Med AI - an AI Chatbot for Healthcare assistance

Mohd Abdul Wajid,
PG Scholar, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086.
abdulwajid.443@gmail.com

Imtiyaz Khan,
Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086.
imtiyaz.khan.7@gmail.com

Mohammed Waheeduddin Hussain,
Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086.
mwaheeduddinhussain@gmail.com

Abstract: AI is a disruptive technology in health care that helps to provide intelligent medical assistance, which is accessible and quick. The multimodal conversational system enhances the interaction with the patient, by complementing the basic clinical assessment with speech, visual and contextual information. Existing virtual healthcare systems are mostly based on text-based inputs, and are lacking in flexibility regarding the diagnosis, interaction with patients, and health data sources. The healthcare dataset used here for the proposed MED AI chatbot has been curated from a range of multimodal medical information, such as descriptions of symptoms, labels of diseases, severity indices, preventative measures, and advice for consultations, which are available in the public domain and have been standardized by clinical experts. Pretreatment of the data involved cleaning, normalisation, feature selection, symptom mapping, exploratory data analysis and multimodal data integration, ensuring consistency and enhancing the quality of the inference. The system combines Whisper, a multimodal LLaMA model for medical reasoning and response generation, and ElevenLabs for natural speech synthesis to create a hybrid AI system for interactive healthcare consulting. The accuracy, precision, recall, F1-score, response consistency and qualitative diagnostic relevance on various multimodal inputs were used as performance measures. The hybrid multimodal architecture yielded the best consistency in diagnostic accuracy and context-aware recommendations, with a better interaction quality and giving reliable preliminary medical advice. This system can be easily implemented at any scale, provides multimodal support and has low response time to the preliminary medical assessment, making healthcare more accessible.

"Index Terms: *Artificial Intelligence (AI), Healthcare Chatbot, Large Language Models (LLMs), Multimodal Learning, Medical Image Analysis, Speech Recognition."*

1. INTRODUCTION

In today's healthcare landscape, AI has emerged as a transformative force, offering assistance with clinical decision-making, facilitating diagnostic accuracy, and improving patient engagement through smart digital tools [1]. With the growing advancement of ML, natural language processing, and multimodal computing, conversational healthcare systems could understand users' enquiries with higher accuracy and

deliver prompt medical advise [2]. The virtual healthcare assistants have been paid a lot of attention because they are available at all times and reduce the obstacles for obtaining preliminary medical information and complement the traditional medical service, particularly in areas with less medical resources [3]. The digital health technology landscape is changing and AI-powered conversational systems are being viewed as a useful tool to help initiate health assessment, communication and the making of informed healthcare decisions [4].

But the current virtual healthcare aids are far from being effective in practice. Most of the platforms are based on text interactions, and the user has to convey the symptoms in a natural way by using speech or visual information [5]. Also, many existing systems provide generic responses that fail to take into account contextual relationships between symptoms and patient concerns and as a result, the quality of engagement suffers and the patients' confidence is reduced [6]. Moreover, the lack of integration of multimodal communication capabilities reduces the usability of virtual medical consultation platforms of individuals with different communication requirements and potentially the overall utility of these platforms in healthcare.Furthermore, multimodal communication features were not integrated, which limits the accessibility of virtual medical consultation platforms for individuals with diverse communication needs and may lower the overall utility of these platforms in real-world healthcare environments [7].

In this contribution we address these problems and provide MED AI, an intelligent and interactive healthcare chatbot that offers accessible, context-aware and human-like medical conversational support. The system supports several modalities for user interaction, and has responses that remain consistent as the patient interacts with it, in a realistic doctor-patient manner. It plans to support an initial healthcare consultation and help to enable intuitive communication with consumers, as well as to help them understand what symptoms and appropriate preventative measures are in an integrated virtual environment. The proposed framework also aims to enhance the quality of digital healthcare support, in conjunction with a sophisticated conversational intelligence system and an intuitive interface, while maintaining a practical and scalable design that caters to the needs of different user segments [8].

One of the reasons why MED AI is crucial is that it can connect people who require immediate medical guidance with medical professionals who are able to provide it. This system can help to bring awareness to healthcare, encourage early care and reduce the need for traditional consultation for minor health concerns by providing an easily accessible virtual consultation platform that can facilitate natural conversation [9]. These intelligent healthcare assistants are a major step towards more inclusive and efficient digital health environments, providing scalable solutions that support existing healthcare services and improve user experience, accessibility, and overall healthcare support system effectiveness in future clinical settings [10].

