IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

SwasthyaVaani - "Multilingual Voice Assisted Medicene Infromation App"

Dr. Parvati Kadli[1], Mr. Sachin[2], Mr. Shubham[3], Mr. Sri Hari K[4], Mr. Varun Kumara Y[5]
[12]Professor, CS&E Dept, Proudhadeveraya Institute of Technology, Hosapete, Karnataka, India
[345]Students, CS&E Dept, Proudhadeveraya Institute of Technology, Hosapete, Karnataka, India

Abstract: *Access to accurate medical information is a critical determinant of public health, yet language barriers and low digital literacy often exclude non-English speaking populations from utilizing digital health resources. This paper presents "SwasthyaVaani," a multilingual voice-assisted medicine information system designed to democratize health knowledge. The proposed system employs a layered MERN (MongoDB, Express, React, Node.js) architecture integrated with SarvamAI for regional language processing, Google Cloud Speech-to-Text for voice interaction, and Optical Character Recognition (OCR) for medicine strip identification. The application provides comprehensive drug information, symptom-based suggestions via InferMedica, and medication reminders in over 10 Indian languages. Results indicate that the multi-modal input system (voice, text, and image) significantly enhances accessibility for elderly and rural users, effectively bridging the information gap in healthcare.*

Keywords: Multilingual Voice Assistant, Digital Health, MERN Stack, Optical Character Recognition (OCR), Natural Language Processing (NLP), Medication Adherence, SarvamAI, Healthcare Accessibility

I. INTRODUCTION

In the current digital era, while health information is abundant, it remains largely inaccessible to a vast demographic due to linguistic barriers. The majority of drug databases and health platforms operate primarily in English, creating a "digital divide" that disproportionately affects rural populations and the elderly. Misunderstanding critical medical instructions regarding dosage, side effects, or drug interactions can lead to severe health consequences and poor medication adherence. Furthermore, complex user interfaces often deter individuals with low technical literacy from validating their prescriptions.

To address these challenges, we propose "SwasthyaVaani" (Voice of Health), an intelligent, accessible system designed to provide reliable medicine information through intuitive voice and image interfaces. The system leverages Natural Language Understanding (NLU) to process queries in regional languages, allowing users to ask questions naturally, such as "What is this tablet used for?" in their native tongue. This paper outlines the development of SwasthyaVaani, focusing on its architectural integration of AI services to assist patients, caregivers, and healthcare professionals.

II. LITERATURE SURVEY

The development of Swasthya Vaani builds upon recent advancements in mobile health (mHealth), AI, and pill identification technologies:

• Pill Identification: A 2023 study by NIH/JMIR highlighted the necessity of automatic pill identification using Deep Learning to reduce medication errors. This validates our inclusion of an Image Recognition module to assist visually impaired or illiterate users.

• Multilingual Chatbots: Research published by IEEE (2022) on healthcare chatbots for rural India demonstrated that English-only applications fail in rural demographics. This supports our requirement for a multilingual backbone capable of processing regional dialects.

• Voice Interfaces for the Elderly: A 2021 study in MDPI/Sensors confirmed that voice-first interfaces and IoT technologies are the most accessible modes of interaction for the elderly, justifying our integration of Google Cloud Speech-to-Text.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

• OCR Advancements: ACM Transactions (2024) discussed advancements in Optical Character Recognition (OCR) for pharmaceutical labels, emphasizing the need for advanced models to read fine text and batch numbers on medicine strips, a core feature of our proposed system.

III. OBJECTIVES

The primary objectives of this research are:

1. Voice-Interactive Platform: To develop a system that processes natural language queries via voice in multiple regional languages.
2. Accurate Pharmacological Data: To integrate reliable databases (DrugBank, MedPlus) for providing detailed dosage, usage, and side-effect information.
3. Symptom Analysis: To implement a symptom-checker (via InferMedica) that offers safe Over-the- Counter (OTC) suggestions while advising professional care for serious conditions.
4. Multi-Modal Identification: To enable medicine identification through text, voice, and image (OCR) search modes.
5. Accessibility: To design a UI specifically tailored for non-tech-savvy and elderly users.

