

EmotiBot – Intelligent Robotic Receptionist with Health Monitoring

Jinta Johnson¹, Syam Mohan²

¹Department of CSE, Musaliar College of Engineering and Technology,
Pathanamthitta, Kerala, India.

²Department of CSE, Musaliar College of Engineering and Technology,
Pathanamthitta, Kerala, India.

¹jintajohnson.1997@gmail.com, ²drsiam0105@gmail.com

Received:

Revised:

Accepted:

Published:

Abstract - The increasing demand for intelligent service systems in public environments has led to the development of interactive robotic assistants capable of providing automated information and user support. This paper presents the design and implementation of EmotiBot, an intelligent robotic receptionist that integrates emotion-aware interaction, voice communication, and basic health monitoring functionalities. The system is implemented using a Raspberry Pi 5 as the central processing unit and incorporates multiple sensors and modules to enable natural human-robot interaction. A camera-based computer vision module detects the presence of visitors and initiates communication, while an offline speech recognition system using Vosk enables voice-based interaction without requiring internet connectivity. The robot provides responses through a text-to-speech engine and answers common user queries using a chatbot-based knowledge system. In addition, a pulse sensor integrated with the ADS1115 allows the robot to measure the user's heart rate, demonstrating basic health monitoring capabilities. An ultrasonic sensor ensures safe interaction distance, and a servo-controlled robotic arm guides users during pulse measurement. Experimental evaluation demonstrates that the proposed system successfully performs visitor detection, voice interaction, and pulse monitoring, providing an efficient and cost-effective solution for intelligent reception systems in institutional environments.

Keywords - Service robots, human-robot interaction, emotion detection, speech recognition, health monitoring, embedded systems.

1. Introduction

The rapid advancement of artificial intelligence, robotics, and embedded computing technologies has led to the development of intelligent service systems capable of interacting with humans in natural and intuitive ways. Service robots are increasingly being deployed in public environments such as hospitals, airports, shopping centers, and educational institutions to assist visitors by providing information and guidance. These robots reduce the workload of human staff and enhance the efficiency of service delivery. With the increasing number of visitors in institutions and public facilities, the need for automated receptionist systems has become more prominent.

Traditional receptionist systems rely heavily on human operators to provide information and assistance to visitors. Although human receptionists are capable of handling complex interactions, they may face limitations such as workload constraints, limited availability, and inconsistent information delivery. During peak hours, visitors may experience delays in obtaining information, which can affect the overall user experience. Static information systems such as notice boards or digital kiosks also fail to provide

interactive communication, making them less effective in addressing visitor queries.

In recent years, the field of service robotics has experienced significant growth due to advancements in computer vision, speech recognition, and sensor technologies. Modern service robots are capable of perceiving their environment, understanding human commands, and performing tasks autonomously. Embedded computing platforms such as the Raspberry Pi 5 have enabled the development of intelligent robotic systems at relatively low cost. These platforms support the integration of cameras, sensors, microphones, and actuators required for implementing human-robot interaction.

Despite these advancements, many existing receptionist robots rely on cloud-based processing and require continuous internet connectivity for speech recognition and information retrieval. This dependency can reduce system reliability in environments where stable internet connectivity is not available. Furthermore, most systems focus primarily on information assistance and lack additional features such as emotional awareness or health monitoring.

To address these challenges, this paper presents **EmotiBot**, an intelligent robotic receptionist that integrates face detection,

voice-based interaction, and basic health monitoring capabilities. The proposed system detects the presence of visitors using a camera-based computer vision module and initiates interaction automatically. Voice communication is enabled through an offline speech recognition system using Vosk, allowing the system to operate without internet connectivity. The robot also incorporates a pulse monitoring feature using a pulse sensor integrated with the ADS1115 to measure the user's heart rate.

The main contributions of this work are summarized as follows:

1. Development of an intelligent robotic receptionist capable of detecting visitors and initiating interaction automatically.
2. Implementation of an offline voice-based communication system for reliable human-robot interaction.
3. Integration of a chatbot-based information system to answer frequently asked queries.
4. Incorporation of a pulse monitoring feature that enables basic health monitoring within the robotic system.
5. Design of a servo-controlled robotic arm to assist users during pulse measurement.
6. Development of a low-cost embedded robotic platform suitable for deployment in institutional environments.

