundry, AP.

Corresponding Author: Dodda Surya Tejaswi

(email-id: suryatejaswi26@gmail.com)

ABSTRACT: In today's digital world, healthcare accessibility remains a major challenge, especially in rural and underserved areas where medical professionals are not easily available. Many people depend on unreliable internet searches to understand symptoms and health conditions, which often leads to misinformation and delayed medical attention.

To address this problem, the proposed system MedSage is developed as a multimodal and multilingual AI-powered healthcare assistance platform. The system combines technologies such as Natural Language Processing (NLP), Retrieval-Augmented Generation (RAG), Computer Vision, and Speech Processing to provide intelligent healthcare support through text, voice, and medical image analysis.

The platform allows users to describe symptoms through text or speech, upload medical images such as X-rays, CT scans, and MRI scans, and receive structured healthcare insights. The system also supports multilingual voice interaction, making it accessible to users from different language backgrounds.

The symptom analysis chatbot uses NLP and RAG techniques to retrieve relevant medical information and generate context-aware responses. The medical image analysis module detects possible abnormalities and estimates severity using computer vision techniques. The voice interaction module enables speech-based communication using speech recognition technology.

The proposed system improves healthcare awareness, reduces dependency on manual symptom searching, and provides an intelligent platform for early-stage health assistance. Although the system does not replace professional medical diagnosis, it acts as a supportive healthcare guidance tool for users.

KEYWORDS: Artificial Intelligence, Healthcare Assistance, NLP, RAG, Computer Vision, Medical Image Analysis, Speech Processing, Multimodal AI, Multilingual System, FAISS.

I. INTRODUCTION

Healthcare accessibility is one of the major problems faced by people in many regions, particularly in rural and remote areas where medical resources and professionals are limited. Many individuals rely on internet searches and self-research to understand symptoms and health conditions. However, such

information is often inaccurate, misleading, or difficult to understand.

Existing healthcare chatbot systems are mostly rule-based and support only text-based interaction. These systems cannot efficiently process medical images, understand complex medical queries, or support multilingual communication. As a result, they provide limited assistance and poor user experience.

To overcome these limitations, the proposed system **MedSage** is developed as an AI-driven healthcare assistance platform. The system integrates multiple artificial intelligence technologies into a single platform to provide better healthcare accessibility and early health awareness.

The MedSage platform consists of:

- A symptom analysis chatbot using NLP and RAG
- A medical image analysis system using Computer Vision
- A multilingual speech interaction module using Speech AI

The system supports text, voice, and image inputs and generates structured healthcare responses for users. The main objective of the project is to improve healthcare accessibility, simplify symptom analysis, and provide intelligent healthcare assistance using AI technologies.

II. Literature Survey:

2.1 Healthcare Chatbot Systems

Healthcare chatbot systems are widely used to provide basic medical guidance and symptom-based assistance to users. Earlier systems mainly depended on rule-based techniques and keyword matching methods. Although these systems could answer simple queries, they lacked contextual understanding and intelligent reasoning capabilities.

Modern healthcare chatbots use Artificial Intelligence and Natural Language Processing techniques to improve response quality and user interaction. These intelligent systems can understand user symptoms more effectively and provide context-aware healthcare suggestions.

2.2 Medical Image Analysis

Medical image analysis systems use Computer Vision and Deep Learning techniques to analyze medical scans such as X-rays, CT scans, and MRI images. These systems help identify abnormalities, disease patterns, and severity levels automatically.

Recent developments in AI-based medical imaging have improved disease detection accuracy and reduced the dependency on manual analysis. Heatmaps and visual indicators are also used to highlight affected regions in medical scans.

2.3 Speech Recognition Technology

Speech Recognition technology converts spoken language into text automatically. It is widely used in virtual assistants, smart applications, and multilingual communication systems.

In healthcare systems, speech-based interaction improves accessibility for elderly users and people who may face difficulty typing symptoms manually. Speech processing also enables multilingual healthcare interaction.

2.4 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation is an advanced AI technique used to improve response accuracy in intelligent systems. Instead of generating responses only from pretrained knowledge, RAG retrieves relevant information from a structured knowledge base before generating answers.

This approach reduces incorrect responses and improves contextual healthcare guidance. FAISS vector similarity search is commonly used for efficient retrieval of medical information in RAG-based systems.

III. System Analysis And Design

3.1 Existing System

Most existing healthcare systems are rule-based chatbot systems that mainly depend on keyword matching and predefined decision trees. These systems are limited in handling complex medical queries and cannot provide intelligent clinical reasoning.

