HAND-MIMIC ROBOT: CNN-BASED SMART GESTURE CONTROLLED ROBOTIC ARM

***Project Reference No.:** 49S_BE_2265*

College	:	*Bangalore Technological Institute, Bengaluru*
Branch	:	*Department of Artificial Intelligence & Machine Learning*
Guide(s)	:	*Dr. M A Dorai Ranga Swamy*
Student(s)	:	*Ms. Eshwari HN*
		Mr. G Udhbav
		Ms. Lakshmi Channabasu Koujalagi
		Mr. Pruthvi Narayana Reddy V

Keywords:

Hand Gesture Recognition, Convolutional Neural Networks (CNN), Computer Vision, Human–Robot Interaction (HRI), Embedded Systems, Assistive Robotics.

Introduction:

Recent advancements in robotics and artificial intelligence have increased the need for intuitive and natural human–machine interaction systems. Conventional robotic control methods such as joysticks, switches, keyboards, and sensor-based gloves require physical contact and prior training, which limits accessibility and flexibility in many real-world environments. Hand gesture recognition offers a more natural and efficient alternative for controlling robotic systems without relying on external control devices.

The proposed project focuses on developing a CNN-based smart gesture-controlled robotic arm capable of mimicking human hand movements using computer vision techniques. The system captures real-time hand gestures through a camera, extracts important hand landmarks, and classifies them using a Convolutional Neural Network deployed on an embedded platform such as Raspberry Pi. The recognized gestures are then translated into servo motor commands that enable the robotic arm to reproduce the corresponding motion smoothly and accurately.

This contactless interaction system has significant applications in assistive technologies, industrial automation, rehabilitation support, and hazardous environment operations where direct human control is difficult or unsafe. By enabling robots to respond naturally to human gestures, the system contributes toward improving human–robot interaction and supports the development of intelligent and accessible robotic solutions.

Objectives:

1. Gesture Recognition: Develop CNN/LSTM models to identify ASL and dynamic gestures, helping hearing and speech-impaired individuals communicate easily with machines and assistive robots.
2. Robotic Arm Control: Create a robotic arm that performs grip, lift, rotate, and release actions to assist in elderly care, home automation, and industrial handling tasks.

3. Voice-Controlled Mobility: Enable wireless RC car movement using ESP32 and voice commands to support hands-free operation for people with mobility challenges or in hazardous areas.
4. Vision-Based Object Detection: Implement color and shape recognition for intelligent sorting and pick-and-place operations useful in waste segregation, manufacturing, and healthcare support.
5. Hybrid Multimodal System: Integrate gesture, voice, and vision-based control to build a versatile robotic system that enhances human–robot collaboration in daily life, hospitals, and emergency services.

Methodology:

The project follows a multimodal human–robot interaction approach that integrates gesture recognition, voice command control, and vision-based object detection with a robotic manipulation platform. The development begins with capturing real-time hand gesture input using a camera connected to an embedded processing unit such as Raspberry Pi. Hand landmarks and gesture features are extracted using computer vision techniques and processed through deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks trained on datasets like ASL and ChaLearn to classify both static and dynamic gestures accurately.

The recognized gestures are mapped to predefined robotic arm movements such as gripping, lifting, rotating, and releasing objects. Servo motors are controlled through PWM signals generated by the embedded controller to ensure smooth and precise motion replication. This enables intuitive contactless control of the robotic arm without the need for wearable sensors or physical controllers.

In addition to gesture control, a voice-command module is implemented using ESP32 to enable wireless hands-free operation. Speech input is processed using lightweight speech recognition libraries and translated into motion commands for robotic mobility control through an RC car platform. This feature supports users with limited mobility and allows remote operation in hazardous environments.

A vision-based object recognition module is also integrated using color and shape detection techniques through OpenCV. The camera identifies objects based on predefined visual characteristics and assists the robotic arm in sorting and pick-and-place operations useful in domestic automation, healthcare assistance, and waste segregation tasks.

