International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

# AI-Driven Gesture Recognition for Robotic Control

[1]D. N. V. S. Vijaya Lakshmi, [2]Dr Prasad Rayi, [3]V. Rambabu, [4]Ch S K Chaitanya

**Submitted:** 10/05/2023 **Revised:** 12/07/2023 **Accepted:**01/08/2023

**Abstract:** In our planet, several individuals use gestures, a really effective communication tool among people. Actions are more instinctive than verbal expressions. Gestural communication is an effective modality. Wireless communication systems are extensively used and implemented in many industrial settings and residential areas. In our gesture control project, a two-wheeled robot is operated by diverse hand gestures. Gesture recognition employs several image processing techniques to identify picture signals. Robots may assume many forms, but some replicate the look of living organisms. These robots engage with people directly, making the interface user-friendly. The first gadgets were primarily used for navigation and robotic control without any organic interface. To address this demand, we have initiated a project that involves sending orders to the robot using hand gestures. This image processing technique enables us to manipulate the robot with our fingertips. We used image processing techniques to collect these directives. Consequently, the robot need to navigate in the specified direction.

***Keywords:*** *Gesture Recognition, Image Processing, Opencv, Human-Computer Interaction (HIC), Python, Machine Learning, Wireless Communication.*

## INTRODUCTION

In the contemporary era, the robotics business has been advancing several technologies to enhance the efficiency, accessibility, and precision of its systems. Fundamental duties may include roles that are detrimental to human well-being, monotonous occupations that induce tedium, or positions that are stressful, among others. Although robots may serve as substitutes for people, they still need human oversight. Wireless communication with robots offers several benefits. However, it also has several drawbacks. A contemporary approach of gesture control has gained significant popularity, surpassing the traditional means of managing robotic systems using physical instruments. Hand gestures provide an alternate method of control, facilitating a more efficient and intuitive interface with the robotic system. Utilizing image processing, diverse machine learning frameworks, and suitable hardware for the issue of gesture control. Numerous systems have been developed in the same domain with diverse methodologies.

We suggested a method that employs a serial wire for the navigation of a robot with hand gesture signals. This technology enables the controller to maneuver a robot in four directions: forward, backward, left, and right as necessary. This technology enables a user to control a robot using an integrated or external camera or webcam connected to a computer, laptop, or any device capable of receiving or interpreting orders. Hand gesture instructions are executed using the fingers. Initially, the visual frame is received as input and then undergoes additional processing. The processed input picture is then used to extract instructions. Upon receiving the gesture command via the camera, a signal is created to transmit the order to the robot. In the subsequent stage, the processed command data is now prepared for upload into the hardware system. The robot now accesses signals via a serial communication line linked to the TX and RX (serial ports) of the Arduino board and a control station. Arduino receives input signals from the device executing the source code over a serial connection and produces output that is sent to the DC gear motors via a motor driver shield. Four distinct input movements produce varying output signals. The motor driver receives digital signals as input and provides output to the DC motors.

[1,2,3,4] *International School Of Technology And Sciences For Women, A.P, India.*

