

Hand Gesture Recognition for Touchless Human–Computer Interaction

Sachin Kumar Gupta*

Abstract

As technology continues to evolve, traditional input devices like the physical mouse can become restrictive, particularly for individuals with physical impairment or those working in hands-free environments. Alternative interfaces have been developed in response to the desire for more accessible and natural ways to interact. To provide a more adaptable and inclusive human–computer interaction experience, this study offers a gesture controlled virtual mouse system that integrates voice command with hand gesture recognition. Through customized hand motions recorded by a webcam, the suggested system allows users to carry out standard mouse functions, including clicking, dragging, and scrolling. Additionally, it integrates speech recognition, which enables users to utilize spoken instructions to control programs. The dynamic toggle function, which enables smooth switching between a conventional mouse and the gesture-voice interface, is a crucial component of the system. To allow users to start with familiar hardware before deciding to enable the sophisticated interface, the system starts up automatically with the virtual controllers off. The system leverages computer vision techniques and machine learning algorithms to interpret gestures and voice commands in real time, ensuring high accuracy and responsiveness. Testing revealed that the system performs reliably across different lighting conditions and user hand sizes, with minimal latency in gesture recognition and voice response. This makes it suitable for both general use and as an assistive technology for individuals with conditions such as arthritis or carpal tunnel syndrome. Overall, the proposed system offers a practical, accessible, and efficient alternative to conventional input methods.

Keywords: Gesture recognition, virtual mouse, human–computer interaction, voice command, computer vision, accessibility, machine learning

INTRODUCTION

The way people interact with computers has significantly evolved over the years, yet traditional input devices, such as keyboards and mice, remain the most commonly used. These tools have become second nature to most users, but they are not without limitations. For individuals with physical impairments, such as arthritis or carpal tunnel syndrome, using a mouse or keyboard for extended periods can be painful or even impossible. Additionally, in situations where hands-free operation is needed—such as while cooking, working in a lab, or presenting—standard input methods fall short. As digital systems

become increasingly integrated into everyday life, there is a growing demand for alternative interaction techniques that are more flexible, intuitive, and inclusive.

In response to this need, the development of gesture- and voice-based interfaces has gained momentum. These technologies offer the possibility of interacting with a computer in a more natural and accessible way. Instead of relying on physical contact with input devices, users can issue commands using hand gestures or spoken

*Author for Correspondence

Sachin Kumar Gupta
E-mail: sachin@gnct.co.in

Assistant Professor, Department of Computer Science & Engineering, Greater Noida College, Greater Noida Uttar Pradesh, India

Received Date: July 02, 2026
Accepted Date: September 02, 2026
Published Date: September 12, 2026

Citation: Sachin Kumar Gupta. Hand Gesture Recognition for Touchless Human-Computer Interaction. Journal of Advances in Shell Programming. 2026; 13(2): 21–29p.

instructions. This form of interaction not only has the potential to reduce physical strain but also offers greater convenience and speed in many scenarios. This study introduces a gesture-controlled virtual mouse, a system designed to enhance the way users interact with computers by combining gesture recognition with voice control. The system allows users to perform basic mouse functions—such as moving the cursor, clicking, dragging, and scrolling—using hand movements detected through a standard webcam. At the same time, it supports voice commands to operate applications, open files, and perform system-level tasks without the need for any physical input. This combination is particularly useful for users who may have difficulty using a standard mouse or keyboard, as well as for those seeking a more efficient or hands-free method of control.

One of the unique aspects of this system is the inclusion of a dynamic toggle switch. This toggle enables users to switch between the traditional mouse and gesture-voice interface with ease. By default, the system launches with the virtual interface turned off, allowing the user to begin with their regular mouse setup. If needed, the user can manually activate the virtual interface. This flexibility ensures that users are not forced into a new mode of interaction but can instead choose what works best for them at any given moment. Additionally, the system remembers the last toggle state even after shutdown, so users do not have to reconfigure their preferences each time the computer restarts. Technically, the system uses machine learning and computer vision algorithms to process video input and recognize specific hand gestures in real time. Voice input is handled using speech recognition libraries that interpret spoken words and match them to predefined commands. The software is lightweight and designed to run on general-purpose computers without the need for additional hardware, making it both practical and cost-effective. The primary aim of this work is to create a more inclusive and user-friendly method of computer interaction. By focusing on accessibility, ease of use, and adaptability, the gesture controlled virtual mouse offers a meaningful improvement over traditional input devices for a wide range of users. In the following sections, we discuss related work in this field, describe the methodology used to build the system, and present the results from testing the system in various environments. Through this study, we hope to contribute to ongoing efforts to make computing more accessible for everyone.

