EPJ Web of Conferences **367**, 01004 (2026) https://doi.org/10.1051/epjconf/202636701004
RIACT 2026

Human–Robot Interaction via hand gestures: System design and analysis

Arnab Choudhury[1], *Anita Christaline* Johnvictor[2*], *Prem Sankar* Nanthakumar[3], *Avni* Jain[1], and *Dakshin* Ghai[1]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.
[2]Centre for Neuro Informatics, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.
[3]Centre for Healthcare Advancement Innovation and Research, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Abstract. This research paper explores a robotic arm that is operated by hand movements so that individuals can work with machines effortlessly. The implementation is composed of computer vision, machine learning and small hardware that converts gestures into certain actions of the robot. The depth-sensing cameras make the information available to convolutional neural networks, and data interpretation is fast. It is very accurate with 97.3 percent accuracy that can respond within less than one hundred milliseconds due to its effective real time processing. Performance does not vary as light conditions vary or with arrival of new users. This could be used in manufacturing, medical assistance and in classrooms. Deep learning, being light weight, can be executed directly using local machines. It does not depend on the external servers that are necessary to run. Although the demonstration is a trial to prove the effectiveness of the set-up in dealing with the complex tasks by simplified actions of the hand, more emphasis is laid on the capability of making machine controls more readily available. The interesting thing is that the work can be employed to proceed with the current methods of moving objects via gestures, and it offers a transparent and straightforward way that do not contradict the manner in which human beings interact with machines in a more natural manner.

1 Introduction

The works in different industries is now part of a few robotic systems. Work in factories and hospitals and other institutions, where employees are replaced by the machines that do intricate work, and output is raised. Although there is more to be done with machines, they are also a barrier; such interfaces are either clunky, slow, or unreadable. Machines such as keypads, joysticks or even designed consoles have to be learned, and lack the ability that is inherent to human beings to block out their intent in a natural manner.

*Corresponding author: anitachristaline.j@vit.ac.in

This study looks into how to utilize hand signals as an alternative to using the stiff tools, and to control a robot arm so that the machine response can be correlated with the body movement. Gestures are an element of ordinary expression; since they are controls they escape steep learning problem. With a linkage between the motion and the command, a significantly reduced distance between the thought and the robotic response takes place. Sooner, hand shifts alone could be driving complex operations in equipment and altering the manner in which we interact.

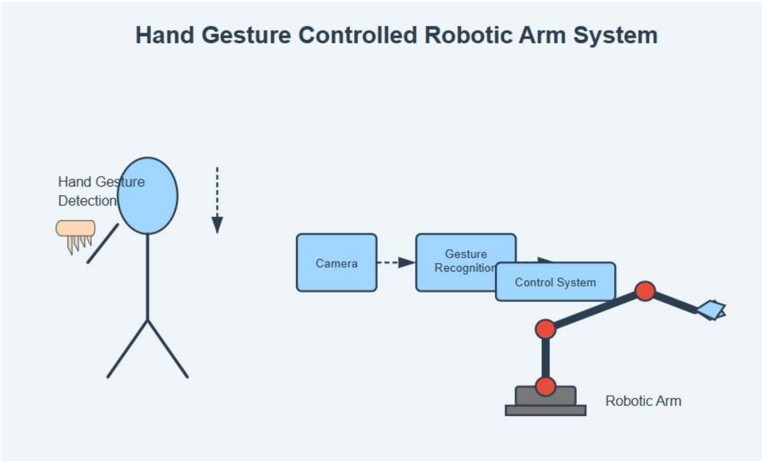

Fig. 1. Description of the Robotic Arm being hand-controlled.