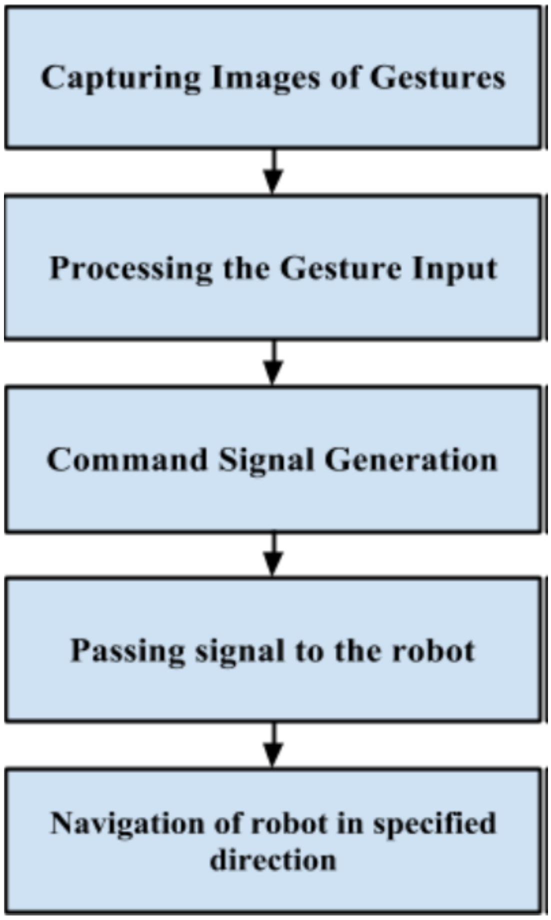



**Figure 1: The basic flow of the system**

**Literature Survey**

A fundamental mechanical arm in [1] was designed to be mostly programmable and has functionalities akin to those of a human arm. The study in [2] detailed the development of a wireless mobile robotic arm capable of performing pick-and-place operations, operated by a wireless PS2 controller. The Arduino Mega platform served as the basis for development, and evaluations of its speed, range, and lifting capacity were undertaken to ascertain its potential.

The objective outlined in [3] was to create a robot capable of recognizing human interaction and doing designated duties. Develop a wearable glove equipped with sensors to detect hand motions and transform the raw mechanical data into electrical signals. This information will undergo additional processing and be formatted for comprehension by the lily pad mounted on the glove. The purpose of this lily pad is to function as a data transmitter for wireless communication. The sent data was analyzed and retransmitted to the microcontroller, which was received by the receiver module. A humanoid robot was developed in [4]. The robot in [5] was an improved version developed using Raspberry Pi and OpenCV, used for item identification based on color, size, and form. A robotic arm using IoT, Bluetooth, and Wi-Fi was deployed in [6].

This article does a research on gesture control. Robotic arms were designed and used for medical surgery in [7,8]. The robotic arm described in [9] used the ZigBee protocol. The literature analysis indicates that the current module use sensors for regulating robotic motions, which are cumbersome and difficult to manage. The utilization of Raspberry Pi is expandable, although it is expensive and somewhat sophisticated in structure. Research using CNN and ANN via ML algorithms is intricate and challenging to comprehend, necessitating a strong understanding of machine learning. Conventional robots operate by remote control, which is ineffective for contemporary industry.

This study presents the design of a robotic arm controlled by hand gestures using artificial intelligence, therefore providing a real-time experience for the user. The design of the suggested robotic arm is shown in Figure 1. It efficiently employs virtual controllers and mitigates distance constraints via the use of a Bluetooth module. The suggested module can accurately identify the location of the right or left hand and fingers, providing real-time information to the user on the screen.

The MediaPipe package in Python is used for hand gesture recognition. Media Pipe Hands is an advanced solution for precise hand and finger tracking. It utilizes machine learning (ML) to deduce 21 three-dimensional landmarks of a hand from a single picture. While current state-of-the-art methods mostly depend on robust desktop settings for inference, the suggested methodology attains real-time performance on a mobile device and is capable of scaling to several hands. We anticipate that offering this hand perception capability to the broader research and development community will foster innovative use cases, inspiring new applications and research directions. Robot vision operates by incorporating one or many cameras into the robotic system.

A camera is affixed to the controller side, serving as the machine's "eye." At the hardware level, Bluetooth receives signals from the software side via a communication channel. These signals are sent to the Arduino, which governs all hardware operations, including the control of servo motors coupled to robotic arms. The servo motors are positioned at 90 degrees to control the fingers. Figure 2 illustrates 21 landmarks used for monitoring human hand movement.

***Dataset Preparation***

The hand gesture collection consists of American Sign Language (ASL) movements, seen in Fig. 2 [23], totaling 700 photos from five people. These photographs exhibit differences in lighting conditions and hand positions, accomplished via the use of image processing tools. The hand photos have a single channel and are 400 by 400 pixels, with the hands positioned center. To accelerate calculations and improve the efficiency of hand motion identification by the system, the pictures have been reduced to dimensions of (64, 64, 1). American Sign Language (ASL) utilizes a series of hand motions that represent 10 numerals, from 0 to 9. Each class has 700 photos, and the dataset is divided into two subsets: one for training (80% of the data) and the other for testing (20% of the data). Moreover, the quantity of photos per category is uniformly allocated between both groups.