REVIEW OF RELATED WORK

The development of gesture- and voice-based human-computer interaction (HCI) systems has seen a significant rise due to the increasing demand for more natural, hands-free, and accessible interfaces. Traditional input devices, such as keyboards and mice, often limit user accessibility, particularly for individuals with physical impairments. To address this issue, various studies have proposed systems utilizing computer vision, machine learning, and sensor-based technologies to enhance user interaction. Several researchers have focused on the implementation of static hand gesture recognition using deep learning techniques. A convolutional neural network (CNN)-based model was introduced [1], showing high accuracy in detecting static hand gestures, which is essential for emulating basic mouse operations such as left click, right click, and scrolling. The importance of data diversity and augmentation in gesture datasets was highlighted [2], where CNNs were trained on expanded datasets to increase recognition accuracy and robustness under different lighting and background conditions. More dynamic gesture recognition systems have also been proposed. A wearable band system using six-axis sensors was developed [3, 4], enabling users to control video games through real-time dynamic gestures. Another system built using flex sensors and machine learning algorithms [5, 6] allowed more detailed movement recognition, broadening the scope of gesture applications. These solutions are particularly beneficial for enhancing the accuracy and responsiveness of gesture-controlled interfaces. In terms of accessibility, glove-based gesture detection systems have shown great promise [7]. A wearable glove that could translate hand movements into system commands was described [8], targeting communication for mute individuals. These wearable solutions demonstrate how gesture recognition systems can be tailored to meet the needs of differently abled users.

Parallel to the development of gesture-based systems, voice-controlled systems have also become

increasingly popular. A voice-controlled home automation setup was proposed [9], allowing users to control appliances without physical interaction. Similarly, a basic voice assistant framework for managing smart systems was introduced [10], further showcasing the practicality of speech recognition in daily tasks. Such systems are especially advantageous for users who find gesture-based controls physically demanding. A growing number of systems now integrate both gesture and voice controls into a unified interface. For example, a smart classroom environment utilizing gesture, voice, and facial recognition was implemented [11], demonstrating the feasibility of combining multiple AI techniques for user authentication and control. Another project [12, 13] introduced an enhanced AI voice assistant using natural language processing, enabling seamless interaction with machines through conversational commands.

A gesture-controlled virtual mouse system incorporating both vision-based gestures and voice control was detailed [14]. The system demonstrated how gesture and voice features could be harmonized to offer an alternative method for standard computer interaction [15]. This aligns closely with the proposed system in this study, which integrates these technologies into a single dynamic framework.

Several gesture-based virtual mouse systems have been developed using open-source technologies. An example is the system designed [16], where OpenCV and Python were used to detect hand gestures and simulate mouse functions. A similar concept was explored [17], where a virtual mouse called “Hand Glide” was introduced, offering both gesture-based control and voice interaction, aimed at improving accessibility for general users. Some studies have explored security-enhanced gesture systems. A hybrid approach combining gesture recognition with biometric verification was introduced [18], aiming to add a layer of user identity verification. This is crucial in scenarios where gesture-based controls might be applied to secure environments or personal devices. Innovative research has also been conducted on non-traditional interaction methods. For instance, eye blink detection was used as a communication medium [19, 20], while Borah et al. (2022) [21] and Bhargavi and Prabhakar (2022) [22] compared various algorithms to detect emotional responses through eye activity. Although different from gesture control, these studies emphasize the importance of alternative communication methods in inclusive design.

Additional research has explored context-specific applications of gesture control. A gesture-controlled video player was developed [23], and a mobile manipulator system controlled by one-handed gestures was demonstrated [24]. An early approach to gesture control for PC applications was outlined [25], which laid the groundwork for many systems developed in recent years. Although prior work has achieved considerable success in gesture recognition or voice-controlled systems independently, few have combined them in a modular and flexible architecture. This study aims to fill this gap by presenting a gesture-controlled virtual mouse with a dynamic toggle feature, offering users the flexibility to switch between traditional input, gesture recognition, and voice commands based on their preference or physical ability.