Three modules—gesture recognition, voice control, and object detection—are combined into a unified multimodal control framework. The integrated system improves flexibility, safety, and usability by allowing multiple interaction methods depending on environmental conditions and user requirements. The final prototype is designed as a compact, cost-effective assistive robotic platform that demonstrates efficient human–robot collaboration for real-world applications.

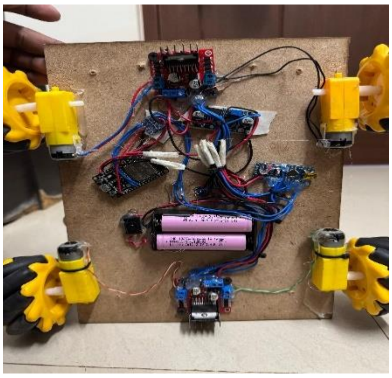 

Fig. Hardware components used for the project

Result & Conclusion:

In conclusion, the developed system demonstrates an effective multimodal human–robot interaction platform that integrates deep learning–based gesture recognition with voice-assisted robotic mobility control. The CNN/LSTM-based gesture recognition module successfully interprets predefined hand gestures in real time using camera input and translates them into corresponding robotic arm movements such as gripping, lifting, rotating, and releasing actions. The robotic arm responds smoothly to gesture inputs through coordinated servo motor control, enabling natural and contactless operation without the need for wearable sensors or external controllers.

The voice-command module implemented using ESP32 enables wireless control of the mobility platform, allowing hands-free navigation in environments where manual interaction is difficult or unsafe. Experimental testing shows reliable gesture classification performance under normal lighting conditions and stable actuation response from the robotic arm during continuous operation. The integration of gesture-based manipulation with voice-assisted mobility improves system flexibility and makes the platform suitable for assistive and semi-autonomous robotic applications.

The prototype validates the effectiveness of combining computer vision, embedded processing, and robotic actuation into a unified interaction system capable of translating human intent into mechanical motion in real time. The results highlight the system's potential for applications in assistive robotics, rehabilitation support, remote handling in hazardous environments, and intuitive human–robot collaboration. Overall, the project demonstrates a practical step toward developing intelligent robotic interfaces that respond naturally to human behaviour and support safer and more accessible interaction with machines.

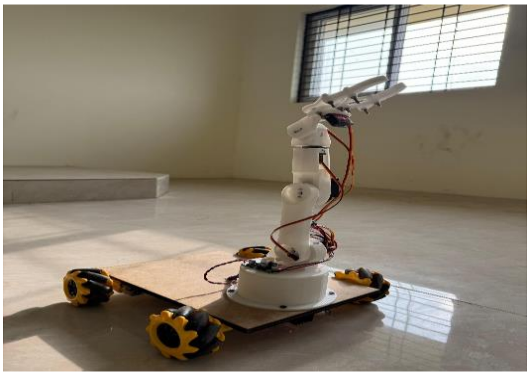 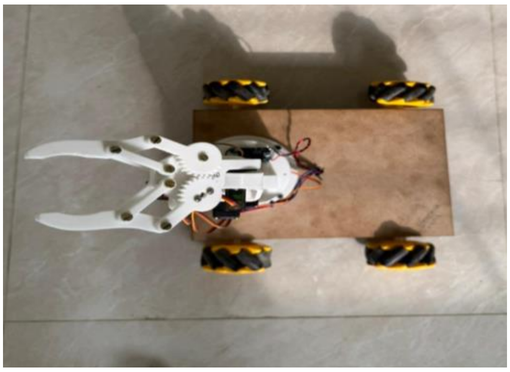

Fig. Fully working prototype.