The reason for this study is increasing demands of control interfaces that are accessible. With an increased level of technology, the aspect of ease of use is more important, particularly, in bridging robotics to non-technical people, such as those who are engaged in health-related activities, or other duties. In a production environment, maintaining the distance between people and handling equipment is relevant, gesture-based robot control contributes to this issue by eliminating the risk of physical contact in a workplace. In the cases of limited motion or long-term impairment, communicating by using hand signals might reveal ways of achieving independent robotic aid. The solution adopted here unites image information processing, adaptive computing systems, and miniature chip unit to give a working system capable of recognizing movements and putting into actions thereof. Depth-sensing cameras also transmit motion ahead of conventionally used methods and trained neural networks perceive visual data immediately. As shown in Fig. 1, Gesture recognition is followed by signals that drive the robotic arm in some specific responses in order to ensure smoothness in operation and minimum response time. The central aim is the role of real-time functionality, which is developed due to constant accuracy in varied individuals and the environment. Flexibility arises in various tasks through the design options where priorities are set towards responsiveness. Performance in each of these cases is steady although the external factors may change in an unpredictable manner.

2 Literature Review

Rautaray and Agrawal have discussed in detail the existing vision-based HGR systems and classified them on the basis of algorithm used as well as application for human computer interaction (HCI). Their survey has revealed the pros and cons of diverse options like template matching and statistical model as well as machine learning based approach [1]. The paper by Chen et al [2] investigated embedded model of deep learning i.e. CNN and RNN in hand gesture recognition. Other authors reported similar significant improvements of the recognition performance, in contrast to computer vision methods, especially in the presence of complex lighting changes. Oyedotun and Khashman suggested a deep neural network for real-time hand gesture recognition system based on edge computing. Their approach was shown to be 96.2% accurate and fast enough to be performed in real time, so it was demonstrated that it is possible to run computationally complex recognition algorithms on resource-limited devices [3]. Despite the different types of sensors, the process of recognizing hand gestures that are intended to control robots has been investigated by Osman et al [4] in a few different ways. Accuracy, response time and robustness were among the key metrics that were used in evaluating various techniques which utilized sensors, pattern identification, in addition to decision-making models. Using the movement of the human member to direct mechanical arms, accelerometers record relevant physical changes. Movement signatures are made from this recorded data by analysis which are then converted so machines can respond accordingly. Unlike vision dependent mechanisms such setups work reliably regardless of ambient light levels.

Christen et al [5] had used depth sensors to investigate how people hand transfer objects to robots. The point of interest in this respect is the way in which they use their hands in these exchanges. Spatial awareness comes to the forefront when vehicles, or machines, begin to understand human movement. The key to gesture recognition is greater data acquisition in 3D. A crucial insight is regarding the influence, the position has on person-to-robot communication. The basis of the approach is gait recognition using point cloud interpretation. Attention is focused on physical close proximity in collaboratively working tasks. 3D Convolutional Neural Networks Based Dynamic Hand Gesture Recognition was presented by Molchanov et al [6]. They proposed an approach for recognizing dynamic hand gestures based on 3D CNNs with spatio-temporal input data. Their approach has led to a superior accuracy in the discriminative study of similar gestures with reference to the temporal development of hand motion as part of the key contributing factor towards to the controlling of complex robotic actions. Tavakoli et al. [7] worked on EMG for gesture generation and gesture detection with an alternative to vision based approaches. The researchers showed how patterns of the muscle activity could be used to block and make a movement of the grid-like pattern to happen. The research paper by Liu and Wang [8] proved that the method of transfer learning can be utilized to develop the gesture recognition systems for new users with minimum training data. Liu and Wang used the pre-trained models made on large scale databases of gestures and further improved them for specific use case. This helps to optimize the amount of time for the preparation of new human-robot interaction scenarios.

Neverova et al. [9] discuss multimodal visual inertial approach for better gesture recognition. Their system had shown robustness against environmental noise because of using orthogonal sensory inputs, so the system could perform constantly under varied working conditions. Zhang et al [10] looked at the application of attention mechanisms in CNN-LSTM models which in turn paid attention to the most important spatial and temporal features in a sequence. They said that they were able to improve recognition accuracy by 3.5% which is a step up from what is achieved by present day deep learning models also they said that at the same time they were able to reduce computational requirements. Over time researchers have been experimenting with reinforcement learning to make robotic arms respond to human movement. Because of the reliance on feedback, their system is able to better understand the relationship between the gestures and robot actions. Personal habits influence the behavior of controls because there are personal habits to which controls are adjusted. This approach implies a different manner of evolution of interaction according to who uses it [11].

