

MedCyBot: Integrating Autonomous Multifunctional Robot for Enhanced Healthcare Communication

Dr. P. Bhuvaneshwari

Professor, Department of Biomedical Engineering
ACS College of Engineering
Bangalore, India
bhuvanasaamuel@acsce.edu.in

Linda Sunil

Student, Department of Biomedical Engineering
ACS College of Engineering
Bangalore, India
linda.sunil01@gmail.com

Mrs. Kavya N.

Asst. Professor, Department of Biomedical Engineering
ACS College of Engineering
Bangalore, India
kavyanacs@gmail.com

Himabindu K.

Student, Department of Biomedical Engineering
ACS College of Engineering
Bangalore, India
bindhuks99@gmail.com

Abstract—MedCyBot is a prototype healthcare assistive robot developed to support basic communication and routine assistance tasks in clinical settings. The prototype combines speech-to-text, text-to-speech, multidirectional movement, and a SCARA-based handling unit in one platform. Mecanum wheels were used in the base to improve mobility, while the upper section supported voice interaction, display output, and small-item handling. In the final prototype, an Arduino Uno was used to test both the base and upper-section functions. The work focused on showing how communication support and simple physical assistance could be brought together in a low-cost prototype. This paper presents the motivation, methodology, implementation, results, and limitations of the system.

Index Terms—Assistive healthcare robot, Arduino Uno, mecanum wheels, SCARA, speech-to-text, text-to-speech.

I. INTRODUCTION

Assistive robotics is gaining increasing relevance in healthcare because it can support communication, reduce repetitive workload, and improve the handling of routine tasks in hospitals and care centres. As shown in Fig. 1, the healthcare assistive robot market increased from \$410.6 million in 2021 to \$514.3 million in 2022. This growth reflects the need for systems that can improve efficiency and accessibility in healthcare environments.

In many healthcare settings, communication barriers, limited accessibility, navigation difficulties, and delays in handling medicines or supplies can affect the quality of care. These challenges become more noticeable in environments where healthcare staff must manage both patient interaction and repeated support tasks.

To address this, we developed MedCyBot as a prototype assistive healthcare robot that combines communication support and physical task support in a single system. The main functions demonstrated in the prototype were speech-to-text, text-to-speech, multidirectional movement, app-based control, and SCARA-based handling. The broader design also considered future expansion toward language translation and

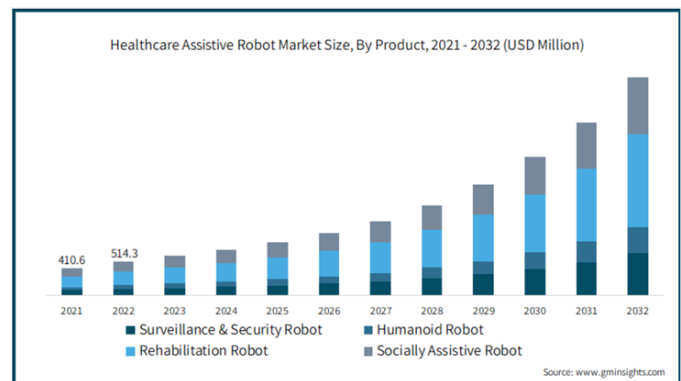


Fig. 1. Healthcare assistive robot market size figure included in the project report.

autonomous navigation. This paper presents the motivation, literature background, problem statement, methodology, prototype implementation, results, and limitations of MedCyBot.

II. LITERATURE SURVEY

Recent work in assistive robotics has focused on communication support, autonomous navigation, translation, and lightweight manipulation. Dhanjal and Singh developed a multilingual speech-to-Indian-sign-language system and reported recognition accuracies of 91%, 89%, and 89% for English, Punjabi, and Hindi, respectively [1]. Other studies on speech recognition and speech synthesis for assistive applications have shown that voice-based systems can improve accessibility and help users with disabilities interact more easily with digital platforms [2], [3], [19]–[21].

