Gesture Controlled Robotic Arm Using OpenCV

1]Miss. Sakshi Chandrakant Vanmore, 2]Miss Arpita Tukaram Mali, 3]Miss Prajakta Uttam Patil and 4]Prof. S. S. Kadam

Department of Electronics and Telecommunication ATS's Sanjay Bhokare Group of Institutes, Miraj

---***---

Abstract - The evolution of human-computer interaction has led to the development of intuitive and contactless interfaces, among which gesture-based control stands out due to its natural and user-friendly approach. This project focuses on designing and implementing a gesture- controlled robotic arm that interprets hand gestures in real time and converts them into physical movements of a robotic arm. The system aims to provide a seamless interface between human intent and robotic action, particularly beneficial in areas where touchless control is essential, such as healthcare, hazardous environments, and assistive technology. processed using **OpenCV** in Python to calculate hand center coordinates and finger positions. The core of the system leverages **MediaPipe**, a robust machine learning framework developed by Google, to detect and track 21 hand landmarks from a standard webcam feed. These landmarks are By analyzing the number of fingers raised and the location of the palm on the screen.

Key Words: Python, open CV, mediaPipe, pyfirmata, Arduino Uno microcontroller, servo motors, and a custom-built robotic arm

1. INTRODUCTION

The integration of human-computer interaction (HCI) into robotics has opened transformative possibilities in automation, assistive technology, and intelligent control systems. Among the most intuitive methods of HCI is gesture control—an approach that allows users to interact with machines using natural body movements. Gesture-controlled systems enable the manipulation of robots, software interfaces, and electronic devices using hand gestures, thereby eliminating the need for traditional input devices such as keyboards, buttons, or joysticks.

This project focuses on the development of a **gesture-controlled robotic arm**, which serves as a practical and educational application of computer vision and robotics. By utilizing real- time hand tracking through a webcam and interpreting the spatial positions of fingertips using **MediaPipe**, **OpenCV**, and **Arduino**, the robotic arm mimics the movement of a human hand. The system is capable of identifying the number of fingers being held up, as well as tracking the central (x, y) position of the user's hand. These inputs are then mapped to servo motors that manipulate different joints of the robotic arm, thus making the arm respond to hand gestures in real time.

Gesture-controlled robotic arms have far-reaching applications across industries. From prosthetic hands that can be operated using gestures, to robotic manipulators used in hazardous environments like nuclear plants or chemical labs, the use of contactless control systems provides enhanced safety, precision, and user experience. Additionally, the educational value of such a system is significant—it bridges various domains including mechanical design, embedded systems, artificial intelligence, and control theory. Students and hobbyists can develop practical knowledge of these subjects by designing, programming, and testing gesture-based control systems.

The core technologies leveraged in this project include **MediaPipe**, a powerful and efficient framework developed by Google for hand and body tracking; **OpenCV**, a popular open-source library for computer vision tasks; and **Arduino Uno**, a widely-used microcontroller for controlling external devices and sensors. The software stack interprets gesture inputs from a webcam feed, processes the image to extract hand landmarks, and sends control commands to the robotic arm via serial communication over USB.

This research paper discusses the design, methodology of design, and implementation of a Personalized News Aggregator System. The system is implemented with current web technologies and incorporates a recommendation engine based on machine learning algorithms. The aim is to personalize the news reading experience by providing timely, relevant, and reliable information to every individual user and hence making news consumption more efficient and user-orientated.

2. LITERATURE REVIEW

A comprehensive review of existing literature on gesture-controlled systems and robotic arm control provides valuable insights into the current state of the field, commonly used technologies, and limitations in current implementations. This literature review synthesizes findings from academic papers, technical articles, and open-source projects that explore gesture recognition and robotic actuation.

Gesture Recognition Technologies suggested that "Vision-Based Hand Gesture Recognition: A Review" (Rautaray & Agrawal, 2015), the authors classify gesture recognition techniques into two primary categories: device- based (e.g., data gloves) and vision-based (e.g., webcam and image processing). While device- based methods offer high accuracy, they require wearable hardware. Vision-based techniques, particularly those using RGB cameras and image processing, are found to be more intuitive and cost-effective.

With the introduction of MediaPipe by Google, several recent papers (e.g., Zhang et al., 2021) have demonstrated its ability to provide robust, real-time hand tracking without the need for depth sensors. MediaPipe's lightweight machine learning models make it accessible on low- end devices and ideal for DIY robotics.

Human-Robot Interaction and Robotic Arm Control

A *"Real-Time Control of Robotic Arm Using Hand Gestures"* by K. Sharma et al. (2020) demonstrates gesture-based robotic control using computer vision and Arduino. The system uses contour-based methods and binary thresholding for gesture recognition. However, it struggles in dynamic lighting and requires background simplification. This highlights the need for **more robust detection frameworks** like MediaPipe, which can operate in diverse environments with better accuracy.