SYSTEM ARCHITECTURE AND METHODOLOGY

The proposed system introduces a gesture-controlled virtual mouse with integrated voice control, designed to offer an accessible and hands-free alternative to traditional input devices. This system is particularly beneficial for users with motor disabilities or in situations where physical contact with input devices is inconvenient or unsanitary. The architecture is modular, allowing users to seamlessly switch between gesture control, voice commands, and traditional input methods using a dynamic toggle feature. The core idea is to provide flexibility and adaptability based on user preferences and requirements. The system is composed of three primary modules: gesture recognition, voice control, and mouse control interface. Each module is independently functional but interconnected to operate cohesively through a control manager that determines the active mode.

Gesture Recognition Module

The gesture recognition module is based on computer vision techniques using the OpenCV library.

The system captures real-time videos from a webcam and processes the frames to detect hand landmarks. This is achieved using a pre-trained MediaPipe model, which provides reliable detection of 21 key points on the hand. Once the hand landmarks are detected, specific gestures are interpreted based on the position and distance between fingertips and palm centers. For instance: index finger up and others down: cursor movement; index and middle fingers up: left click; thumb and index finger touching: right click; palm open: pause mode or no action. These gestures are mapped to corresponding mouse operations. The fingertip coordinates are translated to screen coordinates using a calibration function that accounts for screen resolution and camera orientation. A smoothing algorithm is also applied to avoid abrupt cursor movement caused by minor hand jitters.

Voice Control Module

In addition to gestures, the system supports voice commands to provide users with another mode of interaction. The voice module uses a lightweight speech recognition library to convert spoken commands into text. Common mouse operations such as “left click,” “right click,” “scroll up,” “scroll down,” and “open application” are recognized and processed. A keyword-based detection model filters relevant commands from background noise. Once the command is verified, it is routed to the control manager, which triggers the corresponding mouse function. Voice control is particularly useful in scenarios where the user is unable to perform gestures or requires faster execution of commands.

Mode Toggle Feature

A central feature of the system is the dynamic toggle mechanism that allows users to switch between gesture, voice, and traditional input modes. The toggle can be activated through a designated voice command (e.g., “Switch to gesture mode”) or a specific gesture (e.g., raising all five fingers for two seconds). This functionality ensures that the system is user-centric, providing control based on current physical ability, environmental context, or task-specific requirements. For instance, a user may prefer gestures during presentations but rely on voice commands when their hands are occupied (Figure 1).

Mouse Control Interface

All detected commands from the gesture or voice modules are handled by a unified interface that uses the PyAutoGUI library. This library sends OS-level mouse and keyboard events to simulate real user input. Actions such as cursor movement, clicks, and scrolling are performed in real time, providing the system with a smooth and responsive experience.

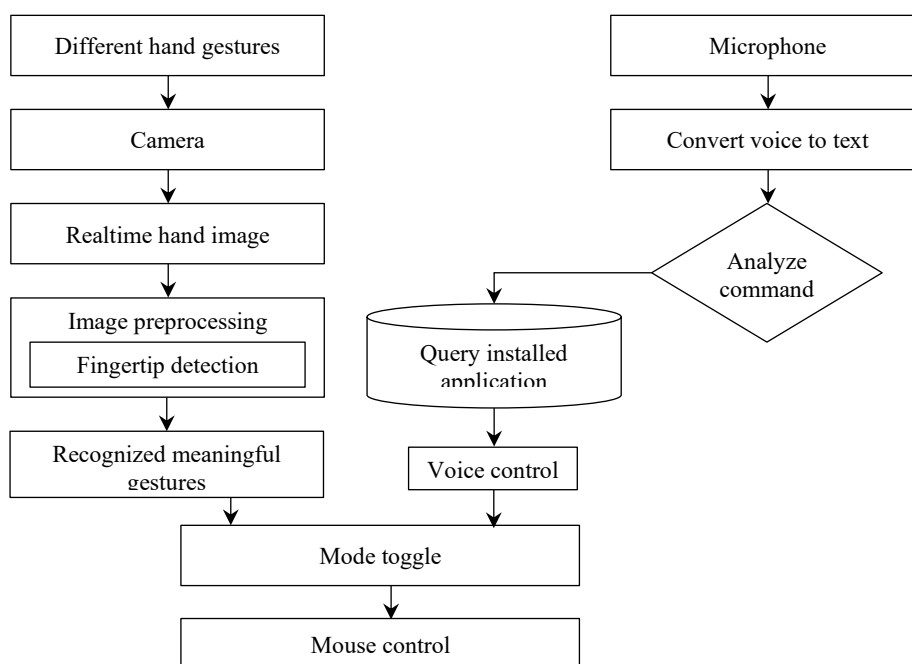


Figure 1. Architectural diagram for hand gesture.

