

Maxbot: A Multimodal Interactive Robot for Personalized Assistance Using Voice, Facial Recognition and Touch Inputs

N.Swathi¹, A. Sandeepraju², K.Madhukeri³, M.Syayagoad⁴

Assistant Professor, Department of CSE, ACE Engineering College, Hyderabad, Telangana, India¹

IV B. Tech Students, Department of CSE, ACE Engineering College, Hyderabad, Telangana, India^{2,3,4}

ABSTRACT: This project focuses on the design and development of an interactive desktop robot built from scratch using a Raspberry Pi. The robot is designed to interact with users through multiple modes such as voice commands, facial recognition, and touch-based inputs. A USB microphone is used for capturing voice commands, while text-to-speech techniques allow the robot to respond audibly. A camera module combined with OpenCV enables face detection and recognition, allowing personalized interactions. Capacitive touch sensors provide physical interaction to trigger actions like image capture or responses. The project demonstrates how embedded systems and AI techniques can be combined to build an intelligent and user-friendly robotic assistant suitable for educational and personal applications.

KEYWORDS: Maxbot, Multimodal, robot, assistant

I. INTRODUCTION

Robotics and artificial intelligence are rapidly advancing fields that are transforming the way humans interact with technology. Modern robotic systems are designed to perform intelligent tasks and assist humans in various environments such as homes, educational institutions, and workplaces. Interactive robots are becoming increasingly popular because they can communicate with users, recognize faces, understand voice commands, and respond to physical interactions.

This project focuses on the development of an **interactive mini desktop robot** that can communicate with users through speech, facial recognition, and touch-based inputs. The system is built using a **Raspberry Pi** as the main processing unit, along with a USB microphone, Raspberry Pi camera module, and capacitive touch sensors. The robot uses technologies such as **OpenCV for face recognition**, **speech recognition for voice commands**, and **pyttsx3 for text-to-speech responses**. These technologies enable the robot to detect users, understand commands, and respond intelligently.

II. LITERATURE SURVEY

Recent studies have investigated the application of Artificial Intelligence, Natural Language Processing (NLP), and Large Language Models (LLMs) in Requirement Engineering to automate tasks such as requirement extraction, analysis, and documentation.

The research paper “LLMs for Requirements Engineering: A Systematic Literature Review” examines the application of LLMs in requirement analysis, documentation, and validation. However, it mostly reviews existing research and lacks implementation.[1]

The research paper “From Requirements to Test Cases with NLP” applies NLP techniques such as Named Entity Recognition (NER) and rule-based extraction to automatically create test cases from requirement descriptions. Although successful, it mostly applies to requirement descriptions in specific formats.[2]

The ReXCL tool is specifically designed to apply heuristic techniques and machine learning algorithms to extract

and classify information from requirement documents. However, it still lacks general validation.[3]In summary, existing research mostly targets requirement extraction and task-specific automation. There is limited research on the automatic generation of structured Product Requirement Documents (PRDs) from unstructured conversations, which is addressed by the proposed AutoPRD system utilizing NLP techniques.

III. METHODOLOGY

The methodology of the interactive desktop robot project describes the systematic process used to design, develop, and implement the system. The development begins with the selection of appropriate hardware components such as the Raspberry Pi, camera module, USB microphone, speaker, and capacitive touch sensors. These components are connected to the Raspberry Pi, which acts as the central controller responsible for managing all system operations.

After setting up the hardware, the software environment is configured using the Python programming language and required libraries. Libraries such as OpenCV and face recognition are used to implement facial detection and identification features through the camera module. Speech recognition libraries are used to capture and interpret voice commands from the microphone, while the pyttsx3 library is used to generate spoken responses from the robot. Additional Python modules are used to handle sensor inputs and system control.

The system then integrates both hardware and software modules to enable smooth interaction between the robot and the user. The robot continuously monitors inputs from the microphone, camera, and touch sensors. When an input is detected, the system processes the data and determines the appropriate action. Finally, the robot responds through speech output, image capture, or other programmed actions, enabling an interactive and intelligent human-robot communication system.

