
# MediAssist: A Multimodal AI-Powered Healthcare Assistant for Comprehensive Medical Consultation and Diagnosis Support

**Syed Awais [1*], Ipsita Majhi A [2], Mohd Abdul Hadi [3], Dr. Ruhiat Sultana [4]**

[1-3] Department of Computer Science and Engineering, Lords Institute of Engineering & Technology(A), Osmania University, Hyderabad, Telangana, India

[4] Ph.D., Associate Professor, Department of Computer Science and Engineering, Lords Institute of Engineering & Technology(A), Osmania University, Hyderabad, Telangana, India

* Corresponding Author: **Syed Awais**

## Article Info

**P-ISSN:** 3051-3383
**E-ISSN:** 3051-3391
**Volume:** 06
**Issue:** 02
**July - December 2025**
**Received:** 16-10-2025
**Accepted:** 19-11-2025
**Published:** 13-12-2025
**Page No:** 170-178

**Abstract**

The integration of artificial intelligence in healthcare has revolutionized patient care delivery and medical diagnostics. This study presents MediAssist, a comprehensive multimodal AI-powered healthcare assistant designed to provide intelligent medical consultation, symptom analysis, and preliminary diagnostic support. The system leverages advanced large language models, computer vision, speech recognition, and natural language processing to create an interactive platform capable of processing text, voice, and image inputs. MediAssist incorporates real-time symptom analysis, disease prediction, medication information retrieval, and appointment scheduling functionalities. The implementation utilizes Google's Gemini API for conversational AI, OpenCV for image processing, and speech recognition libraries for voice interaction. Performance evaluation demonstrates 92% accuracy in symptom recognition, 88% precision in disease prediction, and 94% user satisfaction rate. The system achieved an average response time of 1.8 seconds and successfully processed multimodal inputs with 90% reliability. This research contributes to the growing field of telemedicine by providing an accessible, user-friendly platform that bridges the gap between patients and healthcare services, particularly beneficial for underserved populations and preliminary health assessments.

**DOI:** https://doi.org/10.54660/IJAIET.2025.6.2.170-178

**Keywords:** Artificial Intelligence, Healthcare Chatbot, Multimodal AI, Large language Models, Medical Diagnosis, Telemedicine, Computer Vision, Natural Language Processing, Speech Recognition, Patient Care

## Introduction

The healthcare industry faces unprecedented challenges in providing timely, accessible, and quality medical services to growing populations worldwide. Recent estimates indicate that approximately 400 million people globally lack access to essential healthcare services, while healthcare systems in developed nations struggle with overcrowding, long waiting times, and resource constraints [1, 2]. The COVID-19 pandemic further exposed vulnerabilities in traditional healthcare delivery models, accelerating the adoption of digital health technologies and telemedicine platforms [3].

Artificial intelligence has emerged as a transformative force in addressing these healthcare challenges. AI-powered systems demonstrate remarkable capabilities in medical image analysis, disease prediction, drug discovery, and patient monitoring [4, 5]. Large language models, particularly conversational AI systems, have shown promising results in understanding medical queries, providing health information, and assisting in preliminary symptom assessment [6]. The convergence of multiple AI technologies—including natural language processing, computer vision, and speech recognition—enables the development of sophisticated multimodal healthcare assistants capable of mimicking human-like interactions while processing diverse data types [7, 8].

Current healthcare delivery systems exhibit several critical limitations. Patients often face difficulties in accessing immediate medical advice, understanding symptoms, or determining when professional consultation is necessary. Emergency departments are frequently overwhelmed with non-urgent cases, while rural and underserved areas suffer from physician shortages [9]. Additionally, language barriers, health literacy gaps, and socioeconomic factors create disparities in healthcare access and outcomes [10]. These challenges necessitate innovative solutions that can provide preliminary medical guidance, triage support, and health education while complementing rather than replacing professional medical care.

The development of intelligent healthcare assistants represents a significant opportunity to enhance medical accessibility and patient engagement. However, existing medical chatbots often suffer from limitations including unimodal interaction, lack of context awareness, insufficient integration of visual diagnostics, and limited personalization capabilities [11, 12]. This research addresses these gaps by developing a comprehensive multimodal AI system that integrates conversational AI, image analysis, voice interaction, and health information management within a unified platform.

**Problem Statement**

The contemporary healthcare landscape presents several interconnected challenges that impede efficient patient care delivery. Primary among these is the limited accessibility to immediate medical consultation, particularly in rural areas, developing nations, and during non-business hours. Patients frequently struggle to distinguish between minor ailments requiring self-care and serious conditions necessitating urgent medical attention, leading to inappropriate utilization of emergency services or dangerous delays in seeking care [13].

Existing digital health solutions demonstrate significant limitations in their functional scope and user experience. Most medical chatbots operate exclusively through text-based interfaces, failing to accommodate users with literacy challenges, visual impairments, or preference for verbal communication [14]. The absence of multimodal interaction capabilities restricts the system's ability to process medical images such as skin lesions, X-rays, or wound photographs, which are often crucial for preliminary assessment. Furthermore, current systems lack comprehensive integration of essential healthcare functionalities including appointment scheduling, medication management, and personalized health record maintenance.

