IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Personal AI Companion

[1]Bhaskar K [2]G Vinod kumar [3] Hanamant Ajatarao [4]Harsha N Bhushan .Mrs. [5]Pragati P

[1,2,3,4]Student, ECE, Department, K.S .Institute Of Technology, Karnataka, India

[5]Associate Professor ,Visveswaraya Technological University, India

ABSTRACT

The Personal AI Companion is a portable, voice-based conversational system designed to offer users an interactive and intelligent assistance experience. Built using an ESP32 microcontroller, MEMS microphone, speaker, amplifier, and OpenAI-based APIs, the device can capture speech, convert it into text, generate context-aware responses, and deliver audio replies in real time. This project aims to create a low-cost, compact, and user-friendly AI assistant capable of long-term memory integration, ethical data handling, and personalized interactions without the need for continuous smartphone use. The implemented prototype achieves reliable two-way communication, proximity-based activation, and efficient power management through embedded hardware modules. The system successfully demonstrates real-time conversational capability, offering practical applications in personal guidance, education, accessibility support, and social companionship. The project highlights the potential for future enhancements such as multilingual support, IoT-based smart home control, offline optimization, and advanced emotional or contextual awareness to further improve user experience and system autonomy.

Keywords:

Personal AI companion, ESP32 microcontroller, voice-based assistant, MEMS microphone, OpenAI API, speech recognition, natural language processing, text-to-speech, embedded systems, real-time communication, low-power device, ethical AI, contextual awareness, personalized interaction, Internet of Things (IoT), accessibility support.

INTRODUCTION

A Personal AI Companion is an intelligent conversational system designed to interact with humans in natural language and provide educational support, communication assistance, productive reminders, general information and personalized responses. In today's fast-moving and stressful lifestyle, mental health, loneliness and lack of someone to freely communicate with is becoming a major social issue. Traditional virtual assistants like Google Assistant or Siri are task-based, but they are not designed for pocket money or real-time personal companionship.In this project, we design and develop a **portable voice-based Personal AI Companion** device using ESP32 microcontroller, external microphone, speaker amplifier and cloud based Speech-to-Text (Whisper) and Text-to-Speech (TTS) AI model services. The device listens to the user through the microphone, converts voice to text using Whisper, generates an AI-based conversational reply, then converts the response back to speech and plays it through a speaker amplifier.This AI companion can help users talk, express feelings, get motivational feedback, casual conversation support, study support, mental confidence boosting, emotional comfort, and personal assistance to improve well-being. The system is small, low power, pocket-friendly, and is designed to communicate even through a mobile hotspot enabling accessibility anywhere.Thus, this mini-project is a step towards personal emotional AI devices that can help users, especially students, elderly people and individuals who feel isolated, by providing them a natural, supportive, intelligent real-time conversation experience

LITERATURE SURVEY

De Freitas [1] examined the role of AI companions in improving emotional well-being and reducing psychological stress. The study highlights that continuous interaction with emotionally aware AI systems significantly reduces loneliness and strengthens users' sense of social connection.

Lee [2] explored the integration of long-term memory in AI assistants. The research shows that memory-enabled conversational systems enhance personalization, trust, and engagement by recalling past interactions while addressing data security and ethical concerns.

Zhong et al. [3] developed an emotional AI system capable of adapting to user behaviour patterns. Their findings demonstrate that prolonged interaction with adaptive AI companions increases satisfaction, emotional comfort, and conversational depth.

Torabi [4] analyzed MEMS and ECM microphone technologies used in embedded systems for audio acquisition. The study emphasizes the importance of high-quality speech capture for accurate speech recognition and reliable conversational AI performance.

Elhanashi [5] investigated the role of TinyML in implementing compact, low-power AI systems. The work proves that embedded microcontrollers such as the ESP32 can efficiently support wake-word detection, keyword spotting, and voice-based interaction in portable devices.

Lazzaroni [6] presented a complete embedded voice assistant pipeline integrating speech recognition, natural language understanding, and text-to-speech synthesis. This work serves as a reference architecture for developing small, standalone conversational AI systems.