## METHODOLOGY

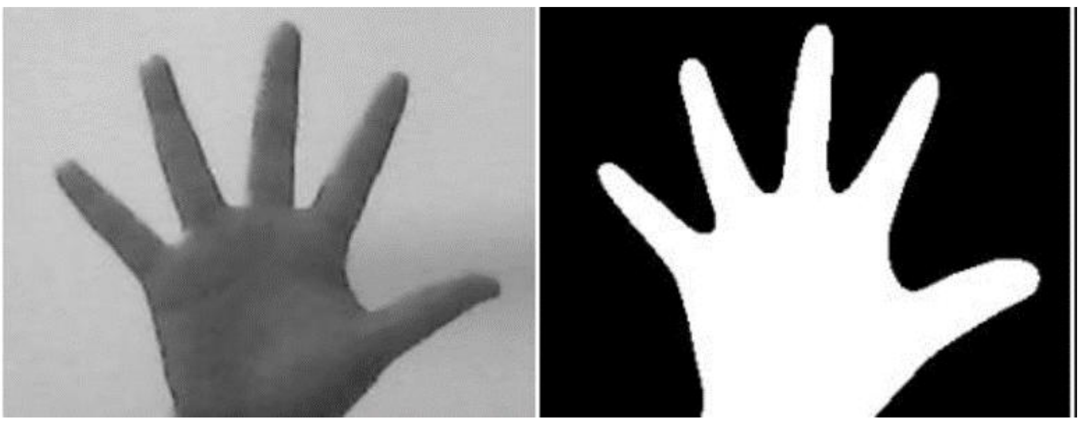
The hand gesture recognition project is executed using OpenCV, an effective computer vision library, together with the Python programming language and its associated libraries, including NumPy, Serial, and Time. Python is a high-level programming language that generates clear and comprehensible system code. The Python library used in this context is Numpy. The webcam is used to collect real-time photos that may be processed to extract necessary data from the backdrop.

### A. Using Hand Segmentation Method

This approach posits that segmentation may transform a grayscale picture into a binary image, contingent upon the presence of certain objects, namely a hand and the backdrop. This technique is used for segmentation, converting grayscale photos into binary representations that represent either the hand or the backdrop. An effective segmentation is essential to determine an appropriate grey level threshold for isolating the hand from the backdrop; there should be no overlap between the hand and the background. The segmentation method is contingent upon the kind of pictures and the corresponding application domains.

### B. Using Haar-Cascade Classifier

To enable the program or source code to perpetually detect the hand, regardless of whether it is stationary or in motion, the code must accurately identify the hand picture. This is where machine learning may be applicable. In this study, the machine learning approach was used to track hand gestures prior to the completion of picture processing. Sander [14] recommends the cascade classifier as the suitable machine learning approach for image processing applications. Consistent with Viola & Jones [15], the cascaded classifiers serve as the program characteristics enabling the robot to interpret the hand picture, which will be classified in each frame of images. We used the Haar-cascade method in this project. Viola and Jones [15] proposed that the Haar-cascade classifier serves as the standard classifier for machine learning algorithms to recognize hand pictures.



## RESULT AND DISCUSSION

The image in fig-10 shows the results of our project where a 2 wheel drive robot is controlled through image processing wirelessly.

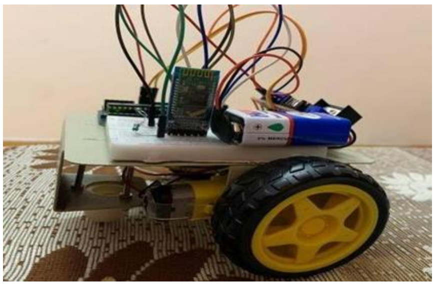


**Figure 3: Image of the wireless control of robot through image processing**

Figure 4 illustrates the binary picture of the user's hand, where contours are often identified. Here, in accordance with Otsu's Binarization technique. The inherent functions of OpenCV and Python automatically determine the threshold of the picture from its histogram. The resultant picture is shown in Figure 4.