References:

1. J. Kumar et al., “A Novel Voice Assisted Internet of Things Based Residential Automation Scheme with Learning Support,” IEEE, 2022.
2. B. Patle et al., “A Review on AI Vision Robotic Arm Using Raspberry Pi,” International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2022.
3. S. Mahadevkar et al., “A Review on Machine Learning Styles in Computer Vision — Techniques and Future Directions,” IEEE Access, 2022.
4. G. Li et al., “A System for Automatic Recognition of Robotic Hand Gesture Based on Raspberry Pi and Convolutional Neural Network by Using Specklegram,” IEEE Sensors Journal, vol. 25, no. 7, pp. 11839–11846, Apr. 1, 2025, doi: 10.1109/JSEN.2025.3537703.
5. Q. Gao et al., “An Efficient RGB-D Hand Gesture Detection Framework for Dexterous Robot Hand-Arm Teleoperation System,” IEEE Transactions on Human-Machine Systems, 2023.
6. H. Ghael, “A Review Paper on Raspberry Pi and Its Applications,” International Journal of Advanced Engineering and Management (IJAEM), 2020.
7. B. P. Kolla, “Content Extraction Studies Using Neural Network and Attribute Generation,” Indian Journal of Science and Technology, 2016.
8. A. Dahiya, D. Wadhwa, R. Katti, and L. G. Occhipinti, “Efficient Hand Gesture Recognition Using Artificial Intelligence and IMU-Based Wearable Device,” IEEE Sensors Letters, vol. 8, no. 12, pp. 1–4, Dec. 2024.
9. Shivani, S. Gaur, P. Khaneja, R. Sharma, S. Kaur, and M. Kaur, “Efficient Approach for Designing Gesture Controlled Robotic Arm,” International Journal of Control and Automation, 2015.
10. S. Verma, “Hand Gestures Remote Controlled Robotic Arm,” Advances in Electronic and Electric Engineering, 2013.
11. B. P. Kolla et al., “Information Extraction in Unstructured Multilingual Web Documents,” Indian Journal of Science and Technology, 2015.
12. N. Kumar et al., “IoT and NLP Based Voice Assistant for Visual Impaired People,” IEEE ICDSAAI, 2025. Hand Mimic Robot: CNN Based Smart Gesture Controlled Robotic Arm 2025-26 Department of AI&ML, BTI, Bangalore-35
13. M. Tang et al., “Mecanum Wheel AGV Trajectory Tracking Control Based on Efficient MPC Algorithm,” IEEE Access, 2024.
14. “Mechanical Design and Path Planning of a Robotic Arm for GIS Pipeline Inspection,” in Proc. IEEE Int. Conf. Mechatronics and Automation (ICMA), Beijing, China, 2020.
15. Q. Wu, “Research Control of a Smart Car Based on Chinese Voice Recognition Technology Through the Raspberry Pi 4B Platform,” IEEE ICPECA, 2023.
16. R. A. Khan et al., “Robotic Arm Controlled by Hand Gesture Using Leap Motion,” University of Sindh Journal of Information and Communication Technology, 2019.
17. H. Shetty, Gouthami B, Sanjana Y. G., Ganesh R. C., G. Kiruba R., and G. T. G. Tanuja, “ROLEX: AI-Driven Voice-Controlled Robotic Systems – Enhancing Human-Robot Interaction,” International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS), Mar. 2025.
18. S. C. Sethuraman, G. R. Tadkapally, A. Kiran, S. P. Mohanty, and A. Subramanian, “SimplyMime: A Dynamic Gesture Recognition and Authentication System for Smart

Remote Control," IEEE Sensors Journal, vol. 24, no. 24, pp. 42472–42483, Dec. 15, 2024, doi: 10.1109/JSEN.2024.3487070.
19. R. Khande and S. Rajapurkar, "Smart Voice and Gesture Controlled Wheel Chair," IEEE ICOEI, 2022.
20. A. S. et al., "Teach Pendant at Fingertips: Intuitive Vision-Based Gesture-Driven Control of Dexter ER2 Robotic Arm," IEEE Access, 2025.
21. Z. M. Yusoff et al., "Wireless Hand Motion Controlled Robotic Arm Using Flex Sensors," Indonesian Journal of Electrical Engineering and Computer Science, 2023.