In addition, safety checks are included to prevent unintended actions. For example, the system pauses input if no hand is detected for a certain time or if voice confidence scores are low. By combining gesture and voice input, the system provides a flexible and accessible human–computer interface suitable for modern interactive applications.

IMPLEMENTATION

The proposed system was implemented as a standalone desktop application capable of responding to both hand gestures and voice commands. The goal was to develop a flexible and user-friendly interface that could interpret human actions and translate them into corresponding computer operations, such as mouse control and application navigation.

Control Flow Algorithm

The complete control flow of the system is described in Algorithm VI. It summarizes the interaction between the two input modalities and outlines how transitions, state retention, and session termination are managed.

Algorithm 1. Gesture and voice command controlled system.

Require: Gesture or Voice Command

Ensure: Mouse Actions or Application Control Operations

```
1: Initialize System
2: Launch Application on Home Screen
3: while System is Active do
4:   if Dynamic Toggle is Enabled then
5:     Activate Gesture Recognition Module
6:     Allow User to Customize Gestures
7:   end if
8:   if Gesture Input is Detected then
9:     Perform Corresponding Gesture Action
10:    if Transition Gesture is Detected then
11:      Deactivate Gesture Module
12:      Activate Voice Command Module
13:    end if
14:  end if
15:  if Voice Command is Detected then
16:    Execute Corresponding Voice Operation
17:    if Transition Voice Command is Detected then
18:      Deactivate Voice Module
19:      Activate Gesture Module
20:    end if
21:  end if
22:  if System Shutdown or Sleep is Initiated then
23:    Save Current Module State
24:    Exit or Enter Sleep Mode
25:  end if
26:  if System Restart is Detected then
27:    Restore Previous Module State
28:    Go to Step 2
29:  end if
30:  if User Terminates Session then
31:    Stop All Modules and Exit
32:  end if
33: end while
```

RESULT AND PERFORMANCE EVALUATION

The gesture controlled virtual mouse system was tested across a range of real-world environments to assess its performance in gesture recognition, voice command accuracy, and overall user interaction. This section presents a detailed evaluation based on five key performance metrics: gesture recognition accuracy, latency, voice command interpretation, switching responsiveness, and user satisfaction. Visual results in the form of graphs and tables are provided to support the findings. First, the gesture recognition module was analyzed under varying lighting conditions, such as well-lit, dim, backlit, and low-light settings. Figure 2 illustrates how the recognition accuracy varies across these conditions. As expected, the system achieved its highest accuracy (95%) in well-lit environments. However, the accuracy dropped noticeably in poor lighting, going as low as 65% in low-light conditions. This suggests the need for adaptive image pre-processing techniques in future versions. Table 1 complements the above graph by summarizing the gesture recognition performance, showing both accuracy and latency for specific gesture types across the same lighting conditions. The latency was the lowest for simple gestures, such as clicks and swipes, while more complex or custom gestures experienced increased response times. Next, we evaluated the latency involved in different gesture actions. Figure 3 displays the average response times for swipes, pinches, clicks, and user-defined custom gestures. Most actions had near real-time responsiveness, but latency increased slightly for custom gestures owing to added processing and user calibration. For the voice command module, accuracy was assessed under three noise scenarios: quiet, moderately noisy, and noisy environments. The system maintained 96% accuracy in quiet environments, dropping to 72% in noisy settings. This highlights a potential area for improvement, such as the incorporation of better noise filtering or enhanced voice models. The performance of the dynamic toggle switch was also tested. The average time taken to switch between gestures and voice modules was measured. On average, it took about 120 ms to switch from gesture to voice, and 110 ms for the reverse.

Table 1. Gesture recognition performance under various conditions.