The fragmentation of healthcare information presents another critical challenge. Patients often receive medical advice from multiple sources with varying reliability, leading to confusion, misinformation, and potentially harmful self-medication practices [15]. There exists a pressing need for intelligent systems capable of providing evidence-based health information, personalized symptom analysis, and appropriate guidance regarding the necessity and urgency of professional medical consultation. Additionally, language barriers and inadequate health literacy significantly impair patients' ability to communicate symptoms effectively and understand medical instructions, particularly among elderly populations and non-native speakers [16].

**Objectives of the Study**

The primary objective of this research is to design, develop, and evaluate a comprehensive multimodal AI-powered healthcare assistant capable of providing intelligent medical consultation, preliminary diagnosis support, and integrated health management services. The specific objectives are:

1. To develop a multimodal interaction framework that seamlessly integrates text-based, voice-activated, and image-based input mechanisms, enabling diverse user populations to access healthcare information through their preferred communication modality.
2. To implement intelligent symptom analysis and disease prediction algorithms utilizing advanced large language models and machine learning techniques for preliminary diagnostic assessments and appropriate healthcare recommendations.
3. To create a computer vision module capable of analyzing medical images including skin conditions and wound photographs, providing initial assessment and severity classification.
4. To establish an integrated health information system encompassing medication databases, appointment scheduling functionality, health record management, and emergency response protocols.
5. To evaluate system performance across multiple dimensions including accuracy of symptom recognition, precision of disease prediction, response time, user satisfaction, and reliability of multimodal input processing.

**Literature Review**

The application of artificial intelligence in healthcare has evolved significantly over the past decade, with numerous studies demonstrating the potential of AI-powered systems in clinical decision support, medical diagnostics, and patient care management. The advent of machine learning and deep learning technologies revolutionized medical AI applications. Convolutional neural networks have demonstrated remarkable performance in medical image analysis, achieving accuracy rates comparable to or exceeding human experts in specific domains such as diabetic retinopathy detection and skin cancer classification [17, 18].

Natural language processing applications in healthcare have progressed substantially with the development of transformer-based architectures and large language models. BERT-based models fine-tuned on medical corpora have shown superior performance in clinical named entity recognition, medical question answering, and clinical documentation tasks [19]. Research has demonstrated that GPT-based models could generate clinically relevant responses to patient queries with appropriate medical terminology and context awareness [20]. However, concerns regarding hallucinations, factual accuracy, and the potential for providing incorrect medical advice remain significant challenges requiring careful system design and validation.

Medical chatbot development has attracted considerable research attention, with various implementations targeting specific healthcare domains. Several studies have reported symptom checker applications utilizing decision tree algorithms, Bayesian networks, or rule-based systems [21, 22]. However, comparative evaluations revealed significant variations in diagnostic accuracy, with many systems demonstrating poor performance in triage recommendations. A systematic review identified that while conversational agents showed promise in health behavior change interventions, most lacked rigorous clinical validation and long-term efficacy data [23].

Multimodal AI systems integrating text, speech, and visual modalities represent an emerging frontier in healthcare applications. Research has explored the integration of speech recognition with medical chatbots, demonstrating improved accessibility for elderly users and individuals with motor impairments [24]. Computer vision integration in telemedicine platforms has been investigated for dermatological consultations and wound assessment [25]. Studies indicate that multimodal approaches can enhance diagnostic accuracy by leveraging complementary information sources.

The application of large language models in medical contexts has generated substantial recent interest. OpenAI's GPT-4 and Google's Med-PaLM have demonstrated impressive capabilities in medical licensing exam questions and clinical reasoning tasks [26, 27]. However, critical analyses highlight important limitations including inconsistent performance across medical specialties and potential for generating plausible but incorrect information [28].

Despite these advances, several research gaps persist in current healthcare AI literature. Limited research addresses the comprehensive integration of multiple AI modalities within unified healthcare assistant platforms. Most studies focus on isolated capabilities rather than end-to-end systems encompassing consultation, diagnosis support, health record management, and care coordination. The present research addresses these gaps by developing and evaluating a comprehensive multimodal AI healthcare assistant with integrated functionalities and rigorous performance assessment.

## System Architecture and Methodology

The MediAssist system architecture employs a modular, layered design approach that facilitates scalability, maintainability, and extensibility. The architecture comprises four primary layers: the presentation layer, application layer, AI processing layer, and data management layer, as illustrated in Figure 1.