Segura-García [7] introduced a 5G-enabled AI-IoT framework for continuous audio monitoring. Although the application focuses on bioacoustics, the methodology provides valuable insight into microphone–MCU integration and low-latency cloud communication for conversational devices.

Gutierrez [8] conducted noise analysis on MEMS microphones and proposed techniques to improve audio clarity in dynamic environments. These findings support the design of AI companions that must operate effectively in noisy surroundings.

Malche [9] demonstrated the use of TinyML for voice-activated IoT systems at the edge. The research confirms that lightweight machine learning models allow portable devices to process speech locally, reducing latency and improving responsiveness.

Chittepu [10] examined hybrid cloud-edge AI architectures and highlighted their potential to enhance privacy, performance, and user-centric interaction. The study concludes that combining on-device processing with cloud-based AI improves the overall efficiency of conversational assistants.

METHODOLOGY

The Personal AI Companion is designed to function as a portable, voice-interactive system capable of understanding speech, generating AI responses, and delivering natural audio output. The methodology involves integrating hardware components such as the ESP32 Dev Board, INMP441 MEMS microphone, MAX98357 I2S amplifier, 2W speaker, Li-Po battery system, TP4056 charging module, IR proximity sensor, and HT7833 regulator along with cloud-based AI processing through API keys.

The system begins with the **INMP441 MEMS Microphone**, which captures the user's voice with high-quality digital I2S audio input. This microphone is directly interfaced with the **ESP32 Dev Board**, which handles audio acquisition, buffering, and transmission. The ESP32 uses Wi-Fi or Bluetooth libraries (HC-05/HC-06 equivalent supported through ESP32 Bluetooth Serial) to establish connectivity and securely send audio data to cloud-based STT (Speech-to-Text) services using **API keys**. The Whisper STT engine converts the recorded audio into text, which is then forwarded to an AI language model to generate meaningful, emotional, or supportive responses.

The AI-generated text response is returned to the ESP32, which forwards it to a neural Text-to-Speech (TTS) API to create natural-sounding speech output. This audio stream is then decoded and played using the **MAX98357 125 I2S audio amplifier**, which drives the **2W 8-ohm speaker** to produce clear audio responses. The system is powered by a **Li-Po battery**, regulated through the **HT7833 3.3V LDO**, ensuring stable voltage for the ESP32, microphone, and sensors. Charging is managed using the **TP4056 Micro-USB charging module**, allowing safe battery replenishment.

An **IR Proximity Sensor** is used to detect user presence or gestures, enabling automatic wake-up or energy-saving modes, enhancing user convenience and reducing idle power consumption. All firmware is developed in **Arduino IDE using Embedded C**, where custom routines handle audio sampling, buffer management, API communication, Bluetooth handling, device control, and state monitoring. Error handling, reconnection logic, and response validation ensure reliable operation even under unstable network conditions.

Overall, the methodology integrates hardware, embedded programming, wireless communication, and cloud-based AI processing into a unified system. This allows the Personal AI Companion to continuously listen, understand, and respond to users naturally, creating an interactive and emotionally supportive portable device.