Lighting Condition	Accuracy (%)	Latency (ms)
Well-lit	95	Swipe: 50
Dim	83	Pinch: 65
Backlight	78	Click: 45
Low-light	65	Custom: 90

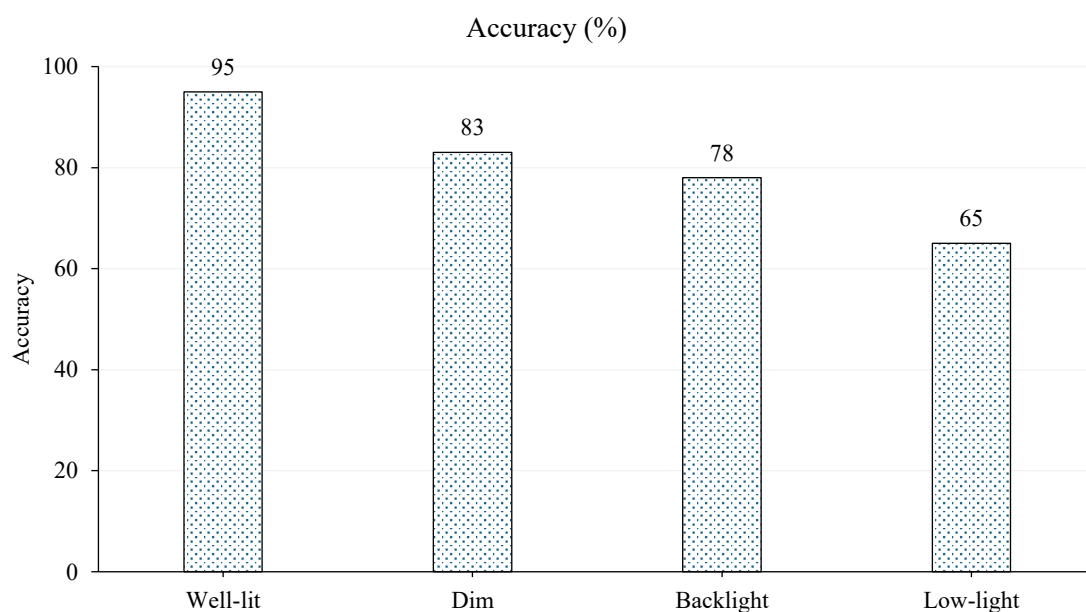


Figure 2. Gesture accuracy under varying lighting conditions.

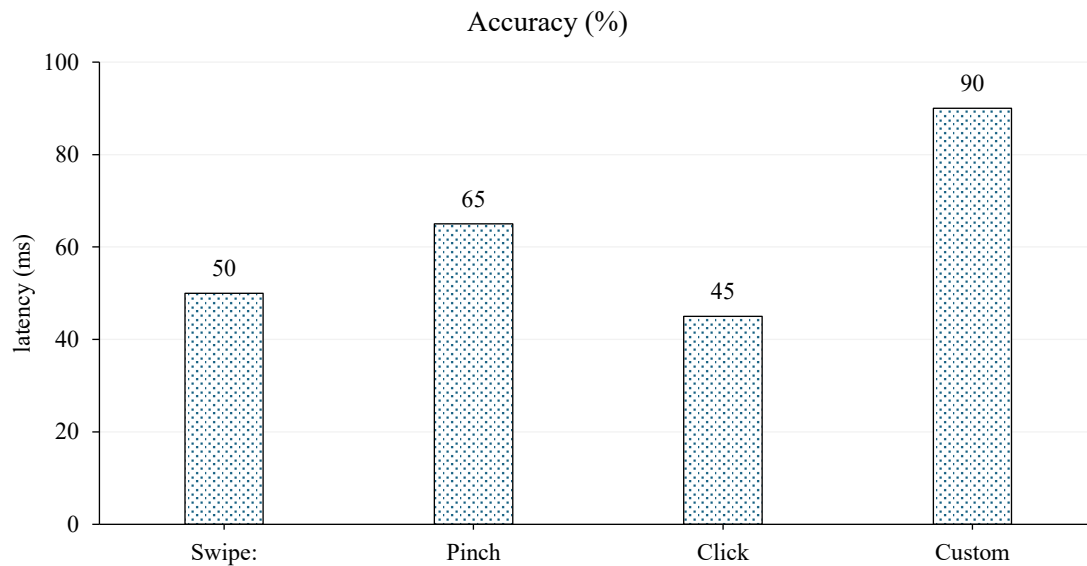


Figure 3. Gesture latency for different action types.

Table 2. Voice command and mode switching performance.