Project Outcome & Industry Relevance:

The proposed project contributes to the advancement of intelligent human–robot interaction by developing a gesture-controlled robotic arm integrated with voice-assisted mobility control using embedded artificial intelligence techniques. The system demonstrates how camera-based gesture recognition combined with deep learning models can enable intuitive and contactless control of robotic manipulators without requiring wearable sensors or traditional control interfaces. This approach supports the development of accessible assistive robotic systems for individuals with limited mobility and improves safety in environments where direct human interaction with machines is difficult or hazardous.

The robotic arm developed in this project can be applied in healthcare support for assisting patients and elderly individuals in handling small objects and performing basic daily tasks. The voice-controlled mobility module enables remote navigation of robotic platforms in industrial inspection, warehouse monitoring, and hazardous-area operations such as chemical handling and disaster-response support. The project also contributes academically by integrating computer vision, embedded systems, and robotic actuation into a unified working prototype that reflects current trends in intelligent automation. Overall, the system demonstrates strong potential for adaptation in assistive robotics, smart automation platforms, and next-generation human–robot collaboration systems across healthcare and industrial sectors.

Working Model vs. Simulation/Study:

The project involves the development of a **physical working prototype** of a gesture-controlled robotic arm integrated with a voice-assisted mobility platform. The system is implemented using a camera-based gesture recognition module powered by CNN/LSTM models running on an embedded processor such as Raspberry Pi, along with servo motor actuation for real-time robotic arm movement. In addition, an ESP32-based voice command module enables wireless control of a mobile robotic platform for hands-free navigation.

The prototype demonstrates real-time gesture interpretation and corresponding robotic motion through hardware integration and embedded processing rather than simulation-only validation. Experimental testing was carried out to evaluate gesture recognition performance, servo response accuracy, and system stability under controlled conditions. Therefore, the project represents a practical implementation-focused assistive robotics system supported by both software intelligence and hardware-level execution.

Project Outcomes and Learnings:

The project resulted in the successful development of a gesture-controlled robotic arm integrated with voice-assisted mobility control using embedded artificial intelligence techniques. The system demonstrated real-time interpretation of hand gestures using CNN/LSTM-based models and their conversion into corresponding robotic arm movements through coordinated servo actuation. The voice-command module enabled wireless navigation control of the mobility platform, improving accessibility and operational flexibility in assistive and hazardous environments.

Through this project, significant learning was achieved in the areas of computer vision, deep learning model implementation, embedded system programming, and robotic actuator control. The team gained practical experience in integrating camera-based perception with real-time motor control using Raspberry Pi and ESP32 platforms. Additional learning outcomes included dataset handling, gesture mapping logic development, hardware interfacing using PWM signals, and debugging of integrated AI–robotic systems. Overall, the project strengthened interdisciplinary skills in intelligent automation and provided hands-on exposure to developing real-time human–robot interaction systems.

Future Scope:

The future scope of this project includes:

1. The system can be extended for assistive robotics to help elderly and physically challenged individuals perform daily tasks independently.
2. It can be adapted for rehabilitation support systems to assist patients during hand-movement therapy and recovery training.
3. The robotic arm can be deployed in hazardous environments such as chemical plants, fire-response areas, and nuclear inspection zones for safe remote handling.
4. The platform can be integrated into industrial automation systems for intuitive operator-controlled manipulation tasks.
5. It can be applied in smart healthcare environments for contactless assistance and patient support applications.
6. The system can be improved by increasing the degrees of freedom of the robotic arm for more natural and flexible motion.
7. Gesture recognition accuracy can be enhanced using larger datasets and advanced deep learning architectures.
8. The voice-control module can be upgraded to support multilingual and offline voice recognition for real-world deployment.
9. Future versions can integrate depth cameras or 3D vision systems for more robust gesture detection.
10. The system can evolve into a fully multimodal human–robot interaction platform for next-generation collaborative robotics applications.