Existing systems mainly support text-based interaction and lack support for medical image analysis and multilingual voice communication. They also require frequent manual updates and cannot efficiently scale for modern healthcare requirements.

Limitations of Existing System

- Limited to text-based communication
- No support for medical image analysis
- Lack of multilingual interaction
- No intelligent contextual reasoning
- Static and outdated knowledge bases

3.2 Proposed System

To overcome the limitations of traditional healthcare systems, the proposed MedSage platform integrates multiple AI technologies into a unified healthcare assistance system.

The system supports:

- Symptom analysis using NLP and RAG
- Medical image analysis using Computer Vision
- Voice interaction using Speech Recognition
- Multilingual healthcare communication
- Structured healthcare report generation

The system accepts text, voice, and image inputs and generates intelligent healthcare insights using AI-based processing techniques.

Advantages of Proposed System

- Provides 24/7 healthcare assistance
- Supports text, voice, and image interaction
- Detects abnormalities in medical images
- Supports multilingual communication
- Uses RAG for accurate healthcare responses
- Improves early health awareness
- Reduces dependency on unreliable symptom searches

3.3 System Architecture

The architecture of MedSage is designed using multiple integrated modules for efficient processing and intelligent healthcare assistance.

The system architecture consists of the following modules:

1. User Interface Module

This module allows users to enter symptoms and upload medical images easily. It shows reports and healthcare results in a simple way.

2. Input Processing Module

This module processes text, voice, and image inputs from users. It prepares the data for analysis by other modules.

3. Symptom Analysis Module

This module analyzes user symptoms using NLP technology. It identifies possible health conditions from the symptoms.

4. Medical Image Analysis Module

This module checks X-rays, CT scans, and MRI images. It detects abnormalities and shows affected areas.

5. Speech Processing Module

This module converts voice input into text. It helps users interact with the system through speech.

6. Knowledge Retrieval Module

This module retrieves medical information from the database. It provides accurate and relevant healthcare guidance.

7. Backend and Database Module

This module manages system operations and stores user data. It connects all modules and handles data flow.

8. Response Generation Module

This module generates healthcare reports and responses. It provides recommendations and health-related suggestions.

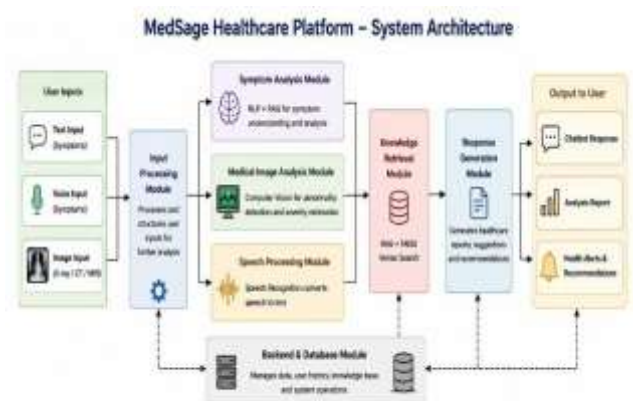


Fig 1. System Architecture of MedSage Healthcare Platform

IV. PROPOSED METHODOLOGY

The proposed methodology explains the working process of the MedSage healthcare platform and how different AI technologies are integrated into a single system.

The platform processes multiple forms of user input including text, speech, and medical images.

4.1 Symptom Analysis

In the symptom analysis process, users enter symptoms through text or voice input. The NLP module processes the input and extracts important medical information.

The RAG-based retrieval system searches the medical knowledge base using FAISS similarity search and retrieves relevant healthcare information. Based on the retrieved context, the chatbot generates structured healthcare guidance.

4.2 Medical Image Analysis

The medical image analysis module processes uploaded X-rays, CT scans, and MRI scans using Computer Vision techniques.

The system detects abnormalities, estimates severity levels, and generates heatmaps to highlight affected regions in the scan images. The module also supports before-and-after scan comparison for visual analysis.



Fig 2. Medical Image Analysis and Comparison Module

4.3 Speech Processing

The speech processing module allows users to communicate through voice interaction. Spoken language is converted into text using speech recognition techniques.

The processed text is then forwarded to the NLP module for symptom analysis and healthcare response generation.

This module improves accessibility for elderly users and users who face difficulty typing symptoms manually.

4.4 Knowledge Retrieval

The knowledge retrieval module uses Retrieval-Augmented Generation techniques to improve response quality and accuracy.

Relevant medical information is retrieved from the vector database using semantic similarity search. FAISS is used to perform fast and efficient vector retrieval operations.

This module helps generate reliable and context-aware healthcare guidance.