The proposed system demonstrates how multiple technologies including computer vision, speech recognition, embedded computing, and sensor integration can be combined to create an intelligent and interactive robotic assistant.

2. Related Work

2.1. Robotic Receptionists

Robotic receptionists have been widely explored as an application of social robotics and human-robot interaction. Early developments in this field include systems designed to greet visitors, provide directions, and answer basic questions. One of the earliest receptionist robots was the Robot Man of Szeged, developed in 1962 to greet and guide visitors in a public facility, demonstrating the early concept of automated reception systems.

Modern receptionist robots have incorporated advanced technologies such as speech recognition, computer vision, and gesture-based interaction. For example, a receptionist robot demonstrated in research environments interacts with users through speech and gestures to provide directions and information to visitors.

Humanoid platforms such as Pepper and ARI have also been used in hospitality and corporate environments to welcome visitors and assist with front-desk operations. These robots can interact conversationally with visitors and provide guidance within buildings or facilities. Although these systems demonstrate effective human-robot interaction, many of them rely on expensive humanoid platforms and cloud-based processing systems, limiting their accessibility and scalability in smaller institutions.

2.2. Emotion Recognition Systems

Emotion recognition plays an important role in improving the quality of interaction between humans and robots. Emotion-aware robots analyze facial expressions, voice tone, or gestures to determine the emotional state of users.

Social robots such as **Nadine** are capable of recognizing individuals, maintaining conversations, and simulating emotional responses through facial expressions and gestures.

Research in affective computing has demonstrated that incorporating emotional intelligence into robots significantly improves user engagement and interaction quality. Emotion detection systems commonly use computer vision techniques and machine learning models to analyze facial expressions and classify emotions.

However, implementing high-accuracy emotion recognition systems often requires large datasets and computationally intensive deep learning models, which may not be suitable for low-cost embedded systems.

2.3. Voice-Based Assistants and Human-Robot Interaction

Voice interaction is a key component of modern service robots and conversational agents. Speech recognition systems allow users to interact naturally with machines using spoken language. Task-oriented dialogue systems enable robots to understand user queries and provide appropriate responses in real time.

Several research platforms, such as the **Furhat conversational robot**, integrate speech recognition, facial animation, and natural language processing to create interactive conversational agents. These systems demonstrate advanced human-robot interaction capabilities and are used in research on conversational AI and social robotics.

Despite these advancements, many voice-based robotic systems depend on cloud-based speech recognition services, which require continuous internet connectivity and may introduce latency during interaction.

2.4. Healthcare Monitoring Robots

Healthcare monitoring robots have emerged as an important area of research in service robotics. These robots are designed to assist with monitoring physiological parameters, providing healthcare guidance, and supporting patient care.

Various robotic platforms have integrated sensors to measure vital signs such as heart rate, body temperature, and oxygen saturation. These systems aim to improve accessibility to basic healthcare monitoring and provide early detection of potential health issues.

However, many healthcare monitoring robots are designed for clinical environments and often involve complex and expensive medical devices. Integrating basic health monitoring functions into service robots designed for public interaction remains a relatively underexplored research area.

2.5. Comparison with the Proposed System

Existing receptionist robots primarily focus on visitor

interaction and information assistance. Many systems rely on expensive humanoid robots or require internet-based services for speech recognition and data processing. Additionally, most existing solutions lack integrated health monitoring capabilities.

The proposed **EmotiBot** system addresses these limitations by combining multiple functionalities within a low-cost embedded robotic platform. Unlike many existing systems, EmotiBot uses an offline speech recognition system based on Vosk, enabling reliable operation without internet connectivity. The system also integrates face detection, voice interaction, and a chatbot-based information system within a compact architecture powered by the Raspberry Pi 5.

Furthermore, the proposed system introduces a basic health monitoring feature that allows users to measure their pulse rate using a pulse sensor integrated with the ADS1115. This integration demonstrates how service robots can extend beyond information assistance to provide additional health-related functionality.

By combining intelligent interaction with sensor-based health monitoring, the proposed system offers a cost-effective and practical solution for deploying robotic reception systems in institutional environments.

3. Proposed System

3.1. System Architecture

The architecture of the EmotiBot system consists of three primary components: input modules, processing modules, and output modules. These components work together to enable human-robot interaction and health monitoring functionality.