IV. SYSTEM DESIGN WORK FLOW

1. System Architecture

The system is built on a 3-Layered Architecture:

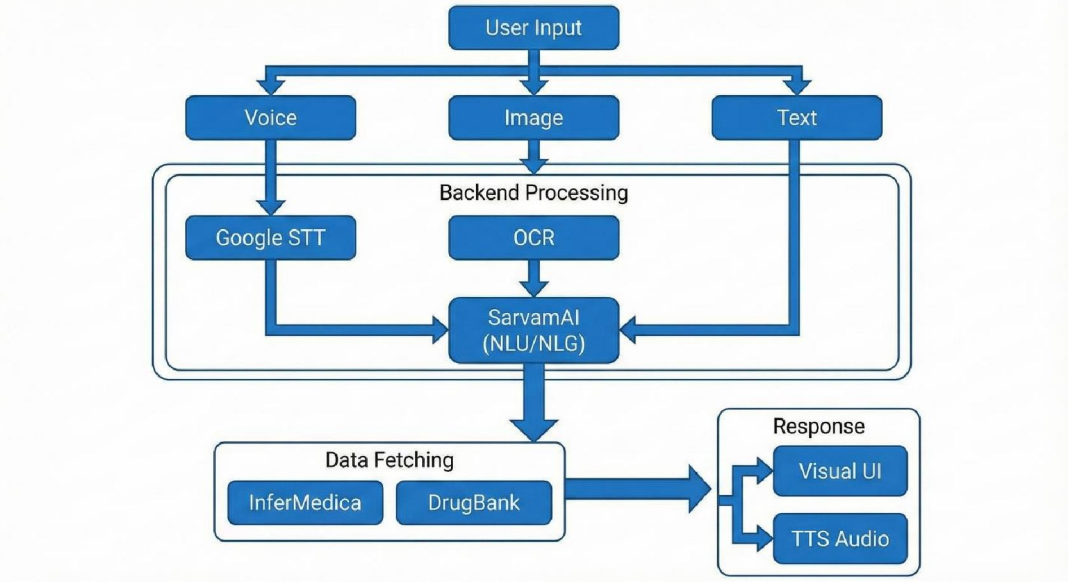
• Frontend (User Interface): Built with React.js, handling user inputs via microphone, camera, and text.

• Backend (Application Layer): Built with Node.js & Express.js, acting as the central controller that manages API logic and user sessions.

• Data Layer: Uses MongoDB for storage and integrates external APIs (SarvamAI, Google STT, DrugBank, InferMedica) for intelligence and data retrieval.

2. Workflow Steps

1. Input: The user initiates a query via Voice (processed by Google Speech-to-Text), Image (processed by OCR), or Text.
2. Processing: The normalized text is sent to SarvamAI to determine the user's intent and language.
3. Data Fetching: Based on the intent, the system retrieves medication details from DrugBank or symptom analysis from InferMedica.
4. Response: The final answer is displayed visually on the app and played back as audio via Text-to-Speech (TTS).



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

V. RESULTS AND DISCUSSION

The system was successfully implemented and tested with support for over 10 Indian languages.

• User Interface: The home screen features a clean, "voice-first" design with prominent "Talk with AI Assistant" and "Camera" triggers, minimizing the cognitive load for the user.