V. IMPLEMENTATION

5.1 Technologies Used

The MedSage platform is developed using modern AI technologies and software frameworks.

Technologies Used

- Python – Main programming language
- React – Frontend development
- Flask – Backend framework
- NLP – Symptom analysis
- RAG – Context-aware response generation
- FAISS – Semantic similarity search
- Speech Recognition – Voice interaction
- Computer Vision – Medical image analysis

These technologies work together to provide an intelligent multimodal healthcare assistance platform.

5.2 Modules

User Interface Module

Provides a user-friendly interface for symptom input, image upload, and healthcare interaction.

Symptom Analysis Module

Processes symptom text using NLP techniques and identifies possible health conditions.

Medical Image Analysis Module

Analyzes medical scans and detects abnormalities using computer vision methods.

Speech Processing Module

Converts speech into text and enables multilingual voice communication.

Knowledge Retrieval Module

Retrieves relevant healthcare information using RAG and FAISS vector search.

Backend and Database Module

Handles system communication, API integration, and data storage.

Response Generation Module

Generates structured healthcare reports and recommendations.

The system was capable of understanding symptom related queries and providing context-aware responses.

The medical image analysis module successfully detected abnormalities from uploaded scans and generated severity heatmaps for visual interpretation.

The speech processing module accurately converted spoken symptoms into text and enabled smooth voice-based interaction.

The graphical user interface provided a simple and user-friendly healthcare interaction experience.

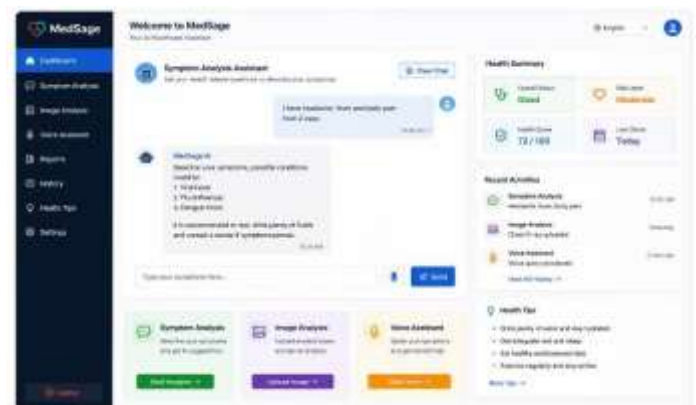


Fig 3: User Interface of MedSage Healthcare Platform

Overall, the proposed system demonstrated effective multimodal healthcare assistance capabilities by integrating AI-based symptom analysis, medical image evaluation, and multilingual speech interaction into a single intelligent platform.

VI. RESULTS AND ANALYSIS

The MedSage platform was successfully tested using different healthcare input formats including text voice, and medical image inputs

The symptom analysis chatbot generated structured healthcare guidance using NLP and RAG-based retrieval techniques.

VII. CONCLUSION

The MedSage healthcare platform successfully combines Artificial Intelligence, Natural Language Processing, Retrieval-Augmented Generation, Computer Vision, and Speech Processing into a single multimodal healthcare assistance system.

The platform improves healthcare accessibility by supporting text, voice, and medical image-based interaction. It helps users understand symptoms, analyze medical scans, and receive structured healthcare guidance more effectively.

The implemented system demonstrated reliable healthcare assistance capabilities while reducing dependency on unreliable symptom searching methods. The multilingual speech interaction module also improved accessibility for users from different language backgrounds.

Although the system does not replace professional medical diagnosis, it can support early health awareness, symptom understanding, and healthcare decision assistance.

In the future, the system can be enhanced with:

- Real-time doctor consultation
- Cloud-based medical record integration
- Advanced disease prediction models
- Mobile healthcare application support
- Improved AI-based clinical reasoning

Overall, the project demonstrates how multimodal AI technologies can be effectively used to build intelligent healthcare assistance platforms for modern digital healthcare environments.

VIII. REFERENCES

- Shashi Kant Ojha, Abhishek Kumar, Tanvi Bhole, Sheenam Naaz, "Rule-Based AI Chatbot," IJIRT, Vol. 11, Issue 5, 2024.
- X. Zhang et al., "Explainable Deep Learning in Medical Imaging: Challenges and Opportunities," IEEE Reviews in Biomedical Engineering, 2025.
- Chen et al., "Multimodal Deep Learning for Healthcare: A Comprehensive Review," IEEE Access, 2023.
- H. Yang et al., "Medical Chatbots for Patient Interaction: Current Trends and Challenges," IEEE Access, 2023.