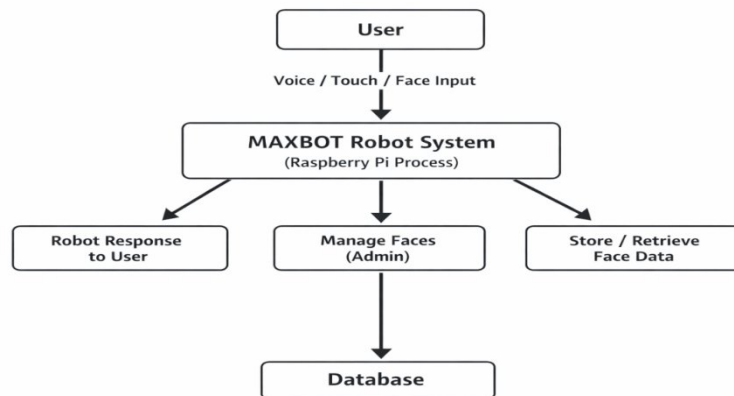
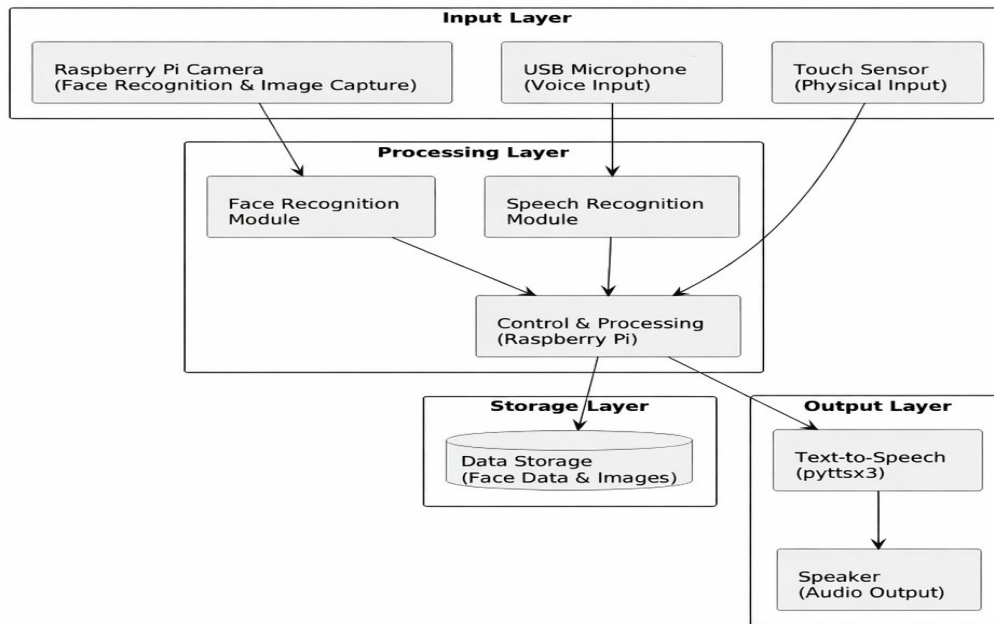


Fig 3.1 Automated Product Requirement Document Generation Methodology

IV. SYSTEM ARCHITECTURE

The system architecture of the interactive desktop robot explains how different hardware and software components are organized and connected to perform intelligent interactions with users. The **Raspberry Pi** acts as the central processing unit that manages all system operations and coordinates communication between input devices, processing modules, and output components.

The input layer of the system consists of devices such as the **USB microphone**, **Raspberry Pi camera module**, and **capacitive touch sensors**. These components collect information from the user in the form of voice commands, facial images, or physical touch interactions. The captured data is sent to the Raspberry Pi where it is processed using software libraries such as OpenCV for face recognition and speech recognition modules for voice command analysis.

