

A Comprehensive Survey on IoT-Enabled and Hand Gesture Controlled Robotic Arm Using Blynk IoT and OpenCV

Yogesh Bhangale¹, Ruchira Pagar¹, Bhaven Bankar^{1,*},
Tushar Kafare²

Abstract

This venture presents an IoT-enabled and hand gesture-managed robot arm that operates in three distinct modes: automated control, IoT-based control via the Blynk app, and gesture-based control using OpenCV. The gadget integrates a NodeMCU microcontroller for wireless verbal exchange and management, with MQTT protocol enabling real-time messaging among the devices. In automated mode, the robotic arm plays predefined tasks autonomously. In IoT mode, users can remotely manage the arm using the Blynk IoT platform, allowing for real-time manipulation over the net. In gesture mode, the pc's camera captures hand gestures, which can be processed by OpenCV to come across precise motions, such as swiping or the hand, and sends control instructions to the robotic arm through the MQTT protocol. The venture combines pc vision, wireless communication, and robotics, making it distinctly versatile for applications in automation, assistive generation, and clever environments. It demonstrates the sensible integration of IoT with robotics, allowing far off control, gesture-based interaction, and autonomous operation in a cost-effective and scalable solution. The capabilities and adaptability of IoT-enabled robotic arms can be greatly increased by using machine learning (ML). These systems can learn from data, adjust to changing conditions, and make wiser judgments by integrating machine learning algorithms. For instance, ML can be applied to enhance gesture detection, allowing for more precise and user-friendly robotic arm control. The arm can carry out activities more effectively and efficiently if ML is employed to optimize task preparation and execution. Additionally, ML can be used to create predictive maintenance models, which can assist minimize downtime and detect possible problems before they arise.

Keywords: IoT enabled robotic arm, gesture recognition, automation, controlled robotic, machine learning

*Author for Correspondence

Bhaven Bankar
E-mail: bhavenbankar.scoe.entc@gmail.com

¹Student, Electronics and Telecommunication, Sinhgad College of Engineering Vadgaon (BK), Pune, Maharashtra, India

²Assistant Professor, Electronics and Telecommunication, Sinhgad College of Engineering Vadgaon (BK), Pune, Maharashtra, India

Received Date: May 05, 2025

Accepted Date: August 04, 2025

Published Date: September 19, 2025

Citation: Yogesh Bhangale, Ruchira Pagar, Bhaven Bankar, Tushar Kafare. A Comprehensive Survey on IoT-Enabled and Hand Gesture Controlled Robotic Arm Using Blynk IoT and OpenCV. Journal of Advancements in Robotics. 2025; 12(3): 35–40p.

INTRODUCTION

With the massive influx of computers and digital devices in society, Human-Computer Interaction (HCI) has become an increasingly critical aspect of our daily lives. As computing, communication, and display technologies continue to advance, the methods used for human-computer interaction must evolve to keep pace with the expanding flow of information. It is widely believed that traditional input methods, such as remote controls and keyboard-based interactions, may soon become limitations in the efficient and effective utilization of modern computing devices. In response, hand gesture recognition has emerged as a promising alternative. This method offers a natural, intuitive

system for interacting with technology, eliminating the need to learn complex control systems and providing an easier, more seamless experience for users. As a form of human-computer interaction, hand gesture recognition can replace standard remote control or keyboard input, offering a more intuitive, gesture-based interface that mimics natural human motion [1].

However, the use of gestures for interaction is not without its challenges. The variability and ambiguity of gestures present significant hurdles, with concepts and gestures often being one-to-many or vice versa. Gestures can be categorized into two primary types: static and dynamic. Static gestures refer to pre-defined stances or poses, while dynamic gestures involve movement over time, and can be broken down into different phases, such as pre-stroke, stroke, and post-stroke. Furthermore, the significance of a dynamic gesture can be better understood by considering its geographic data (the origin of the gesture), pattern data (the design or sign represented by the gesture), and the range or space in which the gesture occurs [2].

The convergence of Internet of Things (IoT) and robotics has introduced a transformative era of intelligent automation, with IoT-enabled robotic arms at the forefront of this revolution. These robotic arms have garnered significant attention due to their potential to enhance operational efficiency across diverse industries, offering both remote control and monitoring capabilities. By seamlessly integrating wireless communication technologies with robotic systems, IoT-enabled robotic arms are able to operate autonomously or be controlled remotely through devices like smartphones or computers, thereby increasing flexibility and adaptability in various applications. One of the critical advancements driving this transformation is the incorporation of hand gesture recognition, which allows for more intuitive and natural control of robotic arms. The integration of gesture-based control enhances the overall user experience by making interactions more fluid and natural, removing the need for complex interfaces, and offering a more direct, humancentric way to interact with robotic systems [3].