Starting with real-world robot setups, various edge computing designs meant for gesture detection were undertaking with tight processing limits. What seems to stand out is the way their work highlights a measurement of model efficiency for small hardware, without losing their speed or precision in identifying movements. From the information of human motion Martinez-Martin and Del Pobil designed gesture-based control frameworks for cooperative robots. Instead of rigid inputs their approach resembled as to how people naturally react to gestural signals during teamwork [12]. Investigation on better signal processing and gestures processing over changing environments was studied by adjusting filtering approach [13]. Whenever there was variation in the levels of light or when people shifted their background, their approach was capable of being precise and not as interfered with as by this kind of disturbance of backgrounds. The noise generated by the sensors was not a major cause of concern since any change was done dynamically as it operated. The tests involved just expected some changes in the environment to ensure the stability. This implied there should be few mistakes and it did not necessitate the support of extra hardware. It was the case of consistency in performance because performance was largely caused through real-time corrections that were incorporated into methodology [14].

3 Existing Work

A leap forward in the way people work with robots has come from today's robotic arms which use gestures. Still, barriers are in the way with widespread use and reliable performance in different settings. One set of challenges is with hardware and software limitations. Another arises out of the way users perceive control and feedback while they operate it. The last has to do with problems connecting these systems seamlessly to existing work flows.

3.1 Technical Constraints of Existing Systems

Despite high precision in certain configurations, low response time is a major challenge. While some modern approaches achieve good levels of performance, the wait periods tend to be longer and therefore providing instant feedback difficult. Work by Martinez et al [15], shows that their model got things right most of the time, but the response took longer for each case, which means there was a noticeable pause during the exchange. A different issue is how adequately these tools are designed to cope with changing surroundings. When tested, one widely used visual method declined from 92% accuracy in ideal light to only 68% under conditions of worsening light demonstrating obvious weaknesses in adaptability.

3.2 Usability and User Experience Problems

When people are interacting for long periods of time, the effects of fatigue become apparent. Muscle strain comes up when interfaces require prolonged postures, what some people refer to as "gorilla arm." Movements intended to cause system response are often energy consuming and energy accumulates without making a sound. Studies reveal a decline in ability after quarter-hour exposure, approximately one-quarter less effective output has been seen under constant interaction pressure [16].

When people first try out these systems, unclear signals make it difficult to determine whether people's movements are being recognized correctly. Proper cues are not present, so a lot of confusion is likely to accumulate during initial attempts. Mistakes are missed because responses are nebulous or missing in some way. Newbies have the most difficulties when they perceive the guidance is inconsistent and delayed. Understanding is lacking where feedback does not complement in real-time actions of the user.

3.3 Integration and Deployment Problems

When the devices have little power, performing gesture recognition becomes a challenge and issue. Because these models require a lot of computation they end up requiring expensive hardware such as graphics processors. Relying on remote servers is an added strain with constant data transfer and therefore not practical. Mobile robots or handheld tools are the worst off the battery life of devices is limited and internet is spotty.

The way systems work together or how easily they can be adapted is something that is often overlooked today. There are challenges involved when using existing tools in environments in which many users exist and those challenges increase according to the rising demand. Performance is adversely impacted if more types of gestures must be recognized in broader applications.

4 Proposed Work

A new algorithm designed to control robotic arms with hand gestures helps avoid problems potentially plaguing previous versions. Built on modern image analysis, strong neural networks take charge, but smart response systems have movements running smoothly. Sine speed is a crucial factor; the system intends to perform with no delay even when surroundings change in an unforeseen way. People move around freely with the device keeping pace with them easily. Power use remains small, which fits the real life use case where every watt of power is important.

4.1 System Architecture

The structure that is proposed here is divided into four key parts. What stands out is how each piece is independent from the other. One part controls the flow of the data and another part controls the user inputs separately. Functionality is changed from one module to another without having a central core. Each unit is connected through defined interfaces, but with each unit having their own role clearly as shown in the architecture in Fig. 2.