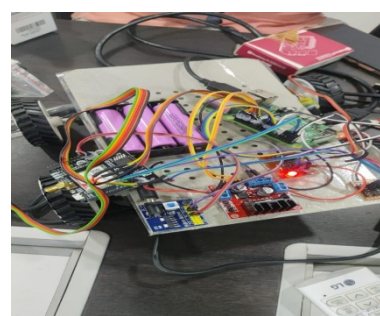


After processing the input data, the system generates appropriate outputs through the output layer. The robot provides responses using a **speaker with text-to-speech technology (pyttsx3)** or performs actions such as capturing and storing images. This architecture ensures that all components work together efficiently, enabling the robot to interact intelligently with users through voice, vision, and touch-based inputs.

V. OUTPUT

The output of the interactive desktop robot system is the response generated after processing user inputs such as voice commands, facial recognition, or touch interactions. Once the Raspberry Pi processes the input data using the required software libraries, the system produces an appropriate response through different output devices. The main output of the system is **voice responses generated using the pyttsx3 text-to-speech library**, which allows the robot to communicate with users in a natural way.

In addition to voice responses, the robot can also produce outputs such as **capturing and storing images through the Raspberry Pi camera module** when a command like taking a photo is given. The system may also trigger specific programmed actions when touch sensors are activated. These outputs ensure that the robot responds effectively to user inputs and provides an interactive and intelligent user experience.



Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJRSET.2026.1504235|

VI. CONCLUSION

In this project, an automated system named AutoPRD was created to produce Product Requirement Documents using Machine Learning and Natural Language Processing approaches. The process of manually writing PRDs is a time-consuming process that involves collecting data from different sources and presenting it in a structured form. The proposed system makes this process easier by automatically analyzing unstructured text inputs and extracting relevant information about key requirements. By leveraging NLP approaches and transformer models, the system is capable of extracting product features, functional requirements, and development tasks. The PRDs produced by the system enable development teams to easily comprehend product requirements and facilitate collaboration between different

VII. FUTURE SCOPE

The interactive desktop robot system can be further improved by adding advanced technologies and additional features to enhance its capabilities. Future improvements can focus on making the robot more intelligent, efficient, and capable of handling complex interactions with users. One possible enhancement is the integration of **advanced artificial intelligence and machine learning algorithms** to improve voice recognition accuracy and facial recognition performance. This would allow the robot to better understand user commands and recognize a larger number of users. The system can also be upgraded to support **natural language processing (NLP)** so that the robot can understand more complex sentences and respond more naturally during conversations.

Another enhancement is the addition of **more sensors and hardware components** such as motion sensors, temperature sensors, or gesture recognition systems. These additions would allow the robot to detect environmental changes and interact with users in more advanced ways. The robot could also be connected to **cloud services or IoT platforms** to access online information, perform smart home control, or store user data securely. Future versions of the robot could also include a **graphical display or touchscreen interface** to provide visual feedback along with voice responses. Improvements in system optimization and hardware upgrades can further enhance performance, making the robot more responsive and capable of handling multiple tasks simultaneously. These enhancements would make the robot more useful for applications such as education, personal assistance, and smart environment interaction.

REFERENCES

1. IEEE Explore Digital Library (2020). Humanoid Robot Research Papers. IEEE Robotics and Automation Society Publications.
2. Honda Motor Co., Ltd. (2000). ASIMO – Advanced Step in Innovative Mobility. Honda Robotics Research Documentation.
3. SoftBank Robotics (2014). Pepper – The Humanoid Robot Designed for Human Interaction. Social Robotics White Paper.
4. Boston Dynamics (2013). Atlas: The World's Most Dynamic Humanoid Robot. Technical Reports on Humanoid Mobility.
5. Craig, J. J. (2005). Introduction to Robotics: Mechanics and Control. Pearson Education.
6. Quigley, M. et al. (2009). ROS: An Open-Source Robot Operating System. IEEE International Conference on Robotics and Automation.
7. Bonaccorso, G. (2017). Machine Learning Algorithms for Robotics Applications. Packt Publishing.
8. Raspberry Pi Foundation (2021). Raspberry Pi 4 Model B – Technical Documentation. Official Datasheet
9. Corke, P. (2017). Robotics, Vision and Control: Fundamental Algorithms in MATLAB. Springer Tracts in Advanced Robotics.
10. OpenAI Robotics Research. (2021). Learning-Based Control and Human–Robot Interaction Models



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details