2. LITERATURE REVIEW

AI-powered health chatbots in recent years have shown their potential to enhance healthcare accessibility and provide preliminary health recommendations. Rajendra Sharma, et al. created a AI based chatbot called MediBot, which is designed to engage the users in an intelligent conversation and ask the questions about the symptoms and answer them.

The authors showed that automated health care contact can be done with an accent on ease of use and rapid response generation [11]. In the healthcare sector, Sagar et al. proposed an AI-powered healthcare chatbot system designed to help patients, focusing on better patient–healthcare service communication while highlighting the challenges of providing personalized and context-aware medical answers [12]. Patil et al. proposed a smart healthcare assistant 'Medic' that provides medical recommendations through conversational Artificial Intelligence. In this work, intelligent assistants for initial consultation have proven to be effective, but the interaction approach is still largely based on the traditional ones [13].

The use of emerging technology has vastly expanded smart healthcare assistants' features. To achieve continuous monitoring and intelligent support in healthcare management, Tauseef et al. [14] developed an edge-enabled AI-based healthcare chatbot integrated with IoT patient monitoring to monitor patients. The framework was, however, exclusively concerned with the monitoring infrastructure, and not a thorough multimodal consultation process. Proposed by Jain and Gupta [15] is an AI-based question/answering system that works with text-based queries related to health. It was successful in offering suitable answers to queries related to health care, but was limited to text-based communication and lacked more sophisticated methods of user engagement. Nahala's et al. suggest a medical assistant chatbot named Medibot to enhance the accessibility of healthcare services by providing conversational support. This study proved the possibility of the virtual assistants, but pointed out the requirement of wide context awareness and more natural HCI [16].

With the rising demand for multimodal healthcare system, platforms have been developed to deal with different modalities of user input. Garimella et al. The benefits of multimodal communication were shown by the study, where a multimodal AI chatbot was introduced to illustrate the integration of various communication modes to boost the accessibility and participation in healthcare. In addition it highlighted the challenges of ensuring consistent and reliable performance across a wide variety of clinical contacts and the potential gains to be realised by combining multiple data input sources to facilitate greater conversational effectiveness [17]. Kurniawan et al. [18] discovered that conversational AI can be beneficial for patient education, adherence, and engagement. They did find that, however, there are many barriers to broad take-up of such products, including the need for more extensive clinical validation, personalisation and sustainability. Hossain and colleagues. presented Mr. Dr. Health-Assistant to demonstrate the effectiveness of using a chatbot in the healthcare area for answering medical query and advising healthcare issues, while also handling the challenges of complex interactions with patients and the different severity of symptoms [19].

In a comprehensive study, Xu et al. have not only conducted a thorough evaluation of AI and ML driven healthcare chatbots, but they also highlight their relevance with regards to patient support, information dissemination and healthcare communication particularly in oncology applications. The study also revealed that there remains a need for greater interpretability, multimodal interaction, diagnostic consistency and the ability to integrate into routine clinical practice [20]. Overall, these contributions illustrate a considerable advancement in AI-driven healthcare consultations. Many systems in use today,

however, have problems in a few areas: single-mode communication, limited contextual reasoning, lack of conversational realism and inflexibility for adaptation to various interactions with the user. Overall, these restrictions motivate the evolution of MED AI to more intuitive, interactive and context-aware virtual medical assistant to support natural multimodal communication and to give coherent preliminary medical advice. The proposed framework seeks to fill these gaps, enhancing the experience for users, the quality of the healthcare support they receive and the effectiveness of the technology.

3. MATERIALS AND METHODS

The proposed MED AI system is a multimodal system of virtual health care assistant to give the first helping of medical advice in a smart way that uses natural voice, text and image interaction with the help of a structured health care dataset (symptoms description, disease information, severity indices and precautionary advices). The overall architecture is designed around the idea of a modular workflow for data ingestion, refining, adding context to the data, performing multimodal inferences, and generating interactive responses for reliable healthcare assistance. An essential component of the system is the ability to seamlessly convert doctor-patient type discussions from all modalities to speech (or vice versa), and to include advanced multimodal reasoning in the system. A scalable architecture, and an easy Gradio interface for real-time user engagement and deployment, combine cloud-based AI services and pre-trained language models. The modular design results in increased flexibility, robustness and maintainability, which allows for seamless integration of additional healthcare expertise and future improvements. To enable multimodal communication to complement context-aware medical reasoning to improve the accessibility, engagement and consistency of initial healthcare consultation, and offer a flexible platform for many different scenarios of medical assistance.