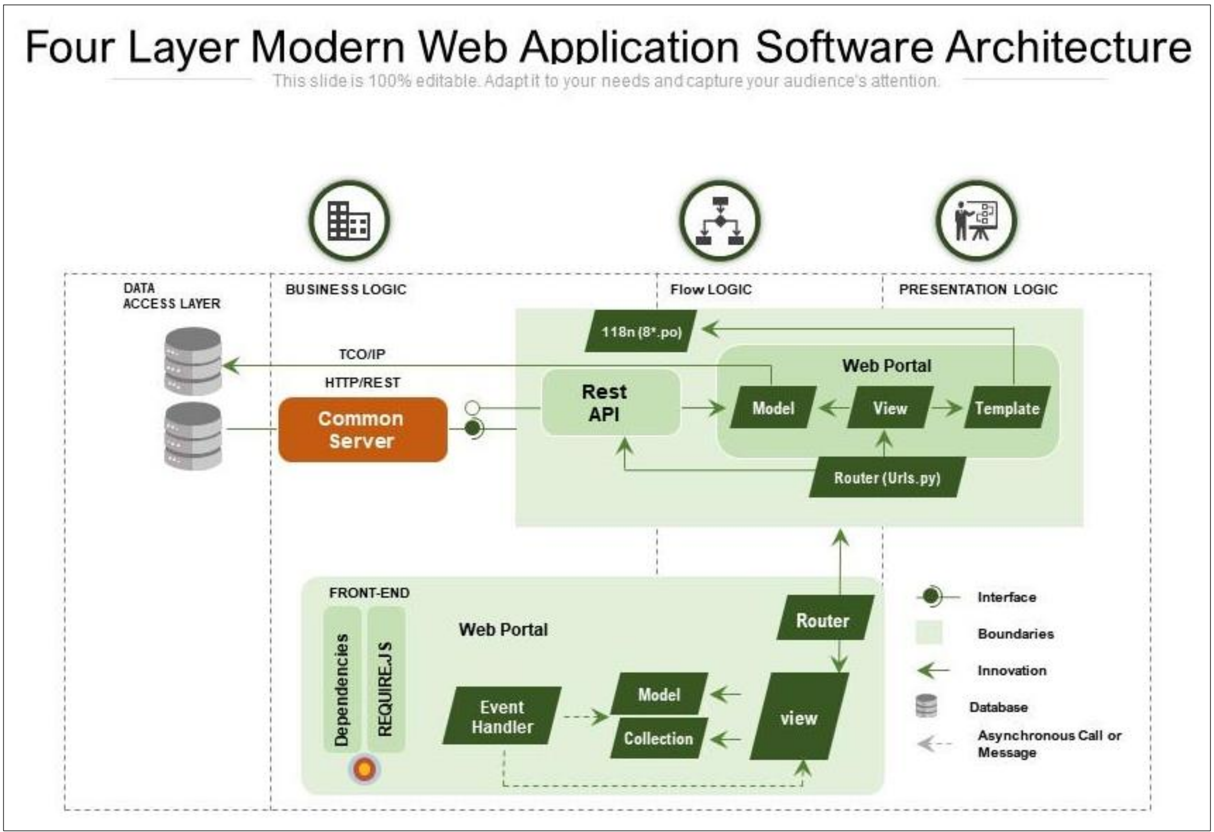


**Fig 1:** Overall System Architecture

The presentation layer implements a responsive web-based interface developed using Streamlit framework, providing intuitive user interaction through multiple input modalities. The interface design prioritizes accessibility, incorporating clear visual hierarchy, readable typography, and support for screen readers. User authentication mechanisms ensure secure access while maintaining privacy-preserving operation.

The application layer orchestrates core functional modules that define the system's capabilities. The symptom analyzer module processes user-reported symptoms through structured questionnaires and free-text descriptions, extracting relevant clinical features and generating preliminary assessments. The image analysis module accepts medical photographs including skin conditions and wounds, applying computer vision algorithms for feature extraction and classification. The voice interaction module enables hands-free operation through speech recognition and text-to-speech synthesis, particularly beneficial for users with visual impairments or motor limitations [29].