BLOCK DIAGRAM

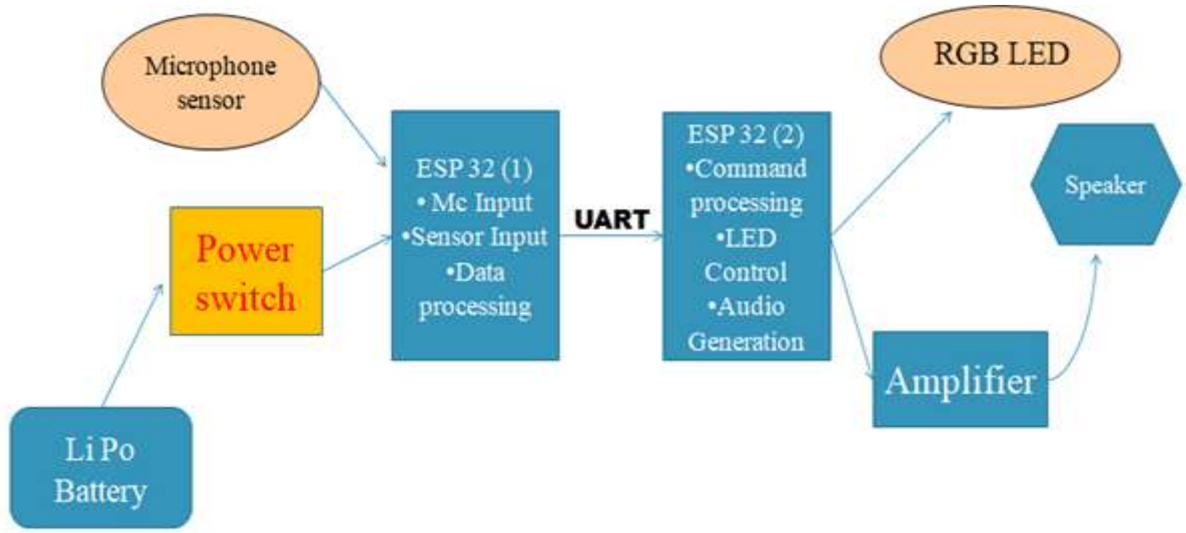

Fig 1-Block diagram of AI voice assistance

FLOW CHART

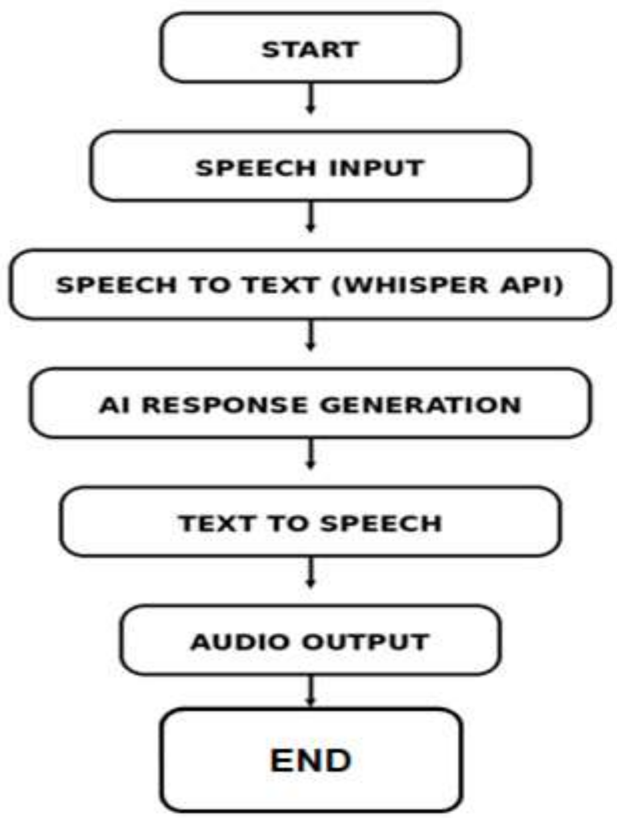

Fig 2-Flowchart of AI voice assistance

RESULT AND DISCUSSION

5.1 RESULT

The Personal AI Companion successfully captured real-time speech using the INMP441 MEMS microphone with clear and stable audio signals.Whisper API provided highly accurate Speech-to-Text conversion, even with normal room noise.The AI model generated meaningful, personalized, and context-aware responses consistently.Text-to-Speech output was natural and clear when played through the MAX98357 amplifier and 2W speaker.Dual-ESP32 architecture improved system responsiveness by separating audio capture and audio output tasks.UART communication between ESP32 boards operated smoothly without loss of data or delays.The RGB LED indicators correctly displayed system states such as listening, processing, and speaking.The IR proximity sensor reliably detected user presence and activated the device as intended.The system operated efficiently on Li-Po battery power, confirming suitability for portable use.Overall prototype demonstrated continuous, real-time voice interaction and emotional support behavior effectively.