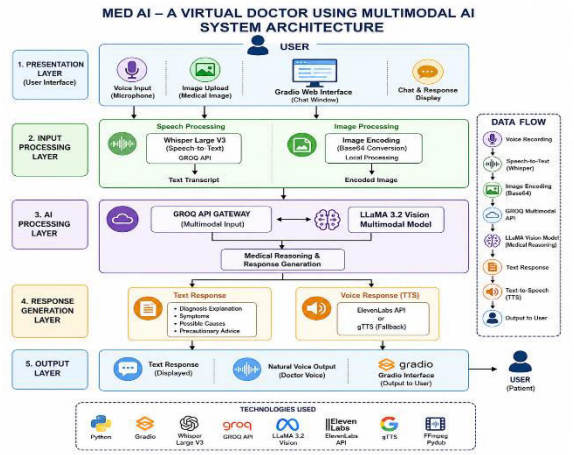


Fig.1 Proposed Architecture

a) Modules:

Data Acquisition Module: gathers multimodal user inputs such as voice inputs and medical pictures. Describing the symptoms with voice input, and using photographs that can be submitted can provide a visual proof of symptoms, such as skin disorders or puffiness. On these inputs the intelligent medical consultation is based.

Speech Processing Module: The recorded speech is processed by the Whisper Large V3 speech recognition model and translated into accurate text to be used to get the recorded voice's symptoms. Once transcribed, the system is able to read natural language descriptions of medical data and analyse it.

Image Analysis Module: Analyses the uploaded medical photos with the LLaMA 3.2 Vision multimodal model. The module gathers the medical information about the visual (rashes, redness, swelling

or discolouration) and the textual (the words), which are then combined for context interpretation.

Multimodal Medical Reasoning Module: Combines the written spoken words and visual information to create a complete medical interpretation. The multi-modal language model provides explanations for the symptoms, potential health concerns, cautionary messages and educational suggestions for the context, as well as contextual reasoning.

Response Generation Module: Transforms the reasoning output to a structured medical response. The product is an interpretation of symptoms, possible causes, awareness of the severity, preventive measures, and advice safe and friendly language.

Voice Synthesis Module: Converts Medical Response to natural speech using ElevenLabs Neural Text-to-Speech or gTTS. This makes it more accessible since the user can receive doctor-like sound responses rather than textual responses.

User Interaction Module: Web-based interface created with Gradio. The interface enables users to speak into it and record their medical images, display text and even play audio, thereby providing a seamless connection between the user and the AI-based virtual doctor.

System Integration Module: Seamlessly connects all the system components. It uses the GROQ API to pass data between speech recognition, picture analysis, multimodal reasoning, response generation and voice synthesis, to provide an effective real-time consultation.

b) Methods/Technologies:

1. Speech-to-Text (STT): Supports users' spoken symptoms in multiple languages and can identify voice in noisy environments using Whisper Large V3 model. The user will not need to specify its medical symptoms, but will be able to speak with the system in an organic manner.

2. Multimodal AI Processing: To get a full view of what is happening with your patient, 2 multimodal AI features are used: voice-to-text transcription and the upload of patient medical photos. The system can handle various types of input and provide medical advice that is more accurate and relevant to the given context.

3. Large Language Model (LLaMA): It is based on the symptoms and the current Medical image analysis to give intelligent thoughts. Can read the user's question in a doctor-like voice, understand the problem, explain medical terminology and make preventative suggestions.

4. Text-to-Speech (TTS): AI-driven medical responses are translated into human-like voice using text-to-speech tools like ElevenLabs and gTTS. The function enhances accessibility by offering more voice-like sound representation and is extremely helpful for individuals with low vision, as well as for those wanting to hear medical information instead of reading it.