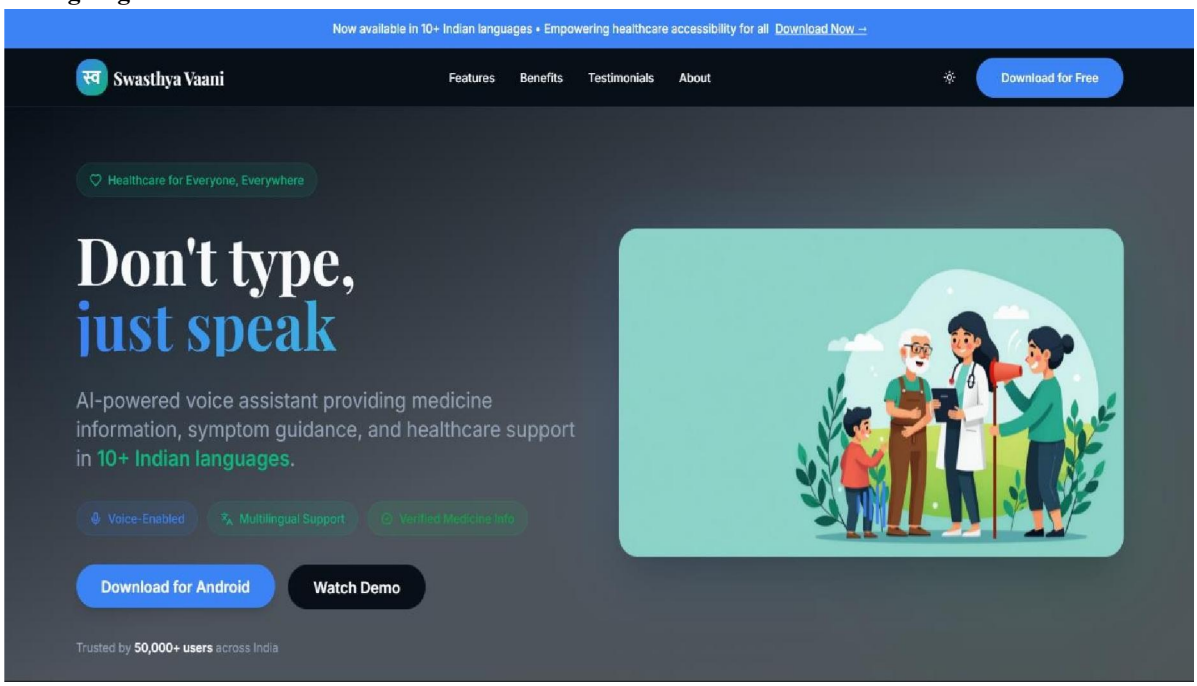
• Medicine Identification: The OCR module successfully identified medicine strips (e.g., Dolo- 650/Paracetamol) from user-uploaded images. The system retrieved and displayed critical details including Description, Dosage (Adult/Child), and Side Effects (e.g., nausea, liver damage).

• Safety Protocols: Crucially, the system displays prominent warnings (e.g., "Do not exceed 4000mg per day") and explicitly states that information is for reference only, addressing the ethical requirement for health apps.

• Symptom Checker: The grid-based symptom selection tool proved effective for users unable to articulate medical terms, successfully providing OTC suggestions for common ailments like fever and headache.