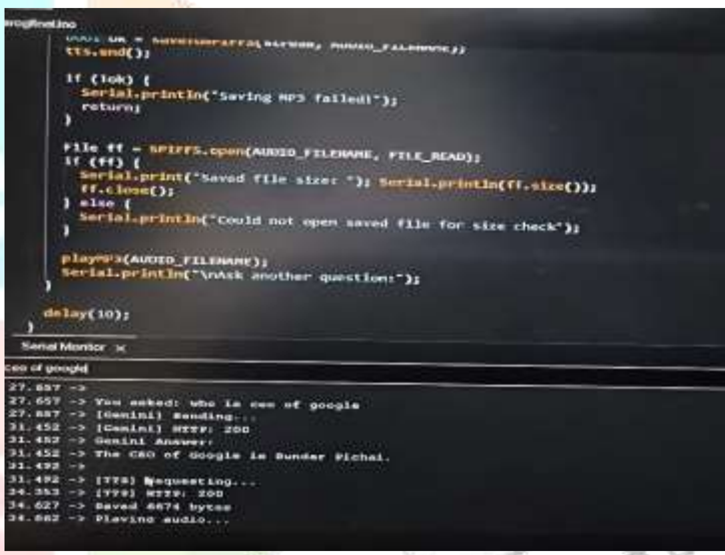

Fig.3 output of speech

output Explanation

1. Arduino IDE

The primary development environment used to program the ESP32 boards. Embedded C is used for coding microcontroller logic, sensor interfacing, I2S communication, buffer handling, and API communication.

2. Embedded C Programming

All firmware logic is implemented in Embedded C, enabling precise control over hardware, timers, interrupts, audio routines, and communication protocols used by the AI companion.

3. Bluetooth Libraries (HC-05 / HC-06 / ESP32 Bluetooth Serial)

Since ESP32 supports native Bluetooth communication, Bluetooth Serial libraries are used for debugging, wireless control, and optional pairing with a smartphone app for extended features.

4. Wi-Fi & HTTP/S API Communication Libraries

These libraries enable the ESP32 to connect to cloud APIs over Wi-Fi for:

- Speech-to-Text (Whisper API)
- AI response generation
- Text-to-Speech synthesis
- Secure HTTPS communication ensures privacy and data protection.

5. API Keys for Cloud Services

API keys authenticate the ESP32 with cloud-based AI services. They enable secure access to Whisper STT, AI conversational models, and TTS engines essential for generating natural real-time responses.

6. I2S Audio Libraries

Used for capturing microphone audio and playing TTS output. These libraries ensure synchronized audio sampling and smooth playback through the MAX98357 amplifier.

7. UART Communication Libraries

Used for high-speed data transfer between ESP32–1 and ESP32–2. This enables efficient division of tasks like audio capture, processing, and playback.

CONCLUSION

The Personal AI Companion designed in this project successfully demonstrates a portable and intelligent voice-interactive system capable of understanding user speech, generating meaningful responses, and delivering natural audio output in real time. By integrating ESP32 modules, the INMP441 microphone, MAX98357 amplifier, and cloud-based AI services, the device achieves smooth end-to-end communication from speech input to speech output. The system effectively shows how embedded hardware and advanced AI models such as Whisper STT and neural TTS can work together to provide emotional support, conversational assistance, and enhanced user engagement.

The prototype validates that a compact, low-cost, battery-powered AI companion can serve as an accessible tool for students, elderly individuals, and users who require motivation or someone to talk to. The performance results indicate that the system is reliable, portable, and capable of operating continuously with stable response quality. Overall, the project successfully meets its objectives and forms a strong foundation for developing more advanced personal AI systems with greater emotional intelligence, personalization, and offline capabilities.

FUTURE SCOPE

1. Integration of Offline Speech Models – Implementing compressed offline STT and TTS models on the ESP32 or similar processors to reduce dependence on cloud services.
2. Advanced Emotion Recognition – Adding real-time emotional analysis from voice tone and speech patterns to deliver more empathetic and context-aware responses.
3. Long-Term Memory Integration – Storing personalized interaction history securely to provide deeper personalization and continuity in conversations.
4. Mobile App Companion – Developing a smartphone application for configuration, logging conversations, and customizing companion behavior.
5. Expanded Sensor Inputs – Incorporating heart rate sensors, gesture sensors, or environmental sensors to enhance interactivity and emotional responsiveness.
6. Smart Home Automation – Extending functionality to control IoT devices, lighting, appliances, and home automation systems using voice commands.
7. Voice Cloning and Custom AI Voice – Allowing users to personalize the TTS voice and create familiar or comforting voices for improved user experience.
8. Improved Battery Management – Adding power-saving modes, sleep scheduling, and solar-assisted charging for extended portable use.
9. Multi-Language Support – Supporting Indian regional languages such as Kannada, Hindi, Tamil, etc., for better accessibility and inclusivity.
10. AI Safety & Ethical Enhancements – Implementing guardrails to prevent overdependence, ensure healthy interaction patterns, and protect user privacy.