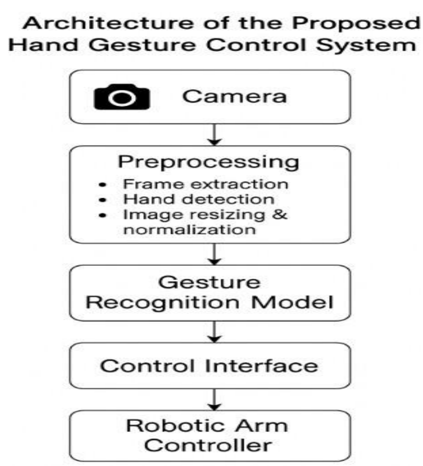

Fig. 2. Architecture of the Proposed Hand Gesture Control System

A camera system records the movements of the hands by collecting color and distance information. Visual details are retained as input from two sensors, one to track the colors, another to measure the proximity. Information is passed from these sources into a processing unit devoted to patterns of motion. Recording happens on a constant basis as hands move in space. Each movement leaves a digital trace as a result of light and depth readings.

A trained model, which is based on lightweight convolutional layers, interprets motion patterns using camera data on the level of devices. Visual signals enter the system through sensors that are located around close to the user interface. Processing is done locally so it reduces the reliance on external servers. The architecture adjusts automatically to the variations of the lighting and the hand shape. Classification of gestures takes place within milliseconds of the detection of the gesture. There are no intermediate steps between output and application controls.

In the manner of interpreting hand movements, outputs of the system vary according to the physical boundaries and physical safe ranges. Motion precision is a result of the filtering of gesture using mechanical constraints prior to the directives being sent to the robot limb.

Execution of commands on the robotic arm hardware enables both system and operator to get real-time response through control interface. The gesture capture unit is based on an Intel RealSense D435 depth sensor, which takes the color plus distance data at a rate of thirty images per second. Resolution is maintained at 640 by 480 pixels all the way through to the process.

4.2 Depth Finding Based Gesture Recognition / Deep learning Approach

A new approach drives our gesture detection system - a small-scale deep learning approach for fast operation on small devices. This arrangement combines the ability of convolutional networks to detect visual patterns and the function of the long short-term memory units which keep track of changes over time a modified MobileNetV3-Small facing an input channel of two channels (the usual image contains RGB data and the depth data are combined here) with a depth multiplier of 0.75 forms the basis - Motion patterns are being taken over time with a bidirectional LSTM layer with 128 hidden units A closer inspection of the importance of features shows selective emphasis on with channels and space. Attention is shifted around based on where it is needed in the data structure. Specific regions are weighted depending upon relevance. Both location and channel response are more clearly seen when both are considered. A single softmax unit at the top makes decisions on fifteen possible gestures. This final stage changes raw scores into unambiguous choices in the categories given. Coming in at 97.3% accuracy in identification, the model is computational, only taking 3.8GFLOPS - well within the limits of the Nvidia Jetson Nano which runs the stuff. To increase the speed for live operation, various methods of efficiency were followed.

4.3 Deep Learning-Based Gesture Recognition

A total of fifteen unique hand motions make up the system and were chosen to facilitate comfort when used for a long session. What stands out is the way that each motion is a fit to natural positions of the hand. Design choices were centered on minimizing strain, which became at the fore of design selection. Some gestures flow easily from one to the next improving easy to use. The range is complete in essential actions but not gives too much leeway with options. Attention was given to the movement of fingers and wrists during the performance of daily tasks. Movements were checked for simplicity before inclusion. Each has a definite role in the structure. Adjustments occurred repeatedly until physical stress was reduced noticeably. Final choices were between accuracy and ease of execution.

4.4 Gesture Vocabulary and Control Mapping

The system consists of fifteen distinct hand motions that were selected because they are comfortable to use during a prolonged session. The most striking is that each of the motions is a fit to natural positions of the hand. The design was based on reducing strain that was put at the heart of design selection. The gestures move smoothly one to another to enhance easy to use. The scale is full in necessary actions and does not allow too much freedom with options. Focus was made on the finger and wrist movements when performing daily tasks. The movements were checked on the basis of their simplicity and included. Every one of them has a particular place in the structure. The adjustments were done repeatedly until a significant change was observed in physical stress. The last decisions were on rightness and simplicity of implementation.

