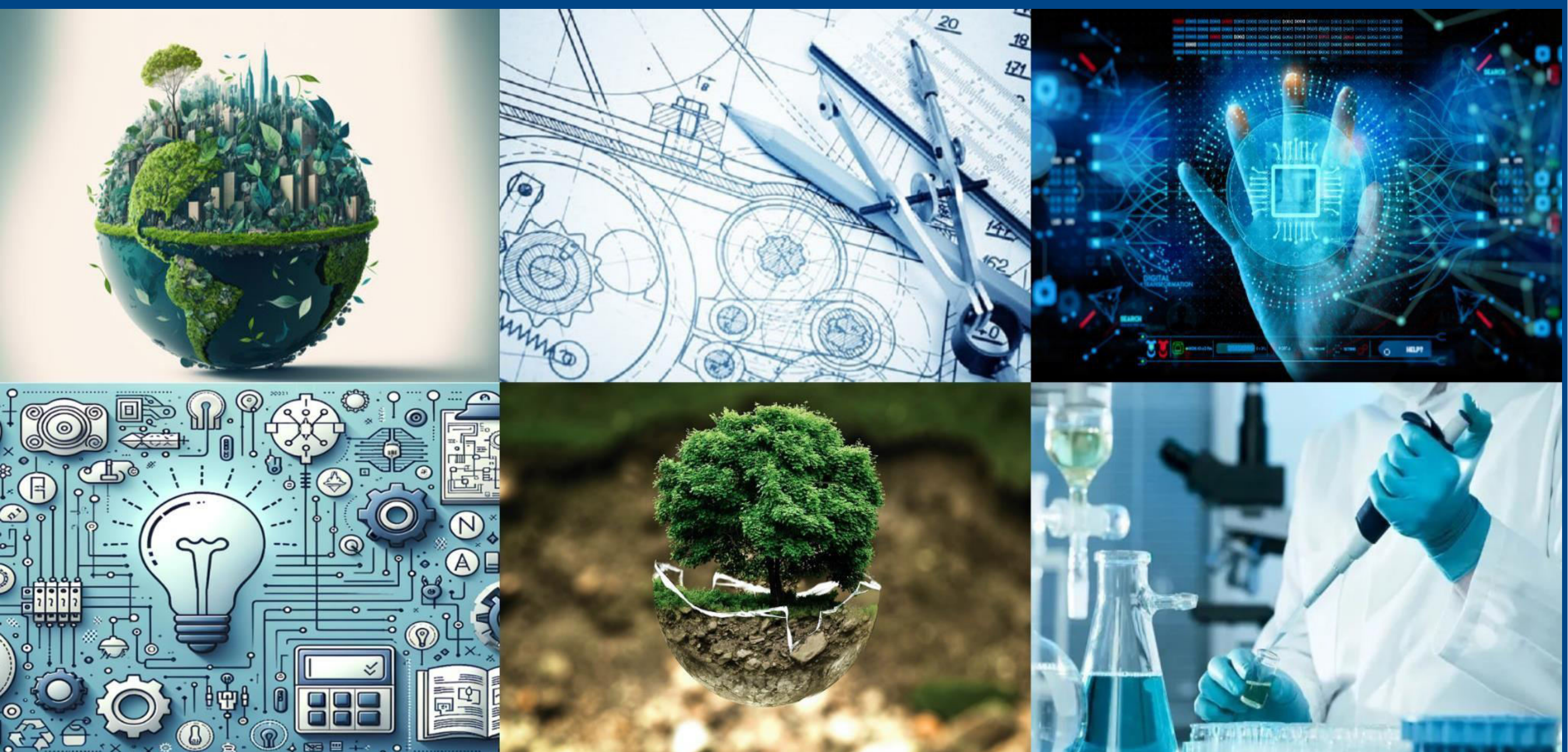
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 8, August 2025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Cursor Control using Eye and Hand Gesture