In addition to this, the Kinect-based system for sign language recognition further demonstrates the potential of gesture-based interactions. This system leverages depth-sensing technology to track hand movements and recognize sign language gestures in real-time, improving accessibility for the hearing impaired. The system's use of machine learning and support vector machines (SVMs) for accurate gesture classification underlines the value of computer vision in enabling intuitive communication interfaces. However, challenges remain in addressing the nuances of gesture recognition, particularly in distinguishing between subtle finger movements or variations in facial expressions that can alter the meaning of a sign language gesture [4].

Moreover, the model-based hand tracking system presents a highly sophisticated approach to gesture recognition, using multiple cues like edges, optical flow, and shading to refine tracking and enhance its accuracy, particularly in dynamic environments where hand movements are subject to occlusions or background interference. This system's ability to adapt in real-time and improve tracking accuracy has important implications for applications like robotic control, where precise and robust gesture tracking is crucial for seamless interaction. The multi-cue tracking methodology offers a reliable solution to some of the most challenging aspects of gesture recognition, ensuring that the system remains responsive to user input even under challenging conditions [5].

The integration of gesture recognition with IoT-enabled robotic systems holds tremendous potential to revolutionize the way humans interact with machines. From sign language translation to robotic arm control, the ability to interpret natural hand movements offers a more intuitive, efficient, and adaptive interface for a variety of applications. As technologies like Kinect, machine learning, and multi-cue hand tracking continue to advance, we can expect the continued growth of gesture-based human-computer interaction, making our interactions with digital systems even more seamless and natural (Figure 1).

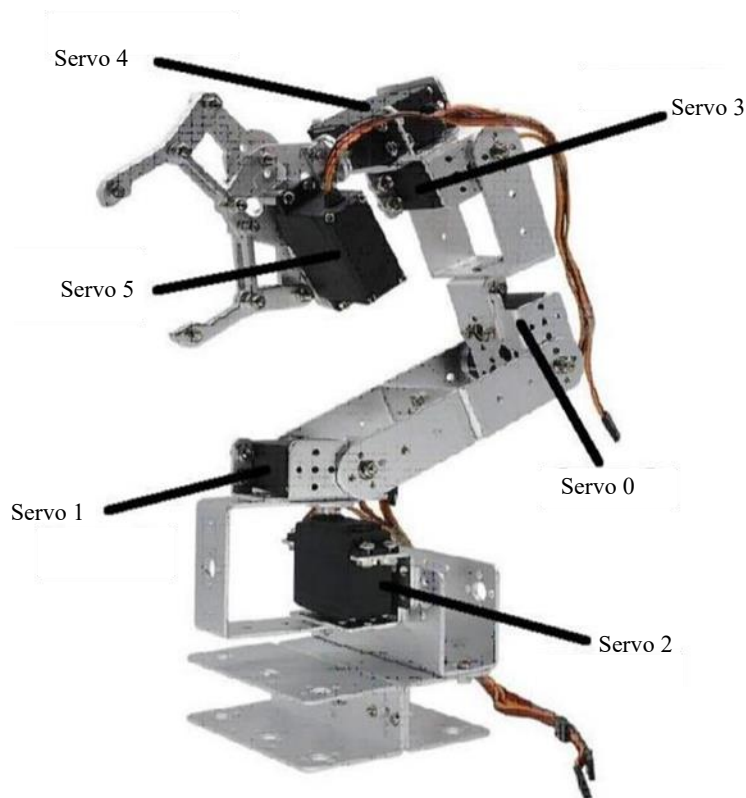


Figure 1. Robotic arm.
Source: roboticsandenergy.com.

LITERATURE SURVEY

The paper by Paterson and Aldabbagh gives an overview of a project focused on controlling a robotic arm using hand and finger movements, proposing a shift from traditional sensor-based gloves to a vision-based motion tracking method. By employing machine learning and computer vision, the approach identifies real-time landmarks on the hand, allowing for the calculation of mathematical parameters essential for arm control. These parameters are then processed and transmitted to a microcontroller, which directs the motors to rotate the robotic arm based on human movements [1]. Additionally, in the paper by Jayakumar *et al.*, the authors propose a cost-effective and scalable surveillance system that leverages a Raspberry Pi along with motion detection algorithms to improve security in various environments. The system integrates a camera module with the Raspberry Pi and utilizes OpenCV to process video feed and detect motion in real-time. By utilizing motion detection algorithms, the system can intelligently manage resources making it energy-efficient and highly adaptable. These concepts are particularly valuable when applied to gesture-controlled robotic arms, where real-time feedback and efficient resource management are critical for system performance. The authors highlight the Raspberry Pi's flexibility and its compatibility with various peripherals [6].

Later, we surveyed on a low-cost, 3D-printed robotic arm that employs gesture recognition via computer vision using a USB camera and the OpenCV library for Human-Computer Interaction. The system interfaces with a Raspberry Pi to control servo motors while monitoring pressure sensors and limit switches, ensuring safe operation and preventing overextension [2].