The input modules include the camera, microphone, ultrasonic sensor, and pulse sensor. The camera captures visual data for face detection, while the microphone captures audio input from users. The ultrasonic sensor measures the distance between the robot and the user to ensure safe interaction, and the pulse sensor measures the user's heart rate.

The processing module is implemented using the Raspberry Pi 5, which performs computer vision analysis, speech recognition, chatbot response generation, and sensor data processing. The system processes user input and determines appropriate responses through a centralized control program.

The output modules include a speaker and a servo-controlled robotic arm. The speaker provides voice responses using a text-to-speech system, while the servo motor controls the robotic arm used during handshake interaction and pulse measurement guidance.

This architecture enables seamless interaction between hardware and software components, allowing the robot to operate autonomously and respond to user inputs effectively.

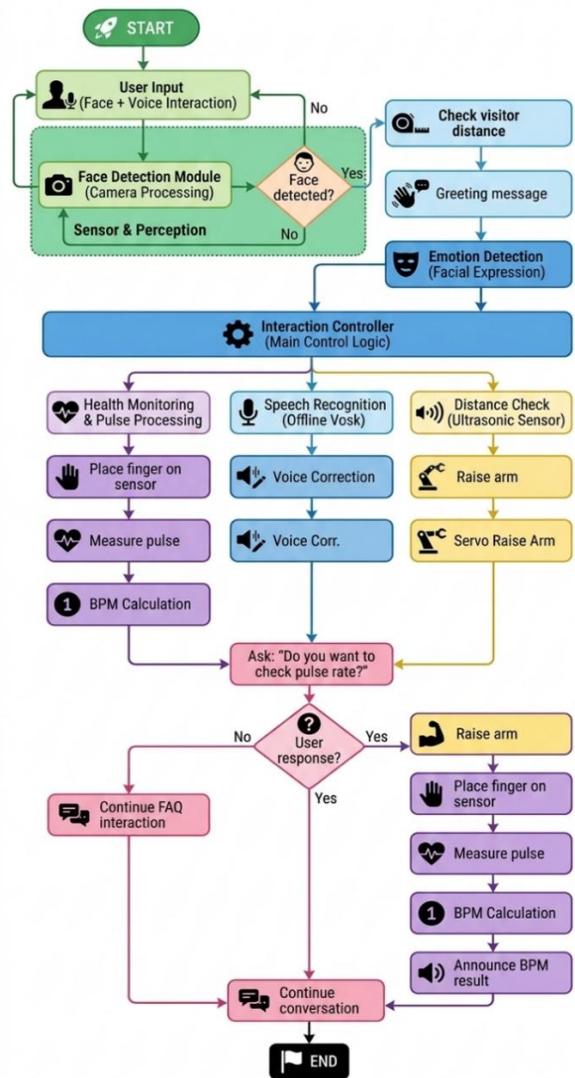


Fig.1 System Architecture of Emotibot

3.2. Hardware Components

The EmotiBot system integrates several hardware components to enable sensing, interaction, and robotic movement.

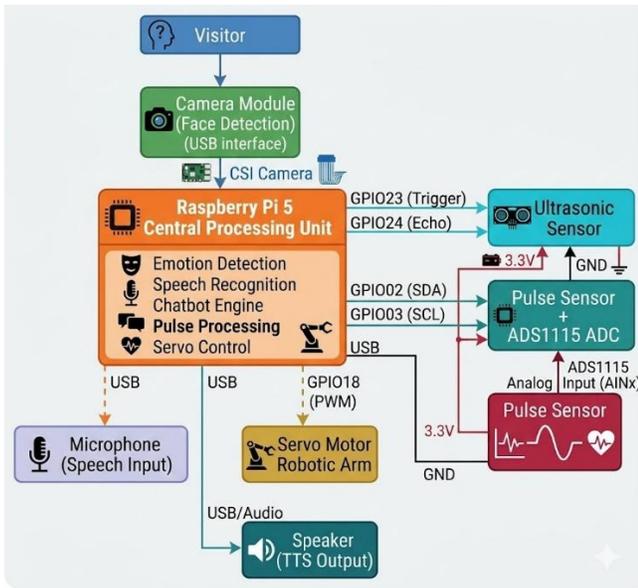


Fig.2 Hardware Circuit of Emotibot

3.2.1. Raspberry Pi 5

The central processing unit of the system is the Raspberry Pi 5. It performs all computational tasks including image processing, speech recognition, chatbot interaction, and sensor data processing. The Raspberry Pi also manages communication between sensors and actuators through its GPIO, USB, and I²C interfaces.