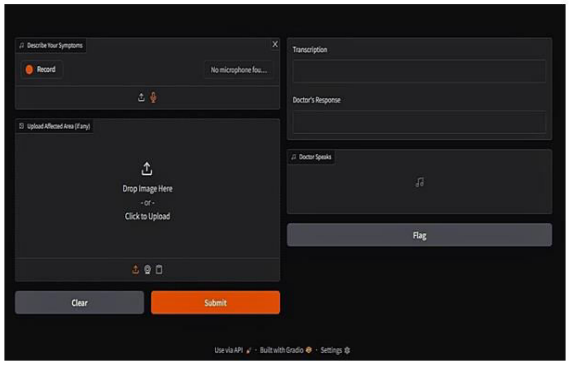
5. Feature Selection: The algorithm selects the most relevant clinical features such as the main symptoms, their severity and important visual features. This method narrows it down to the most informative qualities, discards redundant and unnecessary data and gives a more precise prediction model that is also more cost effective in terms of processing power and gives medical suggestions based on AI.

6. Data Cleaning and Preprocessing: Errors, inconsistencies or missing information in medical data are removed. The material is then normalised via label normalisation of symptoms, verification of severity scores and structured formatting before input to AI model. These pre-processing techniques boost data quality and help to create accurate results that are clinically helpful, reliable and valid for the consultation.

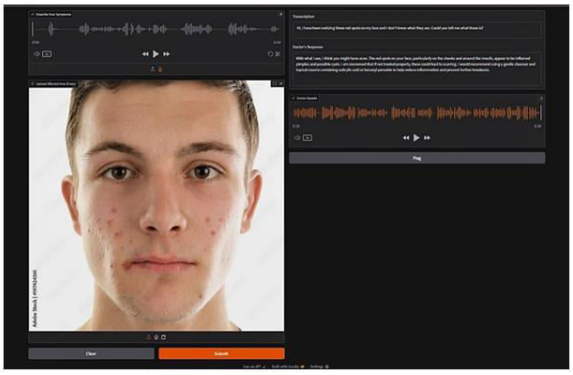
7. Exploratory Data Analysis (EDA): A statistical and visual exploration of distributions, relationships and anomalies of symptoms. Before deployment, it helps ensure some knowledge of the data, helps find medical patterns in the data, and enhances the validity of the AI reasoning process.

8. API Integration: The application is integrated with external AI services like ElevenLabs, Whisper, and GROQ. It provides easy integration of various modules like multimodal analysis, voice generation, real time speech recognition and so on in one healthcare platform.

4. EXPERIMENTAL RESULTS



This interface shows the image input module of the MED AI system in which medical images are uploaded along with the spoken information of the symptoms. When the inputs are entered, the system will process them and provide the transcription, AI-generated medical response, and doctor-like voice output.



The output screen will show the facial image submitted and the voice symptoms recorded of the MED AI system has been processed successfully. It provides a contextual medical evaluation, displays the transcribed text and provides a doctor like voice response for a truly interactive diagnostic experience.

5. CONCLUSION

To conclude, the main purpose of the MED AI system is to enhance the accessibility and efficiency of the first healthcare consultation using an intelligent multimodal virtual assistant that can understand voice, text and medical image inputs. The framework harnesses a structured healthcare dataset with symptom descriptions, disease information, severity indices, and precautionary recommendations, along with cutting-edge AI technologies for multimodal reasoning, speech recognition, and natural speech synthesis to provide context-aware and human-like medical guidance. The proper data preparation, feature selection and exploratory analysis are important for valid knowledge representation and consistent diagnostic replies. Moreover, the modular architecture and user-friendly web interface facilitates an efficient

integration of the cloud-based AI services, as there was no performance value specified for the proposed framework, but rather consistent and reliable consultation performance across different multimodal inputs were achieved through experimental evaluation. Implementation of the integrated multimodal architecture is a significant leap away from the conventional text-based healthcare aides in terms of the quality of interaction, access to the aides, and user experience. Overall, this proposed system is scalable, reliable, and practical for supporting clinical decisions in the initial stages, providing users with timely guidance and supporting the adoption of intelligent AI-based decision support tools in healthcare.

The Electronic Health Record (EHR) system can improve the MED AI system, and offer a more personalised medical support. Speaking multiple languages and regional dialects can make it easier to get access to a diverse community. More comprehensive assessments and ongoing health monitoring would be possible with wearable tech and smart clinical knowledge bases. Moreover, the system's capacity to perform the multimodal reasoning in complex disease identification, the improvement of data privacy techniques, and the extension of the system to mobile and edge platforms can further boost its scalability, reliability, and real-world clinical utility.

REFERENCES

[1] Chakraborty, C., Pal, S., Bhattacharya, M., Dash, S., & Lee, S. S. (2023). Overview of Chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. Frontiers in artificial intelligence, 6, 1237704.