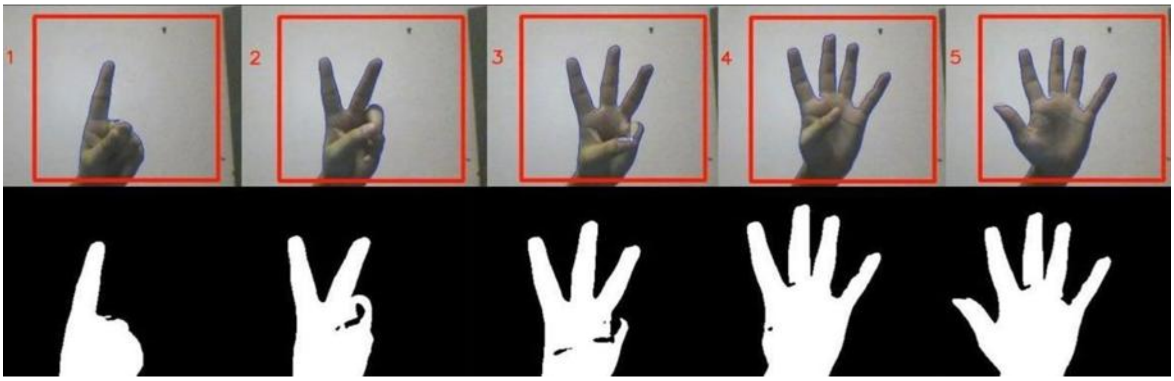



**Figure 4: Five Hand Gestures**

This research identified the number of fingers shown in the illustration. Our preliminary method for developing a gesture recognition system included the technique of backdrop removal. Numerous challenges and accuracy concerns were encountered throughout the implementation of recognition. A system using background subtraction. Background subtraction cannot guarantee stability during abrupt, significant illumination changes, resulting in many discrepancies. The gesture recognition system demonstrated robustness and high accuracy when used against a simple backdrop. This precision was maintained irrespective of the backdrop color, as long as it is a uniform, solid color devoid of any irregularities. In instances when the backdrop was not uniform, the items present were identified as discrepancies in the picture capture process, leading to erroneous results. It is advisable to use this approach against a clean backdrop to optimize results and ensure accuracy.

### Applications of Gesture Recognition in Robotics:

Gesture recognition applications are extensive, including virtual reality, safe driving, healthcare, and gadget control. This section focuses on vision-based hand movements captured by RGB-D and monocular cameras for application in human-robot interaction. This is a list of the main applications of gesture recognition technology. Device control: Gestures may also be used to run intelligent robots. As artificial intelligence progresses, millions of households will progressively adopt home robots or smart appliances, with gesture control becoming more intuitive for consumers than traditional button or touchscreen interfaces. Uses is a firm that develops hardware and software enabling Smart TVs to recognize gestures and finger motions. Gestoo's artificial intelligence platform enables contactless control of music and lighting systems using gesture recognition technology. A single gesture may activate many communication channels using Gestoo, enabling the creation and assignment of gestures via a smartphone or other device.

Healthcare: Personnel and equipment may generate significant noise in tumultuous emergency and operating rooms. Voice instructions are less effective in this context than hand gestures. The clear contrasts between sterile and nonsterile settings exclude the use of touchscreens as well. Microsoft has shown the use of gesture recognition technology for viewing photographs and information during surgical or other procedural contexts. Doctors may use Gesture, a gesture control solution for medical equipment, to examine MRI, CT, and other images using simple gestures instead of scrubbing. By reducing the duration of physical contact between nurses and physicians and patients, this touch-free interaction reduces the risk of cross-contamination.

Understanding of sign language: Sign language serves as the primary communication medium for those with hearing impairments; yet, it may pose difficulties for untrained individuals to understand it. Sign recognition technology for sign language comprehension will significantly enhance the communicative abilities of the deaf and others. Gesture recognition in virtual reality enhances user immersion and experience by facilitating more intuitive interactions and control over virtual environments. In 2016, Leap Motion showcased enhanced gesture detection software that allowed users to track movements in virtual reality while interacting with computers. Mano Motion's hand-tracking software utilizes a smartphone camera (on Android and iOS) to detect 3D motions, applicable in augmented reality and virtual reality environments. This technique is applicable in robotics, consumer electronics, gaming, and Internet of Things devices.