Test Scenario	Accuracy / Time (ms)
Quiet Environment	96%
Moderate Noise	88%
Noisy Environment	72%
Gesture → Voice Switch	120 ms
Voice → Gesture Switch	110 ms

These results confirm that the switch is responsive and allows users to shift modes with minimal delay. A summary of voice command accuracy and mode switching times is presented in Table 2. These findings further emphasize the system's capability to support multimodal interaction seamlessly. Finally, user satisfaction was measured through a survey in which participants rated the system based on usability, accessibility, comfort, and responsiveness. Users appreciated the system's flexibility and accessibility, especially those with mobility challenges. In summary, the system demonstrated strong performance in most conditions, with minor issues in low-light and noisy environments. The ability to switch input modes, retain session states, and support user-defined gestures makes it a promising solution for gesture- and voice-based human-computer interaction.

CONCLUSION AND FUTURE SCOPE

The gesture-controlled virtual mouse system represents a significant advancement in human-computer interaction, offering an intuitive, hands-free solution that enhances both accessibility and user convenience. By combining gesture recognition with voice command functionality, the system enables users to perform standard computing tasks—such as navigating applications, browsing, and controlling media through simple hand movements and natural speech. This multimodal interaction approach is particularly beneficial for individuals with physical challenges, such as carpal tunnel syndrome or arthritis, who may find traditional input devices difficult to use. During testing, the system showed strong performance in gesture recognition and voice command accuracy under standard conditions. The dynamic toggle feature provided seamless switching between gesture and voice modes, allowing users to adapt the interface according to their needs. However, certain limitations were observed, such as reduced gesture recognition accuracy in low-light settings and decreased voice recognition performance in noisy environments. These issues highlight the need for further improvements in environmental adaptability and signal processing. Looking ahead, future development may focus on enhancing

machine learning algorithms to improve recognition reliability and response time. Integrating additional sensors—such as depth cameras or infrared modules—could help maintain performance across various conditions. Expanding support for custom gestures and multilingual voice commands could also improve personalization and user inclusivity. Moreover, the system holds promise for broader applications in fields such as education, virtual and augmented reality, gaming, and assistive technologies. In conclusion, while the current system provides a solid foundation, continued innovation and refinement can further establish this technology as a practical and inclusive tool for modern human–computer interaction.

REFERENCES

1. Babu US, Raganna A, Vidyasagar KN, Bharati SH, Kumar G. Highly accurate static hand gesture recognition model using deep convolutional neural network for human machine interaction. 2022 IEEE 4th International Conference on Advances in Electronics, Computers and Communications (ICAECC), Bengaluru, India. 2022. p.1-6. doi:10.1109/ICAECC54045.2022.9716619.
2. Islam MZ, Hossain MS, Islam RU, Andersson K. Static hand gesture recognition using convolutional neural network with data augmentation. 8th International Conference on Informatics, Electronics & Vision (ICIEV), Washington, DC, USA. 2019. p.324-9. doi:10.1109/ICIEV.2019.8858563.
3. Li Y, Wang T, Khan A, Li L, Li C, Yang Y, et al. Hand gesture recognition and real-time game control based on a wearable band with 6-axis sensors. 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil. 2018. p.1-6. doi:10.1109/IJCNN.2018.8489743.
4. Shubham K, Satyam K, Kapisway P, Memani R, Anand A, Sathiya Kumar CS. IoT-based car driver assistance system using computer vision and machine learning. 2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT). 2024. p.1-6. doi:10.1109/AIIoT58432.2024.10574759.
5. Panda AK, Chakravarty R, Moulik S. Hand gesture recognition using flex sensor and machine learning algorithms. 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Langkawi, Malaysia. 2020. p.449-53. doi:10.1109/IECBES48179.2021.9398789.
6. Borah AR. Detecting background dynamic scenes using naïve Bayes classifier analysis compared to CNN analysis. 2023 4th International Conference on Smart Electronics and Communication (ICOSEC). 2023. p.1031-8. doi:10.1109/ICOSEC58147.2023.10276062.
7. Sun JH, Ji TT, Zhang SB, Yang JK, Ji GR. Research on the hand gesture recognition based on deep learning. 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE). 2018. p.1-4. doi:10.1109/ISAPE.2018.8634348.
8. John K, Sreekumar A, Mathew LR. Wearable gesture detection glove for mute people. 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), India. 2021. p.97-101. doi:10.1109/ACCESS51619.2021.9563296.
9. Roy RM, Sabu B, Nisha A, Aneesh RP. Voice controlled home automation system. 7th International Conference on Bio Signals, Images, and Instrumentation (ICBSII). 2021. p.1-6. doi:10.1109/ICBSII51839.2021.9445150.
10. Selvaganapathy G, Thirugnanam M, Sheela A. Voice assistant system. 2022 1st International Conference on Computational Science and Technology (ICCST), Chennai, India. 2022. p.1025-30. doi:10.1109/ICCST55948.2022.10040374.
11. Khan MR, Al Ahad A, Tania AA, Das T, Das B. Smart classroom automation: a fusion of AI with voice, gesture, and face recognition attendance system. 2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS). 2024. p.1-6. doi:10.1109/iCACCESS61735.2024.10499545.
12. Gowthamy J, Senthilselvi A, Kumar A, Aakash S, Sreedhar G. Enhanced AI voice assistance using machine learning and NLP. 2023 3rd International Conference on Smart Technologies, Communication and Robotics (STCR). 2023. p.1-5. doi:10.1109/STCR59085.2023.10396893.
13. Kasar M, Kavimandan P, Suryawanshi T, Abbad S. AI-based real-time hand gesture-controlled virtual mouse. Aust J Electr Electron Eng. 2024;21(3):258-67. doi:10.1080/1448837X.2024.2313818.