Polvara *et al.* emphasize that while past research on robotic arms exists, the integration of IoT technology presents new opportunities for remote control via Android devices. They recommend exploring different sensors and advanced motors, such as radio control and stepper motors, to enhance the robotic arm's performance and accuracy, ultimately benefiting industrial and smart home automation systems. The paper highlights the benefits of integrating IoT into healthcare applications, where the

robotic arm can provide practical assistance, enhance independence, and improve the quality of life for elderly users. Through smart home automation, the system can be connected to other home devices, allowing for seamless control over various home automation tasks such as lighting, temperature control, and security surveillance, all through a unified platform [3].

We continued our survey on "Sign Language Recognition and Translation with Kinect", which leverages the Kinect sensor to recognize and translate sign language gestures in real-time for the hearing impaired. It features depth sensing for accuracy, machine learning algorithms for gesture classification, and quick processing, although challenges persist in capturing nuances and accommodating different sign languages. The system uses a combination of feature extraction techniques and machine learning classifiers to interpret various sign language gestures. The authors highlight that the use of SVMs (Support Vector Machines) helps in recognizing sequences of gestures with high accuracy. However, despite these advancements, the system faces difficulties when it comes to interpreting the nuances of facial expressions and subtle finger movements that are critical in many sign languages. The challenge of accommodating multiple sign languages is also noted, as sign languages can vary significantly across different regions and cultures, which requires the system to adapt and learn from diverse datasets [4].

Finally, we surveyed a paper that presents a model-based approach for hand tracking using a single camera, integrating multiple cues like edges, optical flow, and shading to refine the model and effectively track complex articulated motions despite challenges such as occlusions and shading variations. This innovative method enhances the accuracy of tracking in dynamic environments, showcasing the potential for improved gesture recognition applications. The multi-cue approach significantly improves tracking performance, particularly in scenarios where the hand is partially occluded or when there are dynamic background changes, which are common in practical applications. The authors highlight that integrating cues from various sources (e.g., edges and flow) provides complementary information that compensates for the loss of accuracy from any single cue. This is especially beneficial for gesture-based control systems like robotic arms or sign language recognition, where precise hand movement tracking is crucial for system responsiveness and user interaction [7].

Additionally, the model refinement technique plays a vital role in adapting the tracking algorithm to dynamic environments. As the hand moves, the model is continually refined, ensuring that the system remains robust to changes in the hand's shape, orientation, or movement speed. This dynamic adjustment helps improve the flexibility and adaptability of the tracking system, making it well-suited for applications in robotic control, where real-time hand gesture recognition is essential [5].

Table 1. Literature survey.

Sr. No.	Research Paper Name	Author	Findings
[1]	Gesture-Controlled Robotic Arm Utilizing OpenCV	Paterson and Aldabbagh	Proposed an effective method for controlling devices using hand gestures, focusing on human-machine interaction.
[2]	A survey on hand gesture recognition techniques, methods and tools.	Patel and He	The survey reveals that advanced hand gesture recognition techniques achieve high accuracy and are applicable in human-computer interaction.
[3]	Vision-based autonomous landing of a quadrotor on the perturbed deck of an unmanned surface vehicle.	Polvara <i>et al.</i>	The paper highlights the use of IoT-based robotic arms and machine learning to improve gesture recognition accuracy for assisting older adults in smart home environments.
[4]	Video surveillance robot control using smartphone and Raspberry pi.	Bokade and Ratnaparkhe	The project translates sign language gestures into spoken or written language using Kinect's motion-sensing technology.
[5]	GPS based smart spy surveillance robotic system using Raspberry Pi for security application and remote sensing	Saha <i>et al.</i>	Multiple cues improve hand tracking by refining models using edges, optical flow, and shading data.

SUMMARY OF LITERATURE REVIEW

The literature shows a strong need for digital solutions in mental health support, as traditional services often face barriers such as cost, stigma, and limited availability [8–10]. Researchers have explored AI chatbots, mobile apps, and virtual platforms as effective ways to provide accessible, low-cost, and timely assistance. Studies highlight the role of multilingual models and speech recognition in reaching diverse populations, while recommendation systems enhance personalization. Overall, the reviewed works suggest that AI-driven tools can complement traditional therapy and expand mental health support globally (Table 1).

CONCLUSION

From the research papers reviewed, it is clear that gesture control and computer vision are key to advancing the interaction between humans and robotic systems. The first paper focuses on replacing traditional sensor-based gloves with vision-based motion tracking for controlling robotic arms. The system uses machine learning and computer vision to track hand movements in real time, converting these into mathematical parameters that are sent to a microcontroller to operate the arm. This method offers flexibility and eliminates the need for physical wearables like gloves, making it more user-friendly.