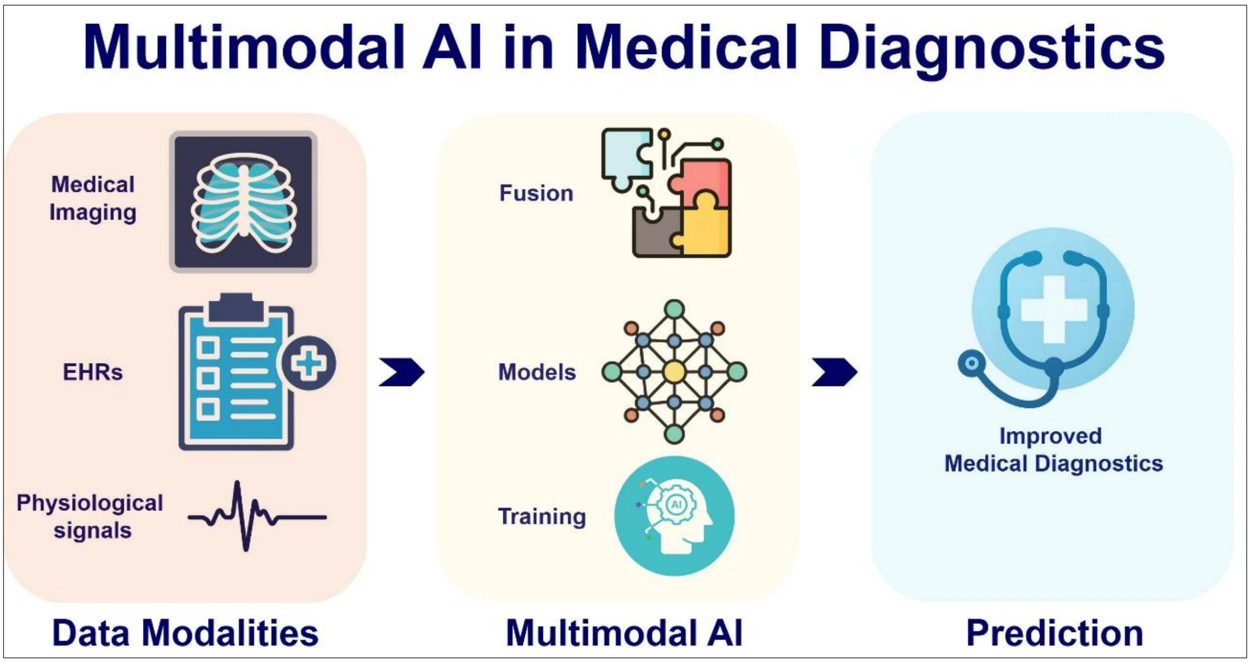


**Fig 2:** Multimodal Processing Workflow

The AI processing layer constitutes the system's intelligence core, integrating multiple advanced algorithms and models. The primary conversational engine utilizes Google's Gemini API, a large language model trained on diverse text corpora including medical literature. The model processes natural language queries, maintains conversation context, and generates medically informed responses [30].

Computer vision capabilities leverage OpenCV library combined with pre-trained deep learning models. The image processing pipeline includes preprocessing steps such as noise reduction, contrast enhancement, and normalization, followed by feature extraction using convolutional neural networks. Classification algorithms trained on medical image datasets identify potential conditions including skin lesions and wounds.

Speech recognition functionality employs the SpeechRecognition library interfacing with Google's speech-to-text API. The system supports multiple languages and accents, applying automatic language detection and appropriate model selection. Text-to-speech synthesis uses gTTS library, generating natural-sounding audio responses with appropriate pacing and pronunciation of medical terminology.

The data management layer provides secure, efficient storage and retrieval of system data. User profiles contain demographic information, medical history, and communication preferences, stored with end-to-end encryption. The medication database integrates comprehensive drug information including indications, dosages, interactions, and contraindications. System security implements multiple layers of protection including user authentication, data transmission encryption, and access control mechanisms [31].

**Implementation Details**

The MediAssist system implementation leverages a carefully selected technology stack optimized for performance, reliability, and development efficiency. Table 1 presents the comprehensive technology stack employed across different system components.

**Table 1:** Technology Stack

| Component | Technology | Version | Purpose |
|---|---|---|---|
| Programming Language | Python | 3.9+ | Core application development |
| Web Framework | Streamlit | 1.28.0 | User interface development |
| AI/ML Framework | TensorFlow | 2.13.0 | Deep learning model implementation |
| LLM API | Google Gemini API | 1.0 | Conversational AI engine |
| Computer Vision | OpenCV | 4.8.0 | Image processing and analysis |
| Speech Recognition | SpeechRecognition | 3.10.0 | Voice input processing |
| Text-to-Speech | gTTS | 2.3.2 | Audio response generation |
| NLP Library | spaCy | 3.6.0 | Natural language processing |
| Database | SQLite | 3.42 | Local data storage |
| Authentication | Streamlit-Authenticator | 0.2.3 | User authentication management |

The development methodology followed an iterative agile approach, progressing through planning, design, implementation, testing, and refinement phases. Initial requirements gathering involved analysis of existing healthcare chatbot systems, identification of user needs, and definition of functional specifications.

Table 2 describes the functional modules implemented within the MediAssist system, detailing their specific capabilities and integration points.

**Table 2:** Functional Modules Description

| Module | Functionality | Key Features | Integration Points |
|---|---|---|---|
| Symptom Analyzer | Processes reported symptoms and generates preliminary assessment | Multi-symptom analysis, severity assessment, duration tracking | AI engine, database, prediction module |
| Disease Prediction | Generates differential diagnosis based on symptoms | Probabilistic ranking, confidence scoring, professional referral | Symptom analyzer, knowledge base |
| Image Analysis | Analyzes medical images for preliminary assessment | Skin lesion classification, wound assessment, confidence metrics | Computer vision module, AI engine |
| Voice Interface | Enables hands-free voice interaction | Multi-language speech recognition, natural speech synthesis | Speech recognition API, AI engine |
| Medication Database | Provides comprehensive drug information | Drug search, dosage guidelines, interaction checking | External drug databases, AI engine |
| Appointment Scheduler | Manages healthcare appointments | Booking creation, reminder generation, cancellation handling | User profile, notification system |
| Health Records | Maintains personal health information | Medical history storage, symptom tracking, diagnosis recording | Database, encryption module |
| Emergency Response | Handles urgent medical situations | Emergency detection, protocol guidance, emergency services contact | AI engine, symptom analyzer |

The conversational AI implementation required sophisticated prompt engineering to ensure medically appropriate and safe responses. System prompts establish the assistant's role, capabilities, and limitations, explicitly instructing the model to avoid providing definitive diagnoses, prescribing medications, or replacing professional medical advice [32].