**Challenges and Future Directions:**

Deep learning serves as an effective catalyst for gesture detection, and as artificial intelligence progresses, these systems will become more accurate and dependable. Future gesture recognition technologies will exhibit increased diversity and utility across a broader spectrum of sectors, including healthcare, education, and entertainment, therefore enhancing user convenience and fostering creativity. As deep learning and artificial intelligence advance, the domain of gesture recognition is set to attain enhanced intelligence. Through training, models may proficiently recognize complex actions with minimum user input, leading to enhanced and intuitive gesture recognition skills. In the future, sophisticated gesture recognition technology will be capable of analyzing and interpreting extensive volumes of motions in real time. This innovation will provide new opportunities for the use of gesture recognition across several domains, including virtual reality, gaming, and medical. As the applications of gesture recognition proliferate, the dependability of this technology becomes more crucial. Extensive testing and validation will be crucial for ensuring consistent and reliable performance of future gesture recognition systems across various situations. Improvements in computer vision and sensor technologies will enhance the precision of gesture detection. Enhanced-resolution cameras and more sensitive sensors may capture finer hand motions, hence augmenting the accuracy of gesture identification. Future gesture recognition systems will be more tailored and capable of adapting to the diverse gesture habits and preferences of users. Users may have the capability to program certain gestures to execute particular tasks or functions.

## CONCLUSION

The Gesture-Controlled Robot using Machine Learning provides an alternative method for robot control. It is a method of manipulating hardware components that enhances the robot's intuitiveness, efficiency, and ease of use. An novel and alternative approach has been developed using gesture-controlled robotic mechanics. A more intuitive and natural manner of control is attained using a gesture control system, enabling efficient and engaging robot operation. We have used a particular technique for inputting motions known as the finger count approach. These instructions are used for a robot, based on finger counts, to facilitate movement in specified directions. This system utilizes a predefined threshold setting to convert the incoming picture frame into a binary format. This method necessitates imposing certain boundaries and constraints on the backdrop, hence requiring a black background. This issue may be resolved using color-based thresholding. In the future, we will strive to enhance accuracy and include more gestures to facilitate more functions.

## REFERENCES


[1] Howard, Andrew & Zhu, Menglong & Chen, Bo & Kalenichenko, Dmitry & Wang, Weijun & Weyand, Tobias & Andreetto, Marco & Adam, Hartwig. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

[2] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li (2009) IEEE Conference on Computer Vision and Pattern Recognition. ImageNet: A large-scale hierarchical image database

[3] Raheja, J. L., Shyam, R., Kumar, U., and Prasad, P. B., "Real-Time Robotic Hand Control using Hand Gestures", Second International Conference on Machine Learning and Computing, 2010.

[4] P. B. Nayana and S. Kubakaddi 2014 Implementation of Hand Gesture Recognition Technique for HCI Using Opencv (International Journal of Recent Dev) vol. 2 no. 5 pp 17–21

[5] Ahuja, M. K., & Singh, A. (2015). Static vision-based Hand Gesture recognition using principal component analysis. Paper presented at the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE).

[6] Bretzner, L., Laptev, I., & Lindeberg, T. (2002). Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. Paper presented at the Proceedings of Fifth IEEE international conference on automatic face gesture recognition.

[7] Chen, F.-S., Fu, C.-M., & Huang, C.-L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. Image and vision computing, 21(8), 745-758.

[8] Dipietro, L., Sabatini, A. M., & Dario, P. (2008). A Survey of Glove-Based Systems and Their Applications. Ieee transactions on systems, man, and cybernetics, part c (applications and reviews), 38(4), 461- 482.

[9] Dong, G., Yan, Y., & Xie, M. (1998). Vision-based hand gesture recognition for human-vehicle interaction. Paper presented at the Proc. of the International Conference on Control, Automation and Computer Vision.

[10] Garg, P., Aggarwal, N., & Sofat, S. (2009). Vision-based hand gesture recognition. World Academy of science, engineering and technology, 49(1), 972-977.

[11] Gupta, S., Jaafar, J., & Ahmad, W. F. W. (2012). Static hand gesture recognition using local Gabor filter. Procedia Engineering, 41, 827-832.

[12] OpenCV Library http://docs.opencv.org/

[13] Arduino http://arduino.cc/en/Guide/HomePage

[14] S. Soo 2014 Object detection using Haar-cascade Classifier (Inst. Comput. Sci. Univ. Tartu) vol. 2 no. 3 pp 1–12

[15] P. Viola and M. Jones 2001 Rapid object detection using a boosted cascade of simple features (IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition. CVPR 2001) vol. 1 pp 1511- 151.