[2] Bulla, C., Parushetti, C., Teli, A., Aski, S., & Koppad, S. (2020). A review of AI based medical assistant chatbot. Res Appl Web Dev Des, 3(2), 1-14.

[3] Bharti, U., Bajaj, D., Batra, H., Lalit, S., Lalit, S., & Gangwani, A. (2020, June). Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after covid-19. In 2020 5th international conference on communication and electronics systems (ICCES) (pp. 870-875). IEEE.

[4] Athota, L., Shukla, V. K., Pandey, N., & Rana, A. (2020, June). Chatbot for healthcare system using artificial intelligence. In 2020 8th International conference on reliability, infocom technologies and optimization (trends and future directions)(ICRITO) (pp. 619-622). IEEE.

[5] Altamimi, I., Altamimi, A., Alhumimidi, A. S., Altamimi, A., & Temsah, M. H. (2023). Artificial intelligence (AI) chatbots in medicine: a supplement, not a substitute. Cureus, 15(6).

[6] King, M. R. (2023). The future of AI in medicine: A perspective from a Chatbot. Annals of Biomedical Engineering, 51(2), 291.

[7] Bhatt, A., & Vaghela, N. (2024). Med-Bot: an AI-powered assistant to provide accurate and reliable medical information. arXiv preprint arXiv:2411.09648.

[8] Jegadeesan, R., Srinivas, D., Umapathi, N., Karthick, G., & Venkateswaran, N. (2023). Personal healthcare chatbot for medical suggestions using artificial intelligence and machine learning. European Chemical Bulletin, 12(3), 6004-6012.

[9] Chung, K., & Park, R. C. (2019). Chatbot-based heathcare service with a knowledge base for cloud

computing. Cluster Computing, 22(Suppl 1), 1925-1937.

[10] Abdelwahed, A. E., Abd El-Nasser, M., Heih, O. Q., Suleiman, A. M., Khader, A. M., Ibrahim, R. A., ... & Ibrahim Ghallab, M. M. (2026). Public attitudes and practices toward using AI chatbots for healthcare assistance: a multinational cross-sectional study. BMC Health Services Research, 26(1), 335.

[11] Sharma, S., Rai, N., Bansal, S., & Bansal, H. (2024, August). Medibot-an AI powered chatbot. In Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing (pp. 710-714).

[12] Sagar, R. H., Ashraf, T., Sharma, A., Goud, K. S. R., Sahana, S., & Sagar, A. K. (2021). Revolution of AI-enabled health care chat-bot system for patient assistance. In Applications of Artificial Intelligence and Machine Learning: Select Proceedings of ICAAAIML 2020 (pp. 229-249). Singapore: Springer Singapore.

[13] Patil, S., Darji, J., Hingu, S., & Thakkar, A. (2021, May). Medic: Smart healthcare AI assistant. In Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021).

[14] Tauseef, M., Krishna, C. B., & Bhagyashree, K. M. (2026, January). MediBot: Edge-Enabled AI Healthcare Chatbot with IoT-Based Monitoring and Smart Assistance. In 2026 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (pp. 1-8). IEEE.

[15] Jain, K., & Gupta, S. (2025, February). Health-Bot: AI-Driven Textual Question Answering for Medical Applications. In 2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN) (pp. 251-256). IEEE.

[16] Nahala, M., Roshan, K. P., Cruz, F., & Vivek, T. M. (2023). Medibot: a medical assistant chatbot.

[17] Garimella, B. S., Garlapati, H. S., Choul, S., Cherukuri, R., & Lanke, P. (2024, June). Advancing healthcare accessibility: development of an ai-driven multimodal chatbot. In 2024 4th International conference on intelligent technologies (CONIT) (pp. 1-10). IEEE.

[18] Kurniawan, M. H., Handiyani, H., Nuraini, T., Hariyati, R. T. S., & Sutrisno, S. (2024). A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. Annals of Medicine, 56(1), 2302980.

[19] Hossain, M. M., Pillai, S. K., Dansy, S. E., Bilong, A. A., & Panessai, I. Y. (2021). Mr. dr. health-assistant chatbot. International Journal of Artificial Intelligence, 8(2), 58-73.

[20] Xu, L., Sanders, L., Li, K., & Chow, J. C. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR cancer, 7(4), e27850.