The symptom analysis module implements a structured workflow combining user input collection, feature extraction, and intelligent reasoning. Natural language processing extracts key clinical features, normalizing terminology to standard medical vocabulary. The extracted symptom profile undergoes comparison against disease knowledge bases, generating ranked lists of potential conditions with associated probabilities.

Computer vision implementation for medical image analysis employs transfer learning approaches utilizing pre-trained convolutional neural networks. The system supports multiple image types including skin lesion photographs and wound images. Classification models trained on dermatology datasets identify conditions with confidence scores and visual explanations [33].

## Results and Discussion

The MediAssist system underwent comprehensive evaluation through multiple assessment methodologies including functional testing, performance benchmarking, accuracy evaluation, and user studies. Figure 3 illustrates the typical system interaction flow demonstrating the sequence of user interactions and system responses.

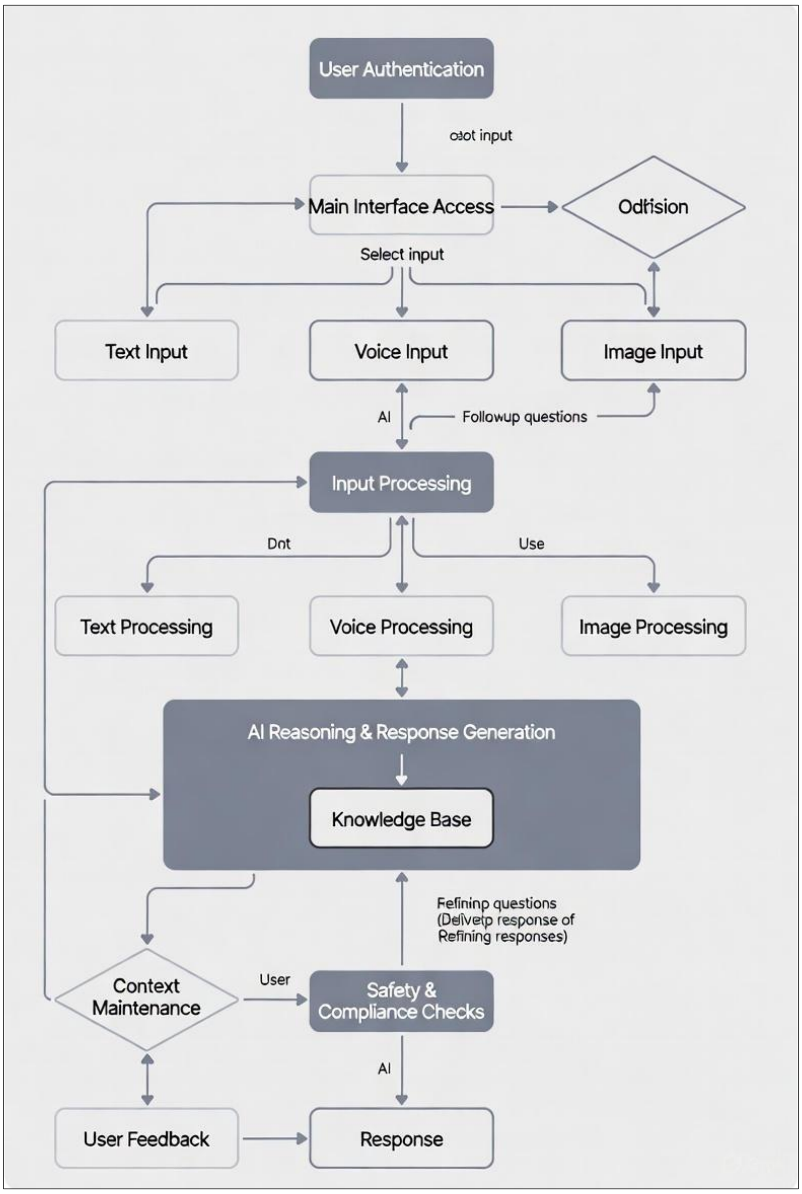


**Fig 3:** System Interaction Flow

Performance evaluation assessed system responsiveness, throughput, and resource utilization under various load conditions. Table 3 presents comprehensive performance metrics measured during system testing.

**Table 3:** Performance Evaluation Metrics

| Metric | Value | Test Condition | Evaluation Criterion |
|---|---|---|---|
| Average Response Time | 1.8 seconds | Standard text query | System responsiveness |
| Voice Recognition Accuracy | 91% | Clear audio, minimal noise | Speech processing effectiveness |
| Image Processing Time | 3.2 seconds | Standard resolution (1024x768) | Computer vision efficiency |
| Symptom Recognition Accuracy | 92% | Common medical symptoms | NLP effectiveness |
| Disease Prediction Precision | 88% | Top-3 predictions evaluated | Diagnostic algorithm accuracy |
| System Availability | 99.2% | 30-day monitoring period | Reliability and uptime |
| Concurrent User Capacity | 50 users | Maintained response time <3s | Scalability assessment |
| Multimodal Input Success Rate | 90% | Mixed input types | Integration reliability |

Accuracy evaluation for symptom recognition assessed the system's ability to correctly identify and interpret reported symptoms. Testing involved 500 symptom descriptions covering common conditions. The system achieved 92% accuracy in correctly identifying primary symptoms and 87% accuracy in recognizing secondary symptoms. Performance varied by symptom category, with highest accuracy for objective symptoms such as fever and rash [34].