Language translation has also been explored as a way to reduce communication barriers. Earlier studies have described translator applications and neural machine translation methods that support multilingual interaction [14]–[16]. This is particularly relevant in healthcare, where even simple instructions and

responses may be affected by differences in language between patients and caregivers.

Speech-based systems have also been studied for medical documentation and workflow support. Sushmita Kulkarni, Dataprasad A. Torse, and Deepak Kulkarni presented a cloud-based transcription system that converted spoken input into text for medical records [4]. In addition, models such as Listen, Attend and Spell showed that end-to-end speech transcription can be improved using neural-network-based learning [5]. Together, these studies highlight the importance of speech-processing modules in assistive healthcare systems.

Navigation and mobility form another important part of the related work. SLAM-based approaches have been used for mapping, localization, and path planning in autonomous robots [6]–[10]. For handling tasks, SCARA-based systems are useful because they provide a relatively simple structure for pick-and-place operations and can be developed economically for controlled motion [11]–[13]. In healthcare robotics, Shubha and Meenakshi presented an assistive robot for medicine delivery to bedridden patients, while Pawan Kadam, Pratik Padalkar, Aniket Mohite, Shantanu Mirajgave, and Santwana Gudadhe discussed a medical assistance robot using ROS for multi-purpose applications [17], [18].

Overall, the literature shows that many existing systems focus on one main function, such as speech interaction, translation, navigation, or item transport. MedCyBot was developed as a student prototype to bring several of these functions together within a single assistive platform.

III. PROBLEM STATEMENT AND OBJECTIVES

Healthcare environments continue to face practical challenges that can affect patient care. These include communication barriers, limited accessibility, difficulty in handling medicines and supplies, and the need to move efficiently through complex spaces. Language differences between patients and caregivers can reduce the clarity of interaction, while inefficient transport of routine items can increase workload and delay simple tasks.

The main objective of MedCyBot was to develop a prototype assistive healthcare robot that could address these challenges through an integrated design. The project aimed to support speech-to-text and text-to-speech interaction, enable multidirectional movement using mecanum wheels, and include a SCARA-based section for handling small items such as medicines or supplies.

A second objective was to combine these functions in a form suitable for prototype-level demonstration. Instead of focusing only on communication or only on mobility, the project explored how voice interaction, movement control, and basic handling could work together within one assistive robotic platform.

IV. METHODOLOGY

The methodology followed in this project was divided into three main stages: acquisition, training, and testing. The overall workflow used in the project is shown in Fig. 2. In the



Fig. 2. Overall methodology followed in the MedCyBot project.

acquisition stage, the required hardware and software elements for movement, communication, and handling were assembled into the prototype. In the training stage, the speech-processing tools were prepared for the intended application. In the testing stage, the integrated functions of the system were evaluated at prototype level.

A. MedCyBot Base

The base of MedCyBot was designed for omnidirectional movement. It used four mecanum wheels driven by four N20 geared motors. An Arduino Uno was used to control the base, while a motor driver managed communication between the controller and the motors. Wireless control was provided through an HC-05 Bluetooth module. The broader design also considered SLAM-based navigation and obstacle avoidance as part of the overall system direction.

The mecanum wheels allowed the robot to move in multiple directions rather than relying only on conventional forward turning. This was useful in hospital-like environments, where the robot may need to pass through tight spaces and move around obstacles during routine assistance tasks.

B. Upper Section and Training Flow

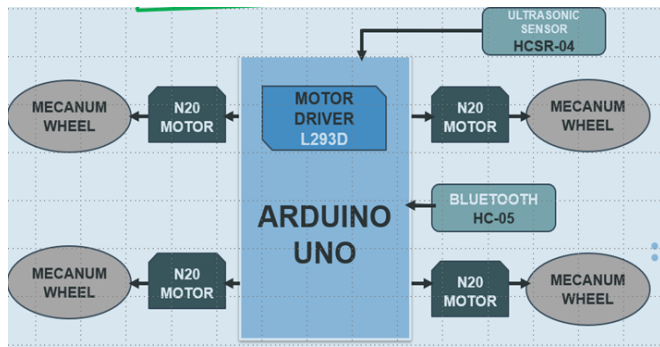
The upper section integrated the communication and handling modules of the robot. It included a MAX9814 microphone module, speaker, a 16×4 display, and a SCARA-based handling mechanism. Speech-to-text was implemented using CMUSphinx, while text-to-speech was implemented using pyttsx3.