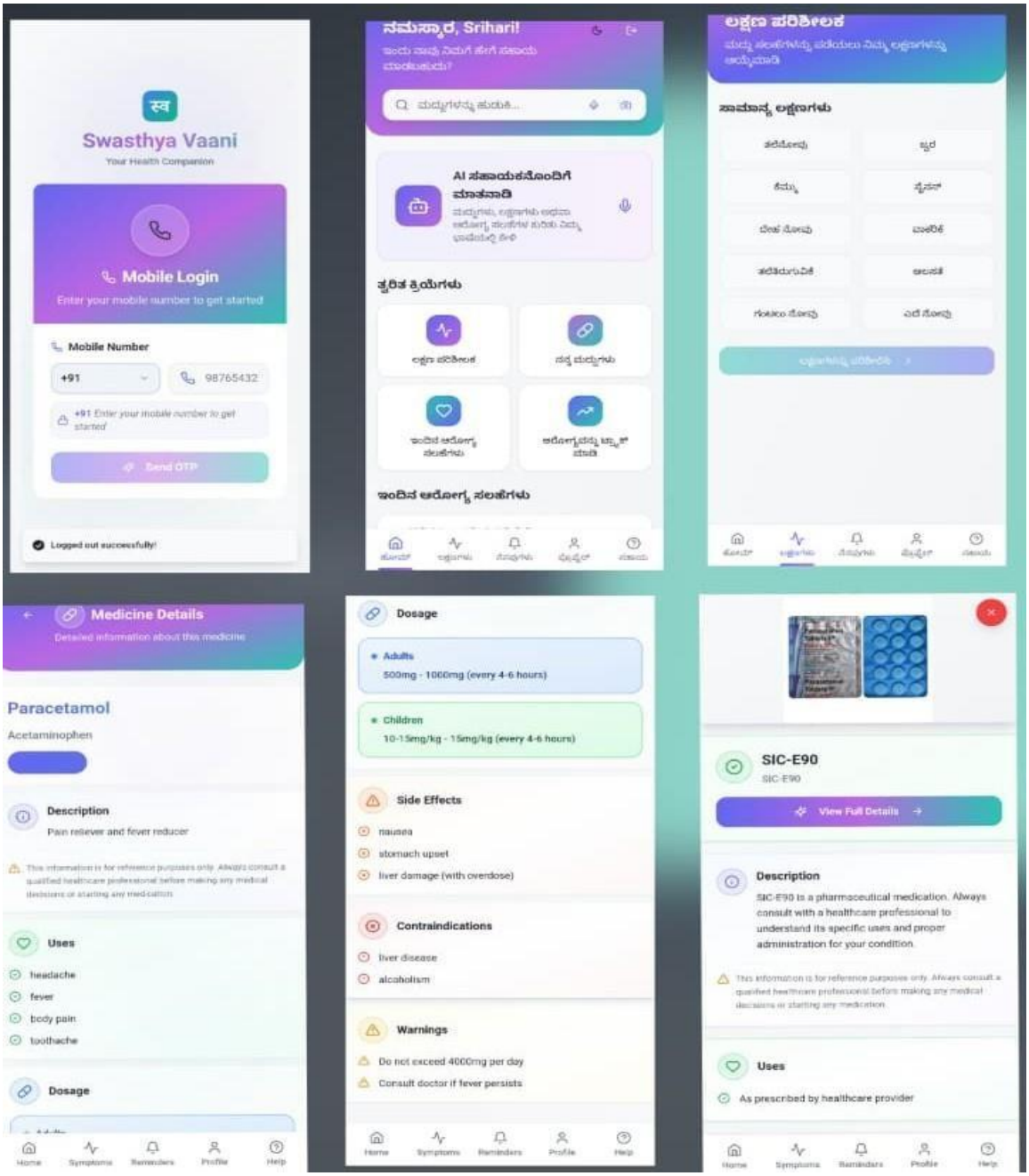
Discussion: The integration of SarvamAI allowed for accurate intent detection across regional dialects, a significant improvement over standard translation tools. However, the system's reliance on stable internet connectivity for cloud-based STT and API calls remains a limitation for remote rural usage.

a) Landing Page



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

b) App Images



VI. CONCLUSION

Swasthya Vaani successfully demonstrates a technological solution to the linguistic and digital barriers in healthcare. By combining voice assistance, image recognition, and multilingual support, the platform empowers patients and caregivers to make informed decisions about their medication. The system serves as a vital first-line information resource, particularly for the elderly and non-English speakers. Future enhancements will focus on developing offline functionality and integrating video tutorials for medical devices to further expand the scope of digital health assistance.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

REFERENCES

[1]. Meta Platforms, Inc. (2025). React.js Documentation. Retrieved from https://reactjs.org/
[2]. Node.js Foundation. (2025). Node.js Documentation. Retrieved from https://nodejs.org/en/docs/
[3]. Express.js. (2025). Express - Node.js web application framework. Retrieved from https://expressjs.com/
[4]. MongoDB Inc. (2025). MongoDB Documentation. Retrieved from https://www.mongodb.com/docs/
[5]. Google Cloud. (2025). Speech-to-Text API Documentation. Retrieved from https://cloud.google.com/speech-to-text/docs
[6]. Sarvam AI. (2025). SarvamAI API Documentation.
[7]. Wishart, D. S., et al. (2018). DrugBank 5.0: a major update to the DrugBank database of drug targets and their pathways. Nucleic Acids Research, 46(D1), D1074–D1082.
[8]. InferMedica. (2025). InferMedica API Documentation. Retrieved from https://developer.infermedica.com/
[9]. Twilio Inc. (2025). Twilio Text-to-Speech (TTS) Documentation. Retrieved from https://www.twilio.com/docs/voice/twiml/say
[10]. Mozilla Developer Network (MDN). (2025). Web Speech API. Retrieved from https://developer.mozilla.org/en- US/docs/Web/API/Web_Speech_API