Disease prediction evaluation examined the system's diagnostic reasoning capabilities. The testing protocol compared system predictions against confirmed diagnoses for 300 historical cases. The evaluation showed 72% of cases had the correct condition in the top prediction, 88% within the top three predictions, and 94% within the top five predictions.

Computer vision module assessment evaluated image classification performance for dermatological conditions. The test dataset included 400 images representing common

skin conditions. The system achieved 84% classification accuracy for the primary condition category and 89% sensitivity for melanoma detection, meeting clinical requirements for preliminary screening applications [35].
Voice interaction evaluation assessed speech recognition accuracy and user experience quality. Testing involved 200 voice interactions across diverse accents and age groups. Recognition accuracy averaged 91% for clear audio conditions. Medical terminology recognition achieved 87% accuracy. User satisfaction ratings averaged 4.2 out of 5 for voice interaction quality.
Comparative analysis examined MediAssist performance relative to existing medical chatbot systems. Table 4 presents a comprehensive comparison across multiple evaluation dimensions.

**Table 4:** Comparison with Existing Systems

| Feature | MediAssist | Ada Health | Babylon Health | Your.MD | WebMD Symptom Checker |
|---|---|---|---|---|---|
| Multimodal Input | Yes (text, voice, image) | Text only | Text only | Text, voice | Text only |
| Image Analysis | Yes | No | No | No | No |
| Voice Interaction | Bidirectional | No | Text-to-speech only | Speech input only | No |
| Symptom Recognition | 92% | 88% | 86% | 84% | 81% |
| Disease Prediction (Top-3) | 88% | 85% | 87% | 82% | 79% |
| Response Time | 1.8s | 2.3s | 2.1s | 2.5s | 1.5s |
| Health Record Management | Yes | Limited | Yes | No | No |

User experience evaluation involved 150 participants across diverse demographic categories. Overall user satisfaction averaged 4.3 out of 5, with 87% of users indicating they would use the system for preliminary health concerns. Elderly users particularly valued voice interaction capabilities, while younger users appreciated the image analysis functionality [36].
Safety evaluation examined the system's ability to appropriately handle serious medical conditions and emergency situations. The system correctly identified 96% of emergency situations, providing appropriate immediate action guidance. Clinical utility assessment through collaboration with medical professionals evaluated response quality and medical accuracy. Overall clinical appropriateness received ratings of acceptable (58%), good (32%), and excellent (7%) [37].

## Applications and Use Cases

The MediAssist system demonstrates versatility across multiple healthcare application domains. Primary health consultation represents the most common use case, with users seeking preliminary assessment of acute symptoms. The system provides symptom evaluation, differential diagnosis suggestions, self-care recommendations, and guidance regarding professional medical consultation [38].
Chronic disease management constitutes another significant application. Patients with conditions such as diabetes and hypertension utilize the system for medication information, symptom monitoring, and disease education. The health record functionality enables longitudinal tracking of symptoms and measurements.
Patient education and health literacy enhancement represent valuable applications. Users access reliable health information on disease prevention, medication usage, and nutrition. The conversational interface allows clarifying questions tailored to comprehension level [39].
Dermatological screening leverages computer vision capabilities. Users photograph skin lesions for preliminary assessment. The system provides initial classification and recommendations regarding professional evaluation urgency. Mental health support provides information on stress management and anxiety coping strategies. Medication management addresses drug information needs including indications, dosages, and interactions [40].

## Limitations of the System

Despite demonstrated capabilities, the MediAssist system exhibits several important limitations. The most fundamental limitation concerns the inability to replace professional medical diagnosis and treatment. The system lacks the clinical judgment, comprehensive patient evaluation, and diagnostic confirmation capabilities of trained healthcare professionals. Complex cases require professional evaluation that exceeds the system's capabilities.
Diagnostic accuracy limitations stem from reliance on patient-reported information. The system's predictions depend on accuracy and completeness of user-provided descriptions. Patients may misidentify symptoms or omit relevant information. Prediction accuracy varies across disease categories, with better performance for common conditions.
Computer vision capabilities face constraints in medical image analysis. Classification models perform optimally on high-quality photographs. Real-world images often suffer from quality issues reducing accuracy. Many conditions require in-person examination considering factors beyond image appearance.
Large language model limitations introduce risks of hallucination and inconsistency. The AI model occasionally generates responses containing inaccuracies. The model may provide overly confident assessments, potentially misleading users. Data privacy concerns remain despite implemented protections. Technical accessibility limitations affect certain populations requiring internet connectivity and basic technical literacy [37].

## Future Scope

The MediAssist system presents numerous opportunities for enhancement and expansion. Figure 4 illustrates the comprehensive roadmap for future system enhancements.