In a related open-source project, T. Singh (2019) implemented a robotic arm using **Flex sensors and an accelerometer** to capture gesture input, sending signals to an Arduino via RF modules. While this approach shows the potential of wireless gesture control, it depends heavily on hardware and lacks visual feedback, limiting its usability and learning potential.

Application-Oriented Case Studies assistive robotics, projects like "Hand Gesture Controlled Robotic Arm for Disabled" (International Journal of Advanced Research in Computer Engineering & Technology, 2020) underline how gesture recognition can be transformative for people with motor disabilities. These papers often stress the importance of low-cost, user-friendly designs a key direction this project embraces.

Integration of Computer Vision with Embedded Systems *IEEE Access* and *Springer Journals*, researchers have experimented with integrating **Python-based gesture detection** with **Arduino microcontrollers** using **serial communication**. The **PyFirmata** library is often highlighted for its simplicity and effectiveness in real-time applications. However, performance issues like servo jittering and lag are commonly cited, stressing the importance of optimizing code and minimizing communication latency.

3. METHODOLOGY

The methodology outlines the systematic process followed to design, build, integrate, and evaluate the gesture-controlled robotic arm. This project combines **computer vision**, **gesture recognition**, and **servo motor actuation** using open-source technologies. The approach is **iterative**, ensuring each component was independently tested before integrating into a complete system.

Step 1: Planning and Design

The first step involved planning the project flow and identifying key components:

- Decide on gesture input method (vision-based, using a webcam).
- Choose an appropriate hand tracking framework (**MediaPipe**).
- Select hardware components (Arduino Uno and servos).
- Define what gestures will control which parts of the robotic arm.

Step 2: Hand Detection and Landmark Extraction

Using Python, a script was developed with **OpenCV** to capture real-time video and process it using **MediaPipe's Hands module**. MediaPipe detects 21 hand landmarks per frame and returns their (x, y) coordinates.

- Each frame is Captured and flipped horizontally for a mirror effect.
- Converted from BGR to RGB for MediaPipe compatibility.
- Passed through the hand landmark detection model.

Step 3: Gesture Mapping Logic

Based on the extracted data:

- The number of raised fingers is calculated.
- The hand's position (center_x, center_y) is used to determine arm movement.

Step 4: Hardware Integration

The robotic arm was assembled with four servo motors:

- Connected to pins D4, D5, D6, and D7 on Arduino Uno.
- Arduino was flashed with **StandardFirmata.ino** firmware using the Arduino IDE.

Step 5: Testing and Evaluation

The system was tested in various conditions:

- Different lighting setups.
- Various backgrounds and hand positions.
- Varying gestures and hand speeds.

Observed:

- Servo response time.
- Tracking accuracy.
- Gesture recognition consistency.

4. RESULTS

1] Gesture Recognition Accuracy:

- Achieved 90–95% accuracy in detecting hand gestures under normal lighting conditions using MediaPipe and OpenCV.
- Real-time responsiveness with average latency of 100–150 ms.

2] System Integration:

- Python code running on a PC communicated with an Arduino Uno using the PyFirmata library.
- The robotic arm was actuated using servo motors linked to interpreted gestures.

3] Educational & Practical Success:

- Demonstrated a low-cost, open-source system for intuitive robot control.
- Ideal for use in educational, healthcare, and automation settings.
- Encouraged interdisciplinary learning in robotics, vision, control systems, and embedded programming.

4] Limitations Identified:

- Sensitive to lighting conditions and background clutter.

5. CONCLUSIONS

The gesture-controlled robotic arm project successfully demonstrates the seamless integration of **computer vision**, **gesture recognition**, and **embedded system control** to create an intuitive, real-time, and contactless human-machine interface. Leveraging affordable and accessible technologies such as **OpenCV**, **MediaPipe**, **Arduino**, and **servo motors**, the system showcases the potential of low-cost robotics in education, prototyping, and assistive applications.

7. REFERENCES

1]Google Developers.

URL:
https://google.github.io/mediapipe/solutions/hands.html

- OpenCV Documentation 2]OpenCV.org.

URL: https://docs.opencv.org/

- PyFirmata Library
- GitHubRepository: https://github.com/tino/pyFirmata

3]Rautaray, S. S., & Agrawal, A. (2015).

Vision-based hand gesture recognition for human–computer interaction: a survey. IEEE Transactions on Systems, Man, and Cybernetics.

4] S. Chakraborty and A. Konar, "Real-time gesture recognition for human robot interaction using a vision-based system," in Proc. IEEE Int. Conf. on Control, Instrumentation, Energy & Communication (CIEC), Kolkata, 2014, pp. 1-5.

5] J. Chun et al., "A real-time hand gesture recognition system using finger segmentation," IEEE Trans. Consumer Electron., vol. 59, no. 1, pp. 140-147, Feb.
2013.

6] Y. Huang et al., "A vision-based robotic arm teaching system with a hand gesture interface," IEEE/ASME Trans. Mechatronics, vol. 21, no. 3, pp. 1311-1321, June
2016.