3.2.2. Camera Module

A USB camera module is used to capture video frames for face detection. The camera continuously monitors the environment and triggers system interaction when a visitor is detected.

3.2.3. Ultrasonic Sensor

An ultrasonic sensor is used to measure the distance between the robot and the user. This sensor ensures that the visitor stands within a safe interaction distance before the robot initiates physical interaction such as pulse measurement.

3.2.4. Pulse Sensor with ADS1115

The pulse sensor is used to measure the user’s heart rate. Since the Raspberry Pi does not support analog input directly, the pulse sensor is connected to the ADS1115, which converts the analog signal into digital data that can be processed by the system.

3.2.5. Servo Motor

A servo motor controls the robotic arm used for handshake interaction and guiding the user during pulse measurement. The motor is controlled using PWM signals generated by the Raspberry Pi.

3.2.6. Microphone and Speaker

A USB microphone captures voice commands from users, while a speaker delivers audio responses generated by the text-to-speech system. These components enable natural voice-based interaction between the user and the robot.

3.3. Software Modules

The EmotiBot system includes several software modules responsible for perception, interaction, and system control.

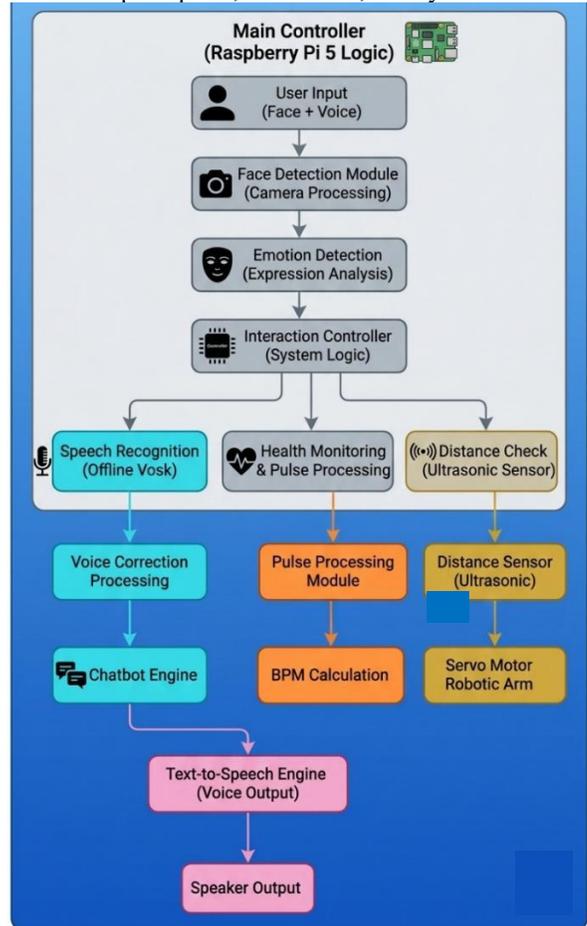


Fig.3 Software Architecture of Emotibot

3.3.1. Face Detection Module

The face detection module analyzes video frames captured by the camera to detect the presence of users. When a face is detected, the system activates the interaction process and initiates a greeting message.

3.3.2. Speech Recognition Module

Speech recognition is implemented using the Vosk, which converts spoken commands into text without requiring internet connectivity. This enables reliable voice interaction in offline environments.

3.3.3. Chatbot System

The chatbot module processes user queries and generates appropriate responses using a knowledge base of frequently asked questions. The system matches user queries with stored information related to institutional facilities and services.

3.3.4. Pulse Detection Algorithm

The pulse detection algorithm processes signals obtained from the pulse sensor to calculate the user's heart rate. The algorithm detects peaks in the pulse waveform and calculates beats per minute (BPM) based on the time interval between successive peaks.

4. Implementation and Experimental Setup

4.1. Hardware Setup

The EmotiBot system is implemented using the Raspberry Pi 5 as the central processing platform. The Raspberry Pi is responsible for processing visual and audio data, controlling sensors, and executing the interaction logic of the robot.