14. Raja M, Nagaraj P, Sathwik P, Khan KMA, Kumar NM, Prasad US. Voice assistant and gesture controlled virtual mouse using deep learning technique. 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). 2023. p.156-61. doi:10.1109/ICSCDS56580.2023.10104619.
15. Subhash S, Srivatsa PN, Siddesh S, Ullas A, Santhosh B. Artificial intelligence-based voice assistant. 2020 4th World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). 2020. p.593-6. doi:10.1109/WorldS450073.2020.9210344.
16. Jain P, Bhandari LS, Sampath GM, Deeksha M. Gesture controlled virtual mouse with voice assistant integration. 2025 International Conference on Artificial Intelligence and Machine Vision (AIMV). 2025. p.1-8. doi:10.1109/AIMV66517.2025.11203344.
17. Khan I, Kanchan V, Bharambe S. Gesture controlled virtual mouse with voice assistant. Int Res J Multidiscip Scope. 2024;5(1):26-35. doi:10.47857/irjms.2024.v05i01.0127. Note: the full article identifies five authors, adding Ayush Thada and Rohini Patil.
18. Sindhu RJ, Satija R, Sarkar S, Malarvel M, Kathiravan M. A hybrid approach to gesture recognition with biometric security. 2022 7th International Conference on Communication and Electronics Systems (ICCES). 2022. p.697-701. doi:10.1109/ICCES54183.2022.9835887.
19. Borah AR, Nischith TN, Gupta S. Improved learning based on GenAI. 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India. 2024. p.1527-32. doi:10.1109/IDCIoT59759.2024.10467943.
20. Kassim M, Zain JM, Yaacob WF, Ruhaiyem NI, Yusoff M, Nisirat MA, editors. Soft computing in data science: 8th International Conference, SCDS 2025, Penang, Malaysia, October 8-9, 2025: revised selected papers. Singapore: Springer Nature; 2026. doi:10.1007/978-981-95-8368-3.
21. Borah AR, Krishna BVS, H SM, Gouda DN, Poojary P. The soul safety. 3rd International Conference on Communication, Computing and Industry 4.0. 2022. p.1-6. doi:10.1109/C2I456876.2022.10051384.
22. Bhargavi MS, Prabhakar B. Deep learning approaches for early detection of Alzheimer's disease using MRI neuroimaging. International Conference on Connected Systems & Intelligence (CSI). 2022. p.1-6. doi:10.1109/CSI54720.2022.9924058.
23. Prathyakshini P, Prathwini. Hand gesture controlled video player application. 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA). 2023. p.67-72. doi:10.1109/ICECA58529.2023.10395578.
24. Xie J, Zhang Y, Xu Z, Gao Y, Bai J, Zeng J, et al. One-handed wonders: a remote control method based on hand gesture for mobile manipulator. 2024 9th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM), Tokyo, Japan. 2024. p.686-91. doi:10.1109/ICARM62033.2024.10715857.
25. Cohen CJ, Beach G, Foulk G. A basic hand gesture control system for PC applications. Proceedings of the 30th Applied Image Pattern Recognition Workshop (AIPR 2001): Analysis and Understanding of Time Varying Imagery, Washington, DC, USA. 2001. p.74-9. doi:10.1109/AIPR.2001.991206.