In the final prototype, an Arduino Uno was used to test both the base and upper-section functions. Although an ESP32 was considered during the design stage, it was not used in the final prototype implementation.

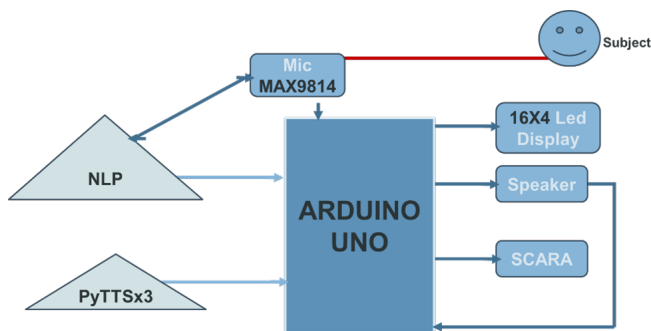
The training stage focused on preparing the speech-processing tools for the intended application. CMUSphinx was associated with Sphinxtrain, and the workflow involved dataset preparation for speech-related processing. The broader project work also used tools such as Jupyter Notebook and Arduino IDE for coding, testing, and system integration during development. The base-level and upper-section arrangements used in the prototype are shown in Fig. 3.

V. IMPLEMENTATION OVERVIEW

The implementation of MedCyBot was organized around three main functions: communication, movement, and handling. The communication block included voice capture through the microphone, conversion of speech into text, display of the recognised output, and spoken output through a speaker. The movement block involved Bluetooth-based control of the mecanum-wheel base through a mobile application.



(a) Base level arrangement used for MedCyBot.



(b) Functional arrangement of the upper section.

Fig. 3. System-level diagrams included in the project documents.

The handling block used a SCARA-based arrangement for the transport of small items.

The robot was developed as a modular structure, with the lower section responsible for locomotion and the upper section responsible for interaction and handling. This made the prototype easier to build and test, while still allowing the different subsystems to work together as part of one integrated platform.

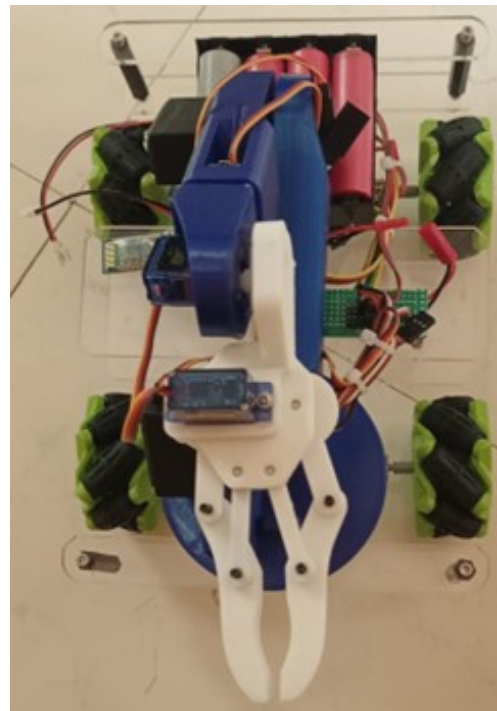
A. Software Environment

The software environment included Arduino IDE for programming and uploading code to the controller, and Jupyter Notebook for development, output review, and testing. These tools supported the integration of the communication functions with the physical hardware used in the prototype.