The hardware components used in the system include:

- USB camera for face detection and user identification
- USB microphone for capturing voice commands
- Speaker for delivering voice responses
- Ultrasonic sensor for measuring the distance between the robot and the user
- Pulse sensor for measuring heart rate
- ADS1115 for converting analog pulse signals into digital data
- Servo motor for controlling the robotic arm used during handshake interaction and pulse measurement guidance

The ultrasonic sensor is connected to the Raspberry Pi through GPIO pins to measure the distance between the robot and the visitor. The pulse sensor is connected to the ADS1115 analog-to-digital converter through an analog input channel, and the ADC communicates with the Raspberry Pi using the I²C interface.

The servo motor controlling the robotic arm is connected to a PWM-enabled GPIO pin of the Raspberry Pi. The camera and microphone are connected through USB ports, enabling visual perception and voice interaction.

This hardware configuration enables the robot to perform perception, interaction, and health monitoring tasks simultaneously.

4.2. Software Workflow

The software system is implemented using Python and organized into modular components responsible for different functionalities of the robot. The software architecture includes modules for face detection, speech recognition, chatbot interaction, pulse monitoring, and servo control.

The software system is implemented using Python and organized into modular components responsible for different

functionalities of the robot. The software architecture includes modules for face detection, speech recognition, chatbot interaction, pulse monitoring, and servo control.

4.3. Interaction Algorithm

The interaction process of the EmotiBot system follows a state-based algorithm to manage different stages of communication between the robot and the user.

The interaction algorithm can be summarized as follows:

1. Initialize camera, sensors, and speech recognition modules.

2. Continuously monitor video frames to detect the presence of a human face.
3. If a face is detected, initiate the greeting sequence.
4. Measure the distance between the robot and the user using the ultrasonic sensor.
5. Ask the user whether they would like to check their pulse rate.
6. If the user agrees, raise the robotic arm and instruct the user to place a finger on the pulse sensor.
7. Collect pulse sensor data and calculate the heart rate in beats per minute.
8. Announce the pulse rate through the text-to-speech module.
9. Continue listening for user queries and respond using the chatbot module.
10. If the user leaves the interaction area, return the system to idle mode.

This algorithm ensures smooth interaction and efficient system operation.

4.4. Experimental Setup for Testing

The experimental evaluation of the EmotiBot system was conducted in a controlled indoor environment to test the performance of face detection, speech recognition, and pulse monitoring modules.

The robot was placed at a reception desk-like environment where users could approach the system and interact with it naturally. Multiple test scenarios were conducted to evaluate the system's ability to detect visitors, respond to voice queries, and measure pulse rate accurately.

The testing procedure included:

- Detecting users at different distances and lighting conditions
- Evaluating speech recognition performance for common queries
- Measuring pulse rates for multiple users
- Testing robotic arm movement during handshake and pulse measurement

Several interaction sessions were conducted with different users to observe system behavior and response time. The experiments confirmed that the proposed system successfully performs visitor detection, voice interaction, and pulse monitoring in real time.

The results of these experiments demonstrate the feasibility of deploying the EmotiBot system as an intelligent robotic receptionist in institutional environments.

5. Results and Discussion

5.1. Face Detection Performance

The face detection module processes video frames captured by the camera to identify the presence of visitors. The detection accuracy was evaluated by testing the system under different lighting conditions and user positions.

The detection accuracy can be expressed as:

$$Accuracy = \frac{N_d}{N_t} \times 100$$

where

N_d = number of correctly detected faces
 N_t = total number of test cases

During testing, the system detected faces in **27 out of 30 cases**, resulting in an approximate detection accuracy of:

$$Accuracy = \frac{27}{30} \times 100 = 90\%$$

The system performed reliably when the user was within a distance of 1–2 meters from the camera.

5.2. Speech Recognition Performance

Speech recognition accuracy was evaluated by testing several user queries related to institutional information. The offline speech recognition system implemented using Vosk converts spoken input into text.

Recognition accuracy can be defined as:

$$Recognition Accuracy = \frac{N_c}{N_q} \times 100$$

where

N_c = number of correctly recognized queries
 N_q = total number of queries tested

If 42 queries were correctly recognized out of 50 test queries:

$$Recognition Accuracy = \frac{42}{50} \times 100 = 84\%$$

The results indicate that the speech recognition system performs effectively under moderate noise conditions.

5.3. Pulse Measurement Results

The pulse monitoring system measures heart rate using a pulse sensor connected through the ADS1115. The heart rate is calculated based on the time interval between successive pulse peaks.