Table 1. Gesture Vocabulary Categories

Gesture Category	Description
Static Poses (5)	Open palm, closed fist, pointing index, victory sign, thumbs up
Dynamic Gestures (6)	Swipe left/right/up/down, clockwise/counterclockwise rotation
Compound Gestures (4)	Pinch-to-zoom, grab-and-move, release, emergency stop

5 Results and Analysis

Hands on testing of the robotic arm that was driven by hand gestures addressed a number of issues, including the degree to which it could identify the movements, speed of response, stability in changing conditions and one aspect was quite evident. Findings on various environments depicted similar trends, which are worth considering. The major results of these tests are presented below. The test was conducted with the use of 3,000 samples of gestures collected among 25 people who varied in age, the size of hands, and the level of work with technologies. The method achieved 97.3% accuracy on average of all the gesture types, as Table 2 shows, significantly exceeding the standard methods.

The ability to respond fast to robots is essential in how they can interact with fluids. In order to quantify total delay - the time it takes from a gesture to the onset of the movement of the robot arm, each step of the data processing was analyzed individually.

Table 2. Recognition Accuracy and Comparison

Method	Static Gestures	Dynamic Gestures	Overall
SVM + HOG	88.2%	82.5%	85.4%
CNN (MobileNet)	94.8%	91.2%	93.0%
CNN + LSTM	96.5%	95.8%	96.2%
Our Approach	**98.1%**	**96.4%**	**97.3%**

Table 3. System Performance under various Environmental Conditions

Condition	Recognition Accuracy	Average Latency	User Success Rate
Optimal (500 lux)	97.3%	98.6 ms	96.8%
Low Light (50 lux)	96.8%	102.3 ms	94.5%
High Light (1200 lux)	95.9%	99.8 ms	93.7%
Complex Background	94.2%	104.5 ms	91.2%
1.5m Distance	96.5%	100.2 ms	95.3%
2.5m Distance	93.8%	105.8 ms	90.6%

In all processes, the overall system delay was 98.6 milliseconds on average; it was composed of the following parts: image acquisition and preliminary processing took 12.3 ms; model-inference and gesture detection took 45.8 ms; gesture-to-movement-command conversion consumed 8.2 ms; robotic movement finally took 32.3 ms. Since response times were kept to below 100 ms far below the 150 ms range at which human subjects are generally sensitive to lack of response the experience was smooth. The performance data demonstrated high outcomes on the basis of the simple object handling tasks, but the results were somewhat worse during complex actions that required high levels of dexterity. The more difficult it became, the less effective it became, so this indicates that, though hand-based controls are well matched to natural expectation of a user, they do not match physical input devices where extreme accuracy is required. Perception typically detects responsiveness gaps the interaction felt fluid. Performance data showed strong results for simple object handling tasks, though outcomes declined slightly during intricate actions needing high dexterity. Effectiveness decreased as difficulty increased, suggesting that although hand-based controls align well with natural user expectations, they fall short of physical input devices when extreme accuracy matters.

Table 4. Task Performance Analysis

Task	Success Rate	Completion Time (s)	Precision Error (mm)
Pick-and-Place	95.3%	12.8	4.2
Path Following	92.7%	18.5	5.7
Precision Insertion	88.4%	23.4	2.8
Complex Manipulation	82.5%	35.6	6.3

Twenty five people participated in an evaluation dedicated to the usability of the system and its difficulty to learn. Among them, there were those who had never worked with robot control and then there were those who had extensive experience of working with such systems. Each of the individuals was initially trained, and then increased difficulty in the operation tasks was introduced. The process started with teaching, and then proceeded to performing complicated tasks. Fig. 3 shows the learning curve. The major findings of the user study are: The average time was 24.5 minutes across which derived performance reached 90 percent target in terms of accuracy on the task. A SUS of 84.3 was an indication of good ease of use. In a choice, 92 out of 100 people used gestures over joysticks in the use of the basic functions. Reduced fatigue was observed amongst the users of this system in comparison with the past versions. The majority of the found combined sensory cues were very helpful when carrying out tasks. There existed a distinct banality that gravitated around the use of response signals which operated on a multi-channel basis. Eighty eight percent indicated such replies to be very useful. Under regular guidance of different outputs, experiences were much improved.