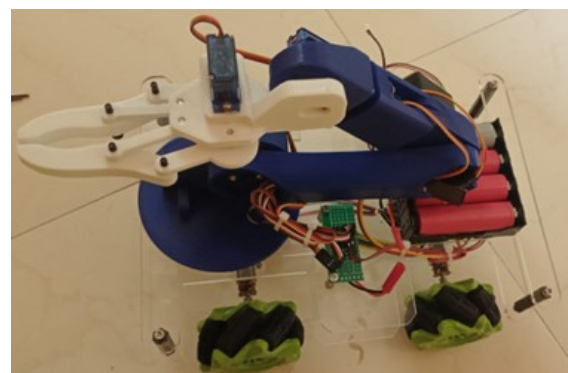
VI. RESULTS AND DISCUSSION

The prototype-level results showed that MedCyBot was able to bring together communication, movement, and handling functions within a single platform. Speech-to-text was implemented so that spoken input could be converted into text, and text-to-speech was used to provide audio output from the system. These features supported basic user interaction through voice and display. The physical prototype is shown in Fig. 4.

The mecanum-wheel base, together with app-based control, allowed the robot to move with multidirectional capability. This improved manoeuvrability and supported movement in confined spaces. The SCARA-based section was used as the



(a) Top view of the MedCyBot prototype.



(b) Side view of the MedCyBot prototype.

Fig. 4. Prototype views included in the project report.

handling component of the prototype and represented the system's task-support function.

The project was also supported by controller application screens, code windows, circuit-level views, and prototype photographs. These elements show that the work involved more than mechanical assembly alone; it also included embedded control, interface-level testing, and software integration. The controller applications used during testing are shown in Fig. 5.

The present results are qualitative rather than fully numerical. The main outcome of the work was successful subsystem integration: voice input, voice output, movement control, and handling were demonstrated as parts of one prototype. Detailed numerical evaluation of recognition accuracy, navigation error, transport repeatability, or response time was not included in the project materials. For this reason, MedCyBot is best presented as a proof-of-concept prototype rather than a fully validated

clinical system.



(a) Robot base controller application.



(b) SCARA controller application.

Fig. 5. Application and software images used during implementation and testing.

VII. LIMITATIONS

The current work has several limitations. First, the prototype was evaluated mainly at a qualitative level, and the project did not include a detailed numerical study of system performance. Second, some ideas discussed during the broader design stage, such as autonomous navigation and translation support, were not fully validated on the final prototype.

Another limitation is that the final prototype used a simplified controller setup, with an Arduino Uno used to test both

the base and upper-section functions. This was suitable for prototype development, but a more advanced version of the system may require a more distributed architecture and more detailed system-level testing.

These limitations do not reduce the value of the project as a student engineering prototype, but they do define the scope of the present paper. The main contribution of this work is the integration of multiple assistive functions into one prototype platform.

VIII. CONCLUSION

MedCyBot was developed as a prototype healthcare assistive robot that combines communication support and basic task assistance in a single platform. The system brought together speech-to-text, text-to-speech, multidirectional motion, and a SCARA-based handling mechanism. This combination makes the prototype relevant to healthcare environments where communication support and routine item handling are both important.

The main contribution of the work lies in showing that these functions can be integrated into one low-cost assistive platform. Although the prototype was not evaluated as a fully validated clinical system, it provides a useful proof of concept for future work in healthcare robotics.

IX. FUTURE SCOPE

Future work may extend MedCyBot in several directions. These include stronger language-translation support, more advanced navigation and obstacle-avoidance capability, sanitation-related features, air-quality sensing, and broader monitoring functions. With further development and testing, the present prototype could be expanded into a more capable assistive system for healthcare environments.