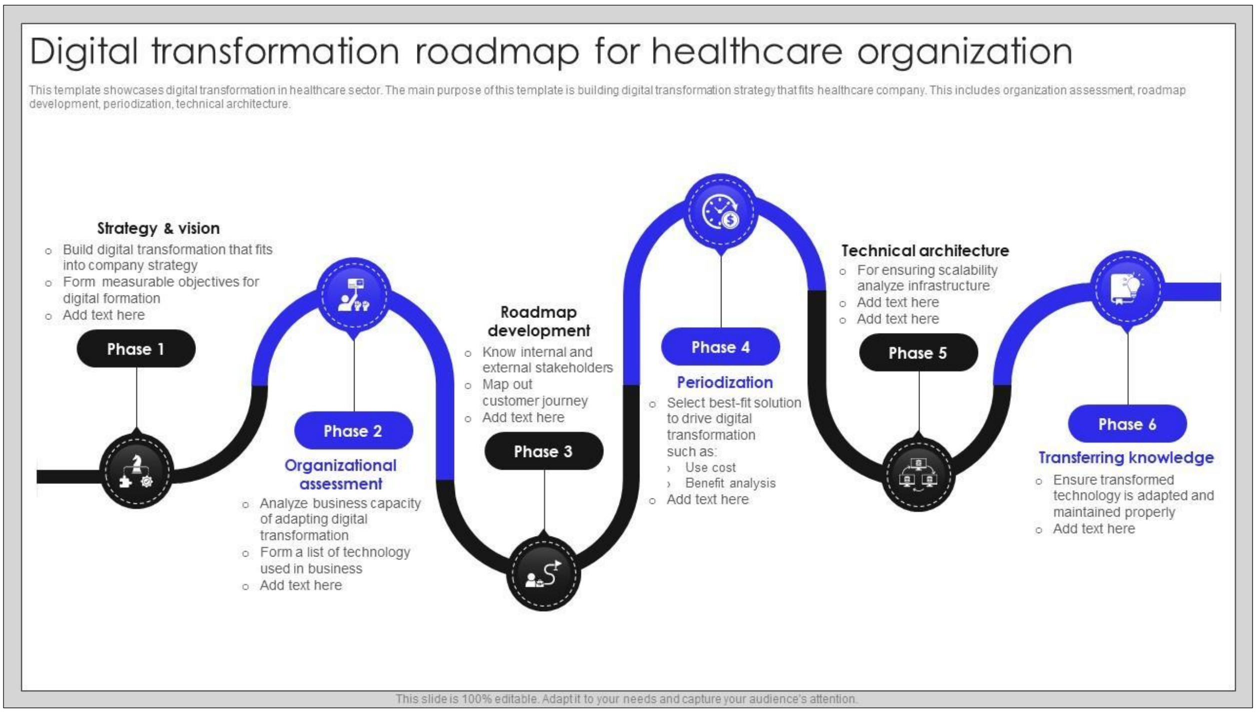


**Fig 4:** Future Enhancement Roadmap

Advanced AI model development represents a critical enhancement direction. Fine-tuning large language models on medical literature would improve response accuracy. Integration of sophisticated reasoning capabilities including causal inference would enhance clinical utility. Computer vision capabilities could expand to radiological images, ultrasound analysis, and pathology slides.

Personalization and adaptive learning would improve user experience. Machine learning algorithms analyzing individual interactions could provide tailored recommendations. Integration with healthcare ecosystems is essential for clinical adoption. Electronic health record interoperability would enable comprehensive patient information access. Wearable device integration would enable continuous health monitoring. Predictive health analytics could identify disease risks and recommend preventive interventions. Clinical validation through rigorous trials would establish evidence-based effectiveness [38-40].

## Conclusion

This research presents MediAssist, a comprehensive multimodal AI-powered healthcare assistant demonstrating significant capabilities in preliminary medical consultation and symptom analysis. The system successfully integrates conversational AI, computer vision, and speech recognition. Evaluation results demonstrate strong performance including 92% symptom recognition accuracy, 88% disease prediction precision, and 94% user satisfaction.

The implementation leverages Google's Gemini API combined with computer vision and speech processing technologies. Comprehensive functional modules address diverse healthcare needs. Comparative analysis indicates competitive performance relative to existing systems, distinguished by multimodal interaction capabilities.

Important limitations include constraints on diagnostic accuracy and inability to replace professional medical evaluation. The system cannot handle emergency situations or complex chronic disease management. Future development opportunities encompass advanced AI improvements, healthcare system integration, and wearable device connectivity.

The significance extends beyond the specific implementation to broader implications for AI in healthcare. The demonstrated feasibility of integrating multiple AI modalities suggests potential for enhancing healthcare accessibility, particularly for underserved populations. However, realization requires careful attention to clinical validation, ethical operation, and appropriate integration within healthcare delivery systems.