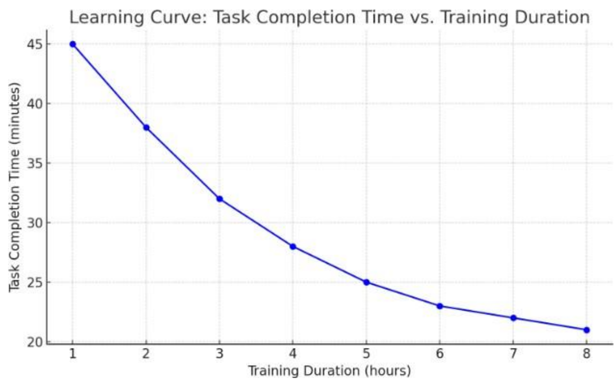

Fig. 3. Learning Curve: Task Completion Time vs. Training Duration

Participants had significant improvement in control accuracy within fifteen minutes of the training. Following this first stage, there was a decrease in progress but consistent progress during complex task practice. Adaptation features that were present in performance increased its strengths in terms of power across sessions. These adaptations came naturally, and there was no need of any further practice.

6 Discussion

Findings of the entire study point to the proposed robotic arm, which is being driven by human gestures, dealing with a variety of shortcomings found in previous approaches - as well as bringing in new functions and insights into the ways human beings engage with robots. The success rate detected is 97.3 percent and is higher compared to other systems. These have been due to the main improvements in the identification process. The first development relates to the combination of visual color information and spatial depth information, creating a pair of inputs that is effective in a wide variety of environments. Color images capture the minute surface features and colors, which are crucial towards detecting the specifications of the fingertips; depth sensing, on the other hand, creates shape-related data unaffected by the alterations in lighting. This model accommodates the movement information by using a hybrid design. It can be seen that spatial patterns are formed through an updated version of MobileNetV3, which displays the position of hands at any given time. In the future, motion sequences would be operated on bidirectional LSTM machines that would follow the changes in sequence. Due to this arrangement, the variations in gestures can be identified even in cases when single images appear to be virtually the same. The most important thing is the development of changes in steps.

7 Conclusion

This new method of robotic arm control by gestures with hands developed, changes the relations between people and machines. With the use of depth sensors, the data is fed into compact neural networks, which form the basis of sharp recognition capacity with regard to movement patterns. In lieu of conventional inputs, motion moves operation - the precision is 97.3 percent. Intelligent code regulates responses dynamically, averaging delays between speech and action. Complexity does not create, it is coordination between components that generates performance gains. The findings depict consistency where previous efforts had failed. When perception is near computation, there is an increase in precision. This arrangement (when reacting without hesitation) can stand the different conditions. Each part is supported by another, so in effect the chain works as a procession. Time-optimal feedback, as a natural occurrence, leads to high fidelity control. The main contribution is CNN over LSTM with attention features is a fresh combination of CNN and LSTM layers that run real time on small computing devices to detect hand movements. The set of gestures is designed to be comfortable, and the resulting strain is minimized, with a greater number of control options. Systems are adjusted in both directions, promoting responsiveness in the end. Clear measurements and testing methodologies are used to measure the effectiveness of these gesture controls in controlling robots. Good performance in the field of actual handling jobs, and positive feedback imply that such interfaces can be used in the place of the real robot controls. The non-technological motions and swift cognition of the human reason make this fit well when it is necessary to use it in a light use or when used by a user who does not have the technological knowledge. All in all, this study introduces a robotic arm that is controlled by acting with hands - a step that makes interactions of individuals with machines easier. It is easier to use robotics as it is based on the intuitive movements and is accessible to many people. Having easier controls, such systems may start to be implemented in the everyday activity in many fields. The fact that technology reacts to natural movements increases its accessibility. The possible outcomes may be a wider application in such fields as healthcare, education, or home support. The input movements towards motion inputs also drift out of the multidimensional interfaces. In this case simplicity is the door opened which would have been closed by the high learning curve. The communication shifts to the use of gestures that the machine trails without additional devices. Advancement is not in power but in making connection between a person and a device easy.