REFERENCES

- [1] A. S. Dhanjal and W. Singh, "An automatic machine translation system for multi-lingual speech to Indian sign language," *Multimedia Tools and Applications*, vol. 81, pp. 4283–4321, 2022.
- [2] R. Saini, "Speech recognition system (speech to text) (text to speech)," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 1, Jan. 2022.
- [3] P. Antoniadis, E. Tsardoulis, and A. Symeonidis, "A mechanism for personalized automatic speech recognition for less frequently spoken languages: The Greek case," *Multimedia Tools and Applications*, vol. 81, pp. 40635–40652, 2022.
- [4] S. Kulkarni, D. A. Torse, and D. Kulkarni, "A cloud based medical transcription using speech recognition technologies," *International Research Journal of Engineering and Technology*, vol. 7, no. 5, May 2020.
- [5] W. Chan, N. Jaitly, Q. V. Le, and O. Vinyals, "Listen, attend and spell," arXiv:1508.01211, 2015.
- [6] S. Jain, U. Agrawal, A. Kumar, A. Agrawal, and G. S. Yadav, "Simultaneous localization and mapping for autonomous robot navigation," in *2021 International Conference on Communication, Control and Information Sciences*, 2021.
- [7] B. Liu, Z. Guan, B. Li, G. Wen, and Y. Zhao, "Research on SLAM algorithm and navigation of mobile robot based on ROS," in *2021 IEEE International Conference on Mechatronics and Automation*, 2021.
- [8] A. Rosinol, M. Abate, Y. Chang, L. Carlone, and K. Carlone, "Kimera: An open-source library for real-time metric-semantic localization and mapping," in *IEEE International Conference on Robotics and Automation*, 2020.
- [9] R. Mur-Artal and J. D. Tardos, "Visual-inertial monocular SLAM with map reuse," arXiv:1610.05949v2, 2017.

- [10] E. Sandstrom, Y. Li, L. Van Gool, and M. R. Oswald, "Point-SLAM: Dense neural point cloud-based SLAM," arXiv:2304.04278v3, 2023.
- [11] M. Shariatee, A. Akbarzadeh, A. Mousavi, and S. Alimardani, "Design of an economical SCARA robot for industrial applications," in *2014 IEEE conference proceedings*, 2014.
- [12] X. Wang, L. Sun, K. Yu, B. Wang, and X. Li, "Research on SCARA robot fault diagnosis based on Hilbert-Huang transform and decision tree," in *2021 IEEE 9th International Conference on Computer Science and Network Technology*, 2021.
- [13] C. Urrea, J. Cortes, and J. Pascal, "Design, construction and control of a SCARA manipulator with 6 degrees of freedom," *Journal of Applied Research and Technology*, Dec. 2016.
- [14] M. Vaishnavi, H. R. Dhanush Datta, V. Vemuri, and L. Jahnvi, "Language translator application," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 7, July 2022.
- [15] N. Verma, A. Jain, A. Basak, and K. B. Saksena, "Survey and analysis on language translator using neural machine translation," *International Research Journal of Engineering and Technology*, vol. 5, no. 4, pp. 3720–3726, Apr. 2018.
- [16] H. Inaguma, K. Duh, T. Kawahara, and S. Watanabe, "Multilingual end-to-end speech translation," in *2019 IEEE conference proceedings*, 2019.
- [17] P. Shubha and M. Meenakshi, "Design and implementation of healthcare assistive robot," in *2019 5th International Conference on Advanced Computing and Communication Systems*, 2019.
- [18] P. Kadam, P. Padalkar, A. Mohite, S. Mirajgave, and S. Gudadhe, "An autonomous assistance robot for multi-purpose medical applications using ROS," in *2023 International Conference on Computer, Electrical and Communication Engineering*, 2023.
- [19] P. Jayawardhana, A. Aponso, N. Krishnarajah, and A. Rathnayake, "An intelligent approach of text-to-speech synthesizers for English and Sinhala languages," in *2019 IEEE 2nd International Conference on Information and Computer Technologies*, 2019.
- [20] R. Luo, X. Tan, R. Wang, T. Qin, J. Li, S. Zhao, E. Chen, and T.-Y. Liu, "Lightspeech: Lightweight and fast text to speech with neural architecture search," in *ICASSP 2021*, 2021.
- [21] R. K. Jaiswal and R. K. Dubey, "Concatenative text-to-speech synthesis system for communication recognition," in *Proceedings of the fifth international conference on electronics, communication and aerospace technology*, 2021.