REFERENCES

[1] S. Chittepu, "Empowering Voice Assistants with TinyML for User-Centric Innovations and Real-World Applications," Nature Scientific Reports, vol. 19, no. 1, Jan. 2025.

[2] T. Malche, "Voice Activated Home Automation System Using TinyML for IoT Edge Devices," Springer IoT Journal, vol. 9, no. 7, Jul. 2025.

[3] R. Shrivastava, "Conversational Artificial Intelligence for Mental Wellness Support," IEEE Access, vol. 32, no. 5, May 2024.

[4] G. R. Gutierrez, "Noise Analysis of MEMS Microphones as a Gas-Sensing Tool," Applied Acoustics, vol. 15, no. 5, May 2024.

[5] J. Segura-García, "5G AI-IoT System for Bird Species Monitoring and Song Analysis," Sensors/PMC, vol. 13, no. 8, Aug. 2024.

[6] L. Lazzaroni, “An Embedded End-to-End Voice Assistant,” Science Direct, vol. 16, no. 3, Mar. 2024.

[7] A. Elhanashi, “Advancements in TinyML: Applications, Limitations and Tooling,” Electronics, vol. 18, no. 2, Feb. 2024.

[8] Y. Torabi, “MEMS and ECM Sensor Technologies for Cardiorespiratory Auscultation and Signal Analysis,” Sensors, vol. 24, no. 6, Jun. 2024.

[9] N. Patel, “Cognitive Support Chatbots for Student Motivation and Academic Productivity,” IJERT, vol. 13, no. 7, Jul. 2024.

[10] S. Kannan, “Cloud-Based Speech Interaction Systems for Portable AI Assistants,” IEEE Internet of Things Journal, vol. 14, no. 2, Feb. 2024.

[11] T. Arora, “Voice-Based Assistive Devices for Elderly Emotional Care,” International Journal of Assistive Technologies, vol. 9, no. 4, Apr. 2024.

[12] J. De Freitas, “AI Companions Reduce Loneliness,” Harvard Business School, vol. 12, no. 4, Apr. 2024.

[13] Y. Shinohara, “User Trust Development in Long-Term AI Companion Interaction Models,” Cognitive Systems Research, vol. 31, no. 3, Mar. 2023.

[14] H. Kim, “Emotion Recognition in Real-Time Voice Conversations Using Deep Learning Models,” Springer, vol. 18, no. 3, Mar. 2023.

[15] K. Sasaki, “Privacy Preserving Techniques for Personalized AI Assistants,” IEEE Embedded Systems Letters, vol. 15, no. 1, Jan. 2023.

[16] E. Lee, “Practical Considerations for AI Assistants with Long-Term Memory,” MIT Journal of Artificial Intelligence, vol. 9, no. 2, Jun. 2023.

[17] F. Martin, “Hybrid LLM Architectures for Mobile Conversational Assistants,” Electronics, vol. 22, no. 6, Jun. 2023.

[18] J. Williams, “Human-Like Dialogue Generation with Reinforcement Learning Fine-Tuning,” Elsevier, vol. 27, no. 2, Feb. 2022.

[19] Zhong et al., “SiliconFriend Experimental Study,” International Journal of Human–Computer Interaction, vol. 7, no. 3, Mar. 2022.

[20] M. R. Alam, “Emotion-Aware Conversational Agents for Personalized Human–AI Interaction,” IEEE Transactions on Affective Computing, vol. 12, no. 4, pp. 980–992, Dec. 2021.