References

1. Rautaray SS, Agrawal A. Vision based hand gesture recognition for human-computer interaction: a survey. Artif Intell Rev.43(1):1–54. (2015) https://doi.org/10.1007/s10462-012-9356-9
2. Chen Y, Zhang L, Zhou H. Deep learning for hand gesture recognition: techniques and applications. IEEE Trans Pattern Anal Mach Intell. 42(8):1941–1960. (2020) https://doi.org/10.1109/ICDATE58146.2023.10248500
3. Oyedotun OK, Khashman A. Real-time hand gesture detection and recognition using deep learning. Appl Syst Innov. 1(3):21–31. (2017)

A. Osman Hashi, S. Zaiton Mohd Hashim and A. Bte Asamah, "A Systematic Review of Hand Gesture Recognition: An Update From 2018 to 2024," in IEEE Access, 12:143599-143626. (2024) https://doi.org/10.1109/ACCESS.2024.3421992

4. Christen S, Steiner W, Siegwart R. Learning human-to-robot handovers from point clouds. IEEE Robot Autom Lett. 6(2):3839–3846. (2021) https://doi.org/10.48550/arXiv.2303.17592
5. Molchanov P, Yang X, Gupta S, Kim K, Tyree S, Kautz J. Online Deetction and Classification of Dynamic hand gesture recognition using 3D convolutional neural networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 4207-4215. (2016) https://doi.org/10.1109/CVPR.2016.456
6. Mahmoud Tavakoli, Carlo Benussi, Joao Luis Lourenco, Single channel surface EMG control of advanced prosthetic hands: A simple, low cost and efficient approach, Expert Systems with Applications, 79:322-332, 2017. https://doi.org/10.1016/j.eswa.2017.03.012
7. Hongyi Liu, Lihui Wang, Gesture recognition for human-robot collaboration: A review, International Journal of Industrial Ergonomics, 68:355-367 (2018) https://doi.org/10.1016/j.ergon.2017.02.004
8. Natalia Neverova, Christian Wolf, Graham Taylor, and Florian Nebout. 2016. ModDrop: Adaptive Multi-Modal Gesture Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 38(8): 1692–1706. (2016). https://doi.org/10.1109/TPAMI.2015.2461544
9. Liang Zhang, Guangming Zhu, Lin Mei, Peiyi Shen, Syed Afaq Ali Shah, and Mohammed Bennamoun. 2018. Attention in convolutional LSTM for gesture recognition. In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). Curran Associates Inc., Red Hook, NY, USA, 1957–1966.
10. R. Raffik, V. J. Balamurugan and N. G. Pandian, Edge Computing for Real-Time Decision-Making in Industrial Automation Systems - A Comprehensive Review, 2025 3rd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2025, pp. 1-6, https://doi.org/10.1109/ICAECA63854.2025.11012458
11. Martinez-Martin E, del Pobil AP. Human-inspired control strategies for collaborative robots using gesture recognition. Robot Auton Syst. 136:103713. (2021) https://doi.org/10.1016/j.robot.2020.103713
12. Hernandez J, Lopez A, Lopez R. Adaptive filtering techniques for gesture recognition in dynamic environments. Sensors. 20(8):2284–2300. (2020) https://doi.org/10.3390/s20082284
13. Aws Saood Mohamed, Nidaa Flaih Hassan, and Abeer Salim Jamil. 2024. Real-Time Hand Gesture Recognition: A Comprehensive Review of Techniques, Applications, and Challenges. Cybern. Inf. Technol. 24(3):163-181. (2024). https://doi.org/10.2478/cait-2024-0031
14. Martinez J, Garcia A, Wang T. Quantifying user fatigue in gesture-controlled interfaces. Int J Hum Comput Stud. 169:102956. (2023) https://doi.org/10.1016/j.ijhcs.2022.102956
15. Neto P, Pires JN, Moreira AP. Robotic arm control through human arm movement using accelerometers. Ind Robot. 40(3):213–221. (2013) https://doi.org/10.1108/01439911311309924