Another paper introduces a Raspberry Pi-based surveillance system that uses motion detection algorithms with a camera to enhance security. The integration of OpenCV allows real-time motion tracking, which is energy-efficient and scalable. The same ideas of real-time processing and efficient resource management can be applied to gesture-controlled robotic arms. This shows how IoT and computer vision technologies can be utilized for gesture recognition and control in a variety of systems, including robotic arms.

We also looked into a low-cost, 3D-printed robotic arm that uses OpenCV and a USB camera for gesture recognition. The system communicates with a Raspberry Pi to control servo motors and includes safety sensors like pressure sensors and limit switches, which help prevent overextension and ensure safe operation. This paper highlights the importance of making robotic arms affordable, while still focusing on the usability and safety of gesture-based control. Further, the paper by Polvara *et al.* emphasizes the growing potential of combining IoT with robotic systems for remote control via mobile devices [3]. They suggest that by exploring advanced motors like stepper motors and integrating IoT, robotic arms can be made more precise and efficient, benefiting both industrial applications and smart home automation. This perspective shows how IoT can bring not just control but also automation and remote accessibility to robotic systems, making them more adaptable for different uses.

The Kinect-based sign language recognition system reviewed uses depth-sensing technology to track hand movements and classify gestures in real-time. It uses machine learning algorithms like Support Vector Machines (SVMs) to recognize gestures accurately, but faces challenges in capturing the nuances of facial expressions and finger movements, which are crucial in sign language. This highlights the importance of gesture recognition in human-computer interaction, where accuracy and the ability to adapt to different types of gestures are essential.

Lastly, a paper on model-based hand tracking using a single camera discusses the integration of multiple cues like optical flow, edges, and shading to enhance tracking accuracy. This method helps address common challenges like hand occlusions or dynamic backgrounds, which can interfere with accurate gesture recognition. This multi-cue tracking approach is especially useful for systems that need to track complex hand movements in real-time, such as gesture-controlled robotic arms, and is a crucial step toward improving the robustness and adaptability of these systems.

This survey highlights the significant advancements and diverse applications of IoT-enabled and gesture-controlled robotic arms, underscoring their potential to enhance automation and user interaction across various fields.

Acknowledgment

I wish to extend my deepest gratitude and sincere appreciation to my mentor and advisor, Dr. T. V. Kafare, Assistant Professor in the Department of Electronics and Telecommunication Engineering at Sinhgad College of Engineering, Vadgaon (BK), Pune-41. His exceptional technical guidance, unwavering support, and constructive feedback have been pivotal in inspiring me to undertake this research with commitment and strive for excellence.

REFERENCES

1. Paterson J, Aldabbagh A. Gesture-controlled robotic arm utilizing OpenCV. In 2021 IEEE 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). 2021 Jun 11; 1–6.
2. Patel N, He SJ. A survey on hand gesture recognition techniques, methods and tools. *Int J Res Advent Technol.* 2018 Jun; 6(6): 939–957.
3. Polvara R, Sharma S, Wan J, Manning A, Sutton R. Vision-based autonomous landing of a quadrotor on the perturbed deck of an unmanned surface vehicle. *Drones.* 2018 Apr 14; 2(2): 15.
4. Bokade AU, Ratnaparkhe VR. Video surveillance robot control using smartphone and Raspberry pi. In 2016 IEEE International conference on communication and signal processing (ICCSP). 2016 Apr 6; 2094–2097.
5. Saha S, Singh A, Bera P, Kamal MN, Dutta S, Gorian U, Pramanik S, Khan A, Sur S. GPS based smart spy surveillance robotic system using Raspberry Pi for security application and remote sensing. In 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). 2017 Oct 3; 705–709.
6. Jayakumar AJ, Muthulakshmi S. Raspberry Pi-based surveillance system with IoT. In *Intelligent Embedded Systems: Select Proceedings of ICNETS2, Vol. II.* Singapore: Springer Singapore; 2018 Feb 17; 173–185.
7. Dziri A, Duranton M, Chapuis R. Real-time multiple objects tracking on Raspberry-Pi-based smart embedded camera. *J Electron Imaging.* 2016 Jul 1; 25(4): 041005.
8. Agarwal D, Rastogi A, Rustagi P, Nijhawan V. Real time RF based gesture controlled robotic vehicle. In 2021 IEEE 8th International Conference on Computing for Sustainable Global Development (INDIACom). 2021 Mar 17; 848–852.
9. Pradipa R, Kavitha S. Hand gesture recognition–analysis of various techniques, methods and their algorithms. *Int J Innov Res Sci Eng Technol.* 2014 Mar; 3(3): 2003–10.
10. Singh D, Zaware P, Nandgaonkar A. Wi-Fi surveillance bot with real time audio & video streaming through Android mobile. In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 2017 May 19; 746–750.