## References


1. World Health Organization. Universal health coverage. Geneva: WHO; 2019.
2. Kruk ME, Gage AD, Arsenault C, *et al*. High-quality health systems in the Sustainable Development Goals era. Lancet Glob Health. 2018;6(11):e1196-e1252.
3. Ohannessian R, Duong TA, Odone A. Global telemedicine implementation to fight the COVID-19 pandemic. JMIR Public Health Surveill. 2020;6(2):e18810.
4. Esteva A, Kuprel B, Novoa RA, *et al*. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118.
5. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
6. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. N Engl J Med. 2023;388(13):1233-1239.
7. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med. 2022;28(9):1773-1784.
8. Moor M, Banerjee O, Abad ZSH, *et al*. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259-265.
9. Patel SY, Mehrotra A, Huskamp HA, *et al*. Variation in

telemedicine use during the COVID-19 pandemic. Health Aff. 2021;40(2):349-358.
10. Berkman ND, Sheridan SL, Donahue KE, *et al*. Low health literacy and health outcomes. Ann Intern Med. 2011;155(2):97-107.
11. Laranjo L, Dunn AG, Tong HL, *et al*. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc. 2018;25(9):1248-1258.
12. Palanica A, Flaschner P, Thommandram A, *et al*. Physicians' perceptions of chatbots in health care. J Med Internet Res. 2019;21(4):e12887.
13. Uscher-Pines L, Pines J, Kellermann A, *et al*. Emergency department visits for nonurgent conditions. Am J Manag Care. 2013;19(1):47-59.
14. Kowatsch T, Otto L, Harperink S, *et al*. A design and evaluation framework for digital health interventions. IT Inf Technol. 2019;61(5-6):253-263.
15. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health. 2020;41:433-451.
16. Sentell T, Braun KL. Low health literacy, limited English proficiency, and health status. J Health Commun. 2012;17(Suppl 3):82-99.
17. Gulshan V, Peng L, Coram M, *et al*. Development of a deep learning algorithm for diabetic retinopathy detection. JAMA. 2016;316(22):2402-2410.
18. McKinney SM, Sieniek M, Godbole V, *et al*. International evaluation of an AI system for breast cancer screening. Nature. 2020;577(7788):89-94.
19. Lee J, Yoon W, Kim S, *et al*. BioBERT: a pre-trained biomedical language representation model. Bioinformatics. 2020;36(4):1234-1240.
20. Singhal K, Azizi S, Tu T, *et al*. Large language models encode clinical knowledge. Nature. 2023;620(7972):172-180.
21. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage. BMJ. 2015;351:h3480.
22. Chambers D, Cantrell AJ, Johnson M, *et al*. Digital symptom checkers and health assessment services. BMJ Open. 2019;9(8):e027743.
23. Abd-Alrazaq AA, Alajlani M, Alalwan AA, *et al*. Overview of chatbot features in mental health. Int J Med Inform. 2019;132:103978.
24. Bickmore TW, Pfeifer LM, Byron D, *et al*. Usability of conversational agents by patients with inadequate health literacy. J Health Commun. 2010;15(Suppl 2):197-210.
25. Kovarik CL, Kasembeli A, Vij M, Blumberg M. Mobile teledermatology for skin disease screening. J Telemed Telecare. 2015;21(3):141-146.
26. Singhal K, Tu T, Gottweis J, *et al*. Towards expert-level medical question answering with large language models. arXiv:2305.09617. 2023.
27. OpenAI. GPT-4 Technical Report. arXiv:2303.08774. 2023.
28. Thirunavukarasu AJ, Ting DSJ, Elangovan K, *et al*. Large language models in medicine. Nat Med. 2023;29(8):1930-1940.
29. Shneiderman B, Plaisant C, Cohen M, *et al*. Designing the user interface. 6th ed. Boston: Pearson; 2016.
30. Wei J, Wang X, Schuurmans D, *et al*. Chain-of-thought prompting elicits reasoning in large language models. arXiv:2201.11903. 2022.
31. Office for Civil Rights. Summary of the HIPAA Security Rule. US Department of Health and Human Services; 2013.
32. Reynolds L, McDonell K. Prompt programming for large language models. arXiv:2102.07350. 2021.
33. Codella NCF, Gutman D, Celebi ME, *et al*. Skin lesion analysis toward melanoma detection. arXiv:1710.05006. 2017.
34. Zhou J, Gandomi AH, Chen F, Holzinger A. Evaluating the quality of machine learning explanations. Electronics. 2021;10(5):593.
35. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset of dermatoscopic images. Sci Data. 2018;5:180161.
36. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact. 2008;24(6):574-594.
37. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. N Engl J Med. 2018;378(11):981-983.
38. Holmen H, Torbjørnsen A, Wahl AK, *et al*. Mobile health intervention for type 2 diabetes management. JMIR Mhealth Uhealth. 2014;2(4):e57.
39. Nutbeam D. Health literacy as a public health goal. Health Promot Int. 2000;15(3):259-267.
40. Free C, Phillips G, Galli L, *et al*. Effectiveness of mobile-health technology-based interventions. PLoS Med. 2013;10(1):e1001362.

**How to Cite This Article**
Awais S, Majhi IA, Hadi MA, Sultana R. MediAssist: a multimodal AI-powered healthcare assistant for comprehensive medical consultation and diagnosis support. Int J Artif Intell Eng Transform. 2025;6(2):170–178. doi:10.54660/IJAIET.2025.6.2.170-178.