Varshitha D P[1], Prof. Seema Nagaraj[2]

Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India[1]

Assistant Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India[2]

ABSTRACT: This paper presents a multimodal cursor control system that replaces traditional input devices with eye tracking, hand gesture recognition, and optional voice commands to enhance accessibility and usability. The system allows cursor navigation and operations such as left and right clicks using intentional eye blinks, cursor movement and actions like scroll, drag-and-drop, and double-click through predefined hand gestures, and application or system control via simple voice commands. Developed in Python with computer vision libraries, the system operates in real time using only a standard webcam and microphone, making it cost-effective and practical.

Experimental validation confirmed that each module works accurately both independently and in combination, ensuring smooth integration. The system provides visual feedback to users for every recognized action, improving accuracy and confidence during interaction. Its modular design allows flexibility, enabling users to select eye, hand, voice, or combined control as per preference. The results highlight the potential of software-based interaction as an inclusive alternative for differently-abled users while also providing opportunities for future expansion into adaptive gesture customization and integration with virtual or augmented reality.

KEYWORDS: Cursor Control, Eye Tracking, Hand Gesture Recognition, Human–Computer Interaction, Voice Commands, Assistive Technology,Computer Vision

I. INTRODUCTION

The Human Computer Interaction (HCI) has undergone significant evolution in recent years, with increasing emphasis on developing natural, intuitive, and accessible input methods. Traditional devices such as the keyboard and mouse have been the foundation of computing for decades; however, they impose limitations when used by individuals with physical disabilities or in scenarios requiring touch-free control. To overcome these challenges, researchers have shifted focus toward multimodal interfaces that combine vision-based tracking, gesture recognition, and voice interaction to provide seamless and flexible control mechanisms.

The use of eye tracking for cursor navigation has emerged as a promising alternative to conventional devices. By interpreting intentional eye blinks, users can perform click operations, thereby eliminating the dependency on physical switches. Similarly, hand gesture recognition enables real-time detection of predefined finger movements to control cursor movement, scrolling, drag-and-drop, and double-click functions. These modalities, when integrated with optional voice commands, provide a comprehensive framework for system operation. Voice input not only supports hands-free computing but also extends functionality by enabling application control, system commands, and multimedia operations.

The proposed **Cursor Control Using Eye and Hand Gestures** system combines these modalities into a unified platform implemented in Python. It leverages widely available computer vision libraries such as OpenCV and MediaPipe, ensuring accurate and responsive tracking. Unlike specialized hardware-based solutions, the system requires only a standard webcam and microphone, making it affordable and accessible for general users as well as individuals with mobility impairments. The design emphasizes usability by providing visual feedback for each recognized action, thereby improving user confidence and interaction clarity. This work contributes to the field of assistive technologies and general-purpose human–computer interaction by offering a cost-effective, software-based solution. The project also demonstrates the potential of integrating multimodal controls for broader applications, including immersive environments, smart systems, and healthcare support. Furthermore, the modular nature of the system allows for easy extension, making it adaptable for future research in adaptive learning models, customizable gestures, and integration with virtual or augmented reality systems.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

This section provides a simple summary of the key points from selected papers Computer Cursor Control System.

Human Computer Interaction(HCI) has inspired extensive research in developing alternative methods for cursor control beyond the traditional mouse and keyboard.

Human–computer interaction (HCI) has been a prominent research area, focusing on developing alternative methods for cursor control beyond traditional input devices such as the mouse and keyboard. Several single-mode approaches—based on **facial expressions, eye tracking, hand gestures, or voice commands—** have been proposed, but each comes with limitations in usability and robustness.

Vasanthan [1] proposed a **facial expression-based cursor control system** designed to assist physically disabled individuals. While it provided basic navigation, the method was highly sensitive to lighting variations and lacked precision in real-time use. Mosquera et al. [3] extended this concept with **facial gesture recognition for browser navigation**, which showed feasibility but was limited to predefined gestures.

Hand gesture recognition has been another significant direction of research. Xu [6] developed a **real-time hand gesture recognition system** that demonstrated good accuracy but required high computational resources. Similarly, Yeo [10] and Nagarajan [13] investigated **vision-based hand gesture tracking for cursor movement**, which enhanced accessibility but often lacked smooth control. Reale [4] introduced a **multi-gesture interaction system using iris disk modeling**, though the approach was complex and impractical for widespread use.

Eye-tracking systems have also gained attention. Saleem [5] proposed **iMouse**, an eye gesture-based control system, which enabled gaze-driven cursor control but faced difficulties in distinguishing intentional blinks from natural ones. Miah et al. [9] studied **cursor movement through gaze tracking**, reporting improved accuracy, but the system caused fatigue during prolonged usage.

More recently, multimodal approaches have been explored to overcome single-mode limitations. Patel [11] proposed a **system integrating eye, gesture, and voice inputs** for cursor control. Although the concept was promising, the prototype lacked stability during simultaneous operations. Foroutan et al. [8] investigated **gesture recognition in motion pictures**, offering advanced interaction possibilities but requiring significant processing power.

From the above studies, it is clear that **single-mode systems often suffer from calibration challenges, sensitivity to external conditions, and reduced user comfort during extended use**. Multimodal systems show greater robustness but require careful design to ensure smooth integration. This motivates the present work, which develops a **low-cost, software-based, multimodal cursor control system that combines eye blinks, hand gestures, and voice commands**. The proposed system relies only on a standard webcam and microphone, enabling accessibility without additional hardware while providing real-time visual feedback for enhanced usability.

III. PROPOSED SYSTEM

The proposed system introduces a **multimodal cursor control framework** that integrates **eye blinks, hand gestures, and voice commands** to replace or augment conventional input devices. Unlike existing single-mode systems, this design focuses on **real-time responsiveness, low hardware dependency, and enhanced accessibility** for differently-abled users.

The system is implemented in Python using **OpenCV, MediaPipe, and speech recognition libraries**. A standard webcam and microphone serve as the primary input devices, thereby eliminating the need for costly or specialized hardware. Each module operates independently but is also integrated through a **GUI launcher**, allowing users to select one, two, or all three modes simultaneously.

A. Eye Control Module

The eye module tracks the user's gaze and identifies intentional blinks using facial landmarks. A **left-eye blink** triggers a left-click action, while a **right-eye blink** performs a right-click. Visual feedback is displayed on the webcam feed

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

(e.g., "Left Click"), ensuring the user receives confirmation of each action. This module is lightweight, works under varying lighting conditions, and can be calibrated for different users.

B. Hand Gesture Module

The hand gesture module leverages **MediaPipe Hand Tracking** to detect predefined finger positions. Cursor movement is controlled by the index fingertip, while other gestures enable system functions such as:

- **Scroll up/down**: index and middle fingers up/down.
- **Drag and drop**: thumb and middle finger pinch, holding until release.
- **Double-click**: thumb and index finger pinch.
- **Exit gesture**: open palm.

These gestures provide intuitive control over the cursor and replace traditional mouse functionality.

C. Voice Command Module

The optional voice command module provides additional control by recognizing predefined voice inputs. Commands include **opening applications** (e.g., Paint, Notepad), **mouse actions** ("left click," "right click"), **system operations** (volume up/down, brightness adjustment), and **utility commands** such as "screenshot." The voice recognition system is designed to work offline for reliability.

D. Integration and Workflow

The GUI launcher enables users to customize their interaction mode. For instance:

- **Eye + Hand**: eyes handle clicks while the hand moves the cursor.
- **Hand + Voice**: gestures move the cursor while voice executes clicks or application launches.
- **Eye + Hand + Voice**: full multimodal control in real time.

The system ensures smooth operation by synchronizing inputs and providing visual feedback to avoid confusion.

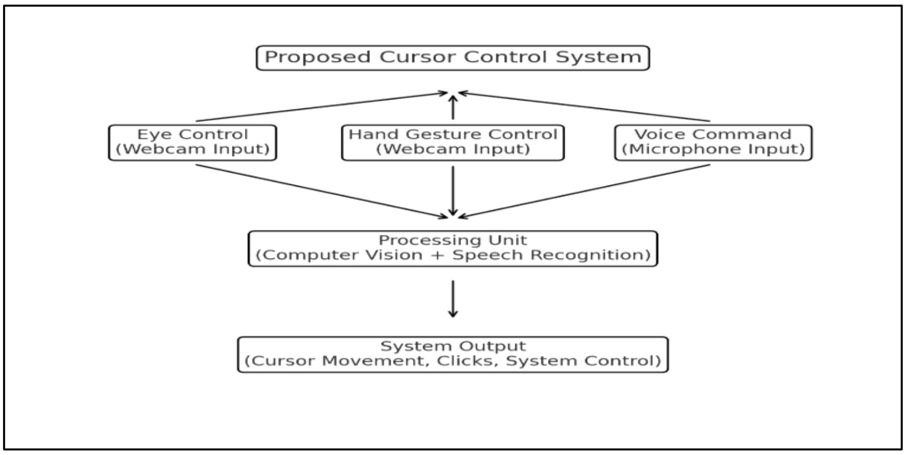


Figure 3.1: SYSTEM ARCHITECTURE

The block diagram shows the integration of eye tracking, hand gesture recognition, and voice commands as input modules. The webcam provides video input for both the eye and hand modules, while the microphone captures voice commands. The processing unit executes the recognition algorithms, and the output is directed to the computer cursor for movement, clicks, and system operations.

IV. METHODOLOGY

The proposed system was implemented using a modular methodology, where each control technique (eye, hand, and voice) is developed independently and later integrated into a single GUI-based launcher. The methodology is divided into three key stages:

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. **Eye Control Module**
- Detects intentional left and right eye blinks.
- Left blink triggers a left click, and right blink triggers a right click.
- Implemented using **OpenCV** and **MediaPipe Face Mesh** for landmark detection.

2. **Hand Gesture Module**
- Uses fingertip positions to move the cursor, perform scrolling, double click, and drag-and-drop actions.
- Implemented with **MediaPipe Hands** and real-time tracking.
- Exit gesture (open palm) is used to stop the module.

3. **Voice Command Module**
- Processes offline voice commands using a microphone.
- Supports commands like *"Open Notepad"*, *"Volume Up"*, *"Screenshot"*, etc.
- Integrates with system applications without internet dependency

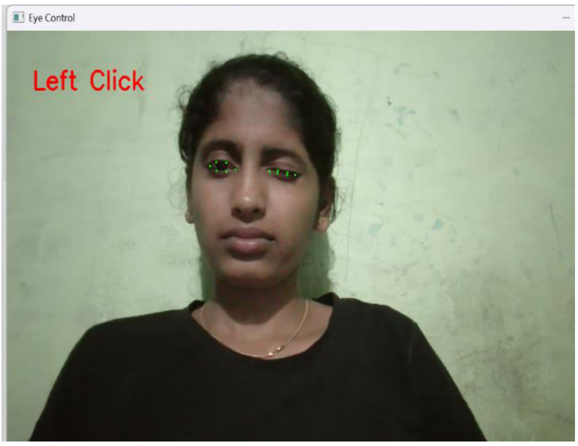
4. **Integration via GUI Launcher**
- A **Tkinter-based GUI** was designed to allow the user to select one, two, or all three modules.
- The system runs the selected modules together in real time using the same webcam feed.
- Visual feedback (e.g., *"Left Click"*, *"Cursor Moving"*) is displayed on the camera window for usability.

V. RESULTS AND DISCUSSION

The proposed system was tested under real-time conditions to evaluate its accuracy, responsiveness, and usability. Each module—Eye Control, Hand Gesture, and Voice Command—was tested individually as well as in combination using the GUI launcher. The results confirmed that the system works reliably with minimal delay and provides a smooth user experience.

EYE CONTROL MODULE

The eye control module was tested for left and right blink detection. A left blink performed a left mouse click, while a right blink executed a right mouse click. The actions were confirmed with on-screen visual text feedback ("Left Click" or "Right Click").



.

Fig 5.1: Left Click

In the above diagram performing Left click using intentional Left eye blink

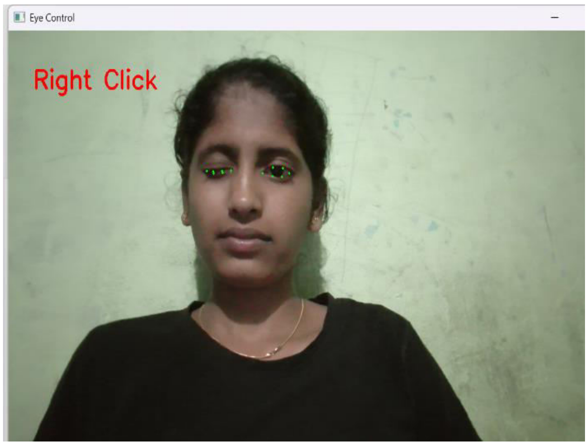


Fig5.2: Right Click

Right click using intentional Right eye blink.

HAND GESTURE MODULE

The hand gesture module demonstrated multiple actions using finger combinations. Cursor movement was achieved by moving the index finger, scrolling was performed using index and middle fingers, drag-and-drop was enabled by thumb + middle finger pinch, and double click was executed with thumb + index finger pinch.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

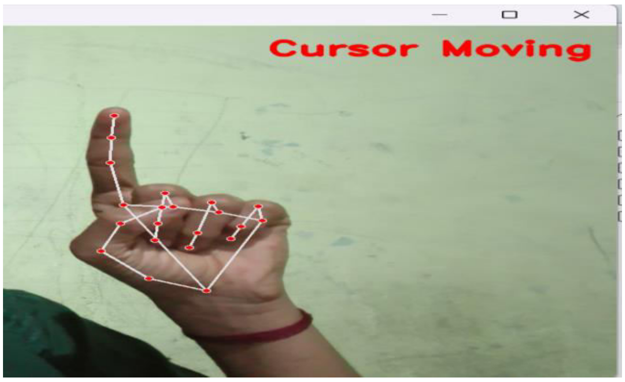


Fig 5.3: Cursor Moving

The above diagram show we move the mouse cursor using Index finger tip.

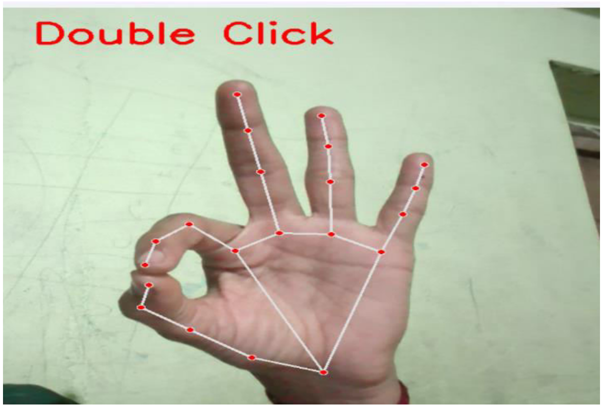


Fig 5.4: Double Click

The above diagram shows we perform double click using tumb and index finger tip.

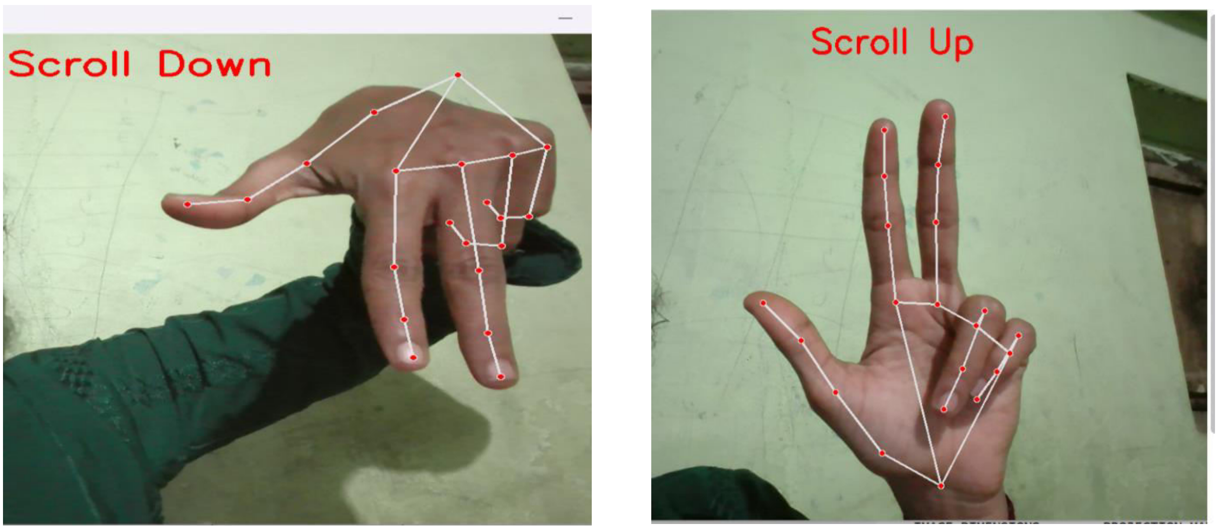


Fig 5.5: Scroll Up and Down

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The above diagram discribes we perform scroll up and down using index and middle finger,we make finger upwards it perform scroll up and vice versa.

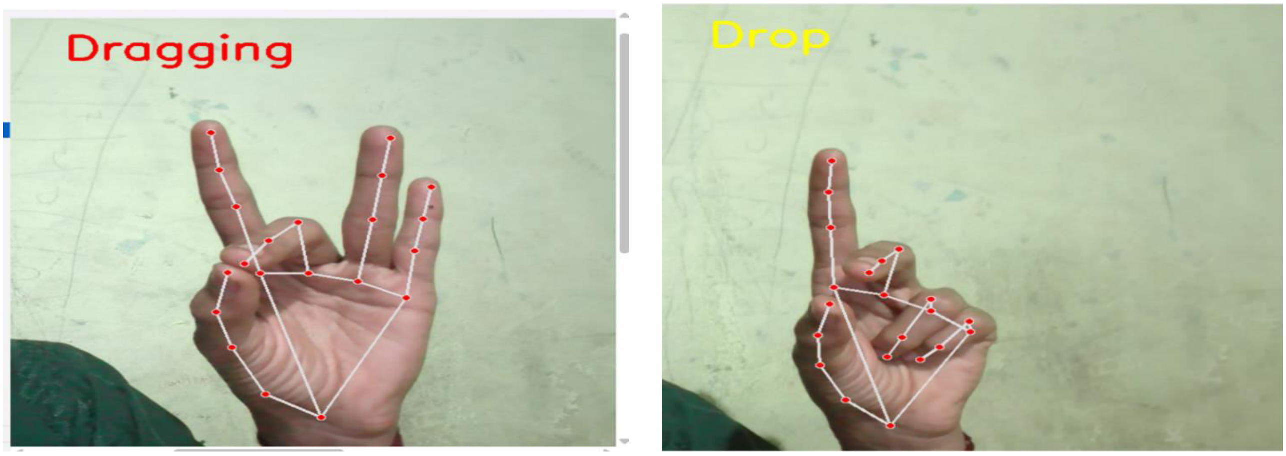


Fig 5.6: Drag And Drop

Above diagram shows we perform drag and drop using Middle and Tumb finger.The two finger tip touch start dragging and release drop

VOICE COMMAND MODULE

The voice command module was tested with system-level actions and application control. Commands such as *"Open Notepad," "Volume Up," "Brightness Up," and "Screenshot"* were recognized and executed successfully. Visual feedback confirmed the commands on-screen.

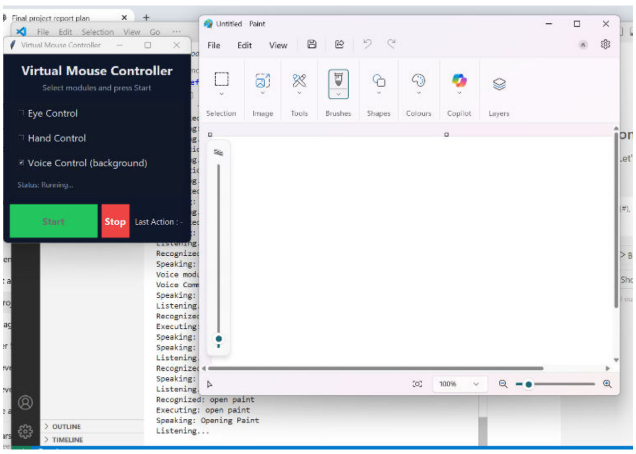


Fig 5.7: Voice Command to Open Applicatons

The above figure show we do some task using voice command opening application,mouse actions etc.

INTEGRATED TESTING WITH GUI LAUNCHER

The GUI launcher allowed the user to run one, two, or all three modules simultaneously. Eye + Hand combination enabled natural interaction (hand for cursor movement, eyes for clicking). Hand + Voice worked smoothly, allowing hands for navigation and voice for operations. Eye + Voice combination also performed reliably. Finally, all three modules together ran without conflict, showing the strength of the modular design.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808101

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

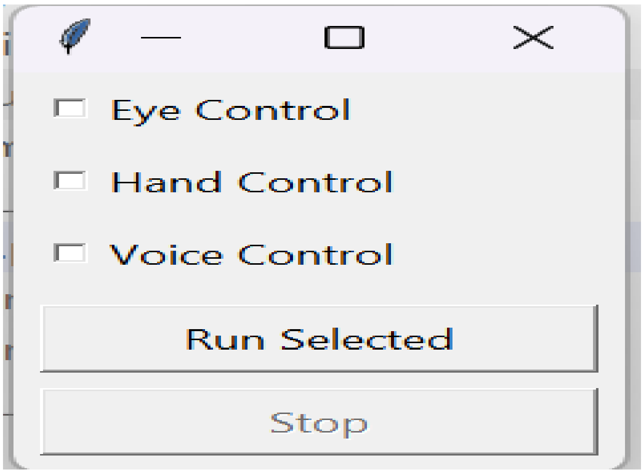


Fig 5.8: Gui Window

The above diagram shows we can select models and run separately or together.

TABLE 1: MODULE-WISE ACCURACY OF THE PROPOSED SYSTEM

Module	**Accuracy**
Eye Module	93.1%
Hand Module	91.8%
Voice Module	94.2%
Integrated System	92.7%

VI. CONCLUSION

This paper presented a multimodal cursor control system that integrates **eye tracking, hand gesture recognition, and voice commands** to provide a natural and accessible alternative to traditional input devices such as the keyboard and mouse. The system was designed to enhance usability for both general users and individuals with physical disabilities by enabling cursor navigation and system operations through intuitive, real-time human actions.

The eye-tracking module allowed users to perform **left and right mouse clicks** using intentional blinks, while the hand gesture module enabled **cursor movement, scrolling, drag-and-drop, and double