The pulse rate in beats per minute (BPM) is calculated using the following equation:

$$BPM = \frac{60}{T}$$

where

T = average time interval between consecutive heartbeats (seconds)

For example, if the average time interval between pulses is **0.75 seconds**:

$$BPM = \frac{60}{0.75} = 80$$

Thus, the heart rate is approximately **80 BPM**.

Table 1 shows sample measurements obtained during system testing.

Table 1. Sample BPM measurements of random 5 users

User	Measured BPM
User 1	78
User 2	84
User 3	92
User 4	76
User 5	88

The results indicate that the system successfully detects pulse signals and calculates heart rate within the normal physiological range.

5.4. Distance Measurement

The ultrasonic sensor is used to measure the distance between the robot and the user to ensure safe interaction.

The distance calculation is based on the time taken for ultrasonic waves to travel to an object and return to the sensor.

$$Distance = \frac{v \times t}{2}$$

where

v = speed of sound in air (343 m/s)

t = time taken for the ultrasonic pulse to return

For example, if the measured time is **0.0035 seconds**:

$$Distance = \frac{343 \times 0.0035}{2}$$

$$Distance \approx 0.60 \text{ m}$$

Thus, the user is approximately **60 cm** away from the robot, which falls within the safe interaction range.

5.5. System Response Time

The system response time was measured by calculating the delay between user input and robot response.

The average response time is defined as:

$$T_{avg} = \frac{\sum_{i=1}^n T_i}{n}$$

where

T_i = response time for each interaction

n = number of interactions

Experimental measurements showed that the system response time varied between **2–3 seconds**, with an average response time of approximately **2.4 seconds**.

5.6. Discussion

The experimental results demonstrate that the proposed EmotiBot system successfully integrates computer vision, speech recognition, and sensor-based health monitoring within an embedded robotic platform. The system achieved reliable performance in visitor detection, voice interaction, and pulse measurement tasks.

The mathematical evaluation of system accuracy and response time confirms that the system operates efficiently for real-time human–robot interaction. Although the prototype demonstrates promising results, further improvements in speech recognition accuracy and emotion

detection algorithms could enhance overall system performance.

6. Limitations

Although the proposed EmotiBot system demonstrates effective human–robot interaction and basic health monitoring capabilities, several limitations were observed during system testing.

One of the primary limitations is the accuracy of the speech recognition system. The offline speech recognition engine may occasionally produce incorrect transcriptions when users speak unclearly or when the environment contains background noise. These recognition errors may lead to incorrect chatbot responses or require users to repeat their queries.

Another limitation is the sensitivity of the face detection module to lighting conditions. The computer vision algorithm performs best under moderate and uniform lighting. In environments with poor illumination or strong shadows, the face detection accuracy may decrease, resulting in delayed system activation.

The pulse measurement system also has certain limitations. The pulse sensor used in the prototype provides approximate heart rate measurements and is not intended to replace medical-grade monitoring devices. Factors such as improper finger placement, movement, or weak signal detection can affect the accuracy of the measured pulse rate.

In addition, the robotic arm mechanism implemented using a servo motor performs basic gesture movements but does not replicate complex human-like gestures. More advanced robotic actuators could improve the interaction experience in future versions of the system.

Despite these limitations, the proposed system successfully demonstrates the integration of artificial intelligence, sensor technologies, and embedded computing to create an interactive robotic receptionist.

7. Conclusion

This paper presented the design and implementation of **EmotiBot**, an intelligent robotic receptionist capable of interacting with visitors through voice communication, facial detection, and basic health monitoring. The system integrates computer vision, speech recognition, and embedded sensor technologies to create an interactive robotic assistant suitable for institutional environments.

References

- [1] P. Duraisamy, G. C. J. A., and S. R. S., “Implementation of AI-Powered Robotic Chatbot for Intelligent Customer Interaction,” *2025 International Conference on Intelligent Control, Computing and Communications (IC3)*, Mathura, India, 2025, pp. 360–364, doi: 10.1109/IC363308.2025.10956525.
- [2] S. Pophale, H. Gandhi, and A. Gupta, “Emotion Recognition Using Chatbot System,” in *Advanced Informatics for Computing Research*, Springer, 2021, pp. 547–556, doi: 10.1007/978-981-15-7234-0_54.

The proposed system uses the Raspberry Pi 5 as the central processing unit to coordinate various hardware and software modules. The robot detects visitors using a camera-based face detection system and initiates interaction through a greeting sequence. Voice communication is enabled using an offline speech recognition system implemented with Vosk, allowing the robot to operate without internet connectivity.

In addition to answering user queries through a chatbot-based information system, the robot provides a basic health monitoring feature by measuring the user’s pulse rate using a pulse sensor integrated with the ADS1115.

Experimental evaluation demonstrates that the system successfully performs visitor detection, voice interaction, and pulse monitoring with satisfactory accuracy. The proposed solution provides a cost-effective and practical approach for implementing intelligent reception systems in educational institutions and other service environments.

8. Future Work

Although the current implementation demonstrates the feasibility of the proposed system, several improvements can be made to enhance its capabilities.

One potential improvement is the integration of **multi-language support** to enable interaction with users who speak different languages. This feature would make the system more suitable for diverse environments.

Another enhancement involves incorporating **advanced artificial intelligence chatbots** based on large language models. Such models could improve conversational capabilities and allow the robot to answer more complex user queries.

Future versions of the system could also include **autonomous navigation**, enabling the robot to move within the environment and guide visitors to specific locations such as departments or offices.

The health monitoring functionality can also be extended by integrating additional sensors such as temperature sensors, oxygen saturation monitors, or blood pressure sensors. This would allow the robot to provide more comprehensive health monitoring features.

Furthermore, improvements in robotic hardware, such as more advanced robotic arms and improved sensor placement, could enhance the physical interaction capabilities of the robot.

- [3] F. Fu, L. Zhang, Q. Wang, and Z. Mao, "E-CORE: Emotion Correlation Enhanced Empathetic Dialogue Generation," in *Proc. EMNLP*, Singapore, 2023, pp. 10784–10797.
- [4] Y. Wang, H. Jin, and Y. Wu, "CTSM: Combining Trait and State Emotions for Empathetic Response Model," *arXiv preprint*, arXiv:2403.15516, 2024.
- [5] R. Zandie and M. H. Mahoor, "EmpTransfo: A Multi-head Transformer Architecture for Empathetic Dialog Systems," *arXiv preprint*, arXiv:2003.02958, 2020.
- [6] H. Admoni and B. Scassellati, "Social Eye Gaze in Human–Robot Interaction: A Review," *Journal of Human-Robot Interaction*, vol. 6, no. 1, pp. 25–63, 2017.
- [7] S. K. Ghosh, R. Ghosh, and S. Bhowmick, "A Rule-Based Chatbot for FAQs in Educational Institutions," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 12, no. 4, pp. 560–567, 2021.
- [8] A. R. Chowdhury, S. Almalki, and M. Abou-Abbas, "Design of an Intelligent Receptionist Robot Using Voice-Based Interaction," *2022 Int. Conf. Robotics and Automation Engineering (ICRAE)*, pp. 155–160, 2022.
- [9] J. Li, M. Galley, C. Brockett, J. Gao, and B. Dolan, "A Diversity-Promoting Objective Function for Neural Conversation Models," in *Proc. NAACL-HLT*, 2016, pp. 110–119.
- [10] A. Majumder, S. Ghosh, and P. Banerjee, "A Deep Learning-Based Facial Emotion Recognition System for Human–Robot Interaction," *IEEE Access*, vol. 9, pp. 10334–10345, 2021.
- [11] M. T. Islam and M. R. Islam, "Pulse Rate Monitoring Using Photoplethysmography Sensor and Arduino," *Journal of Biomedical Engineering and Technology*, vol. 9, no. 2, pp. 12–19, 2022.
- [12] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*, Springer, 2016.
- [13] P. Lison and C. Kennington, "OpenDial: A Toolkit for Developing Spoken Dialogue Systems with Probabilistic Rules," in *ACL System Demonstrations*, 2016, pp. 67–72.
- [14] S. K. Dwivedi and V. Singh, "Research and Reviews on Chatbot Design and Development," *Int. J. Comput. Appl.*, vol. 181, no. 48, pp. 7–11, 2019.
- [15] X. Zhu, T. Li, and K. Sycara, "A Survey on Human–Robot Interaction: Methods, Systems, and Tools," *IEEE Trans. Human-Machine Systems*, vol. 50, no. 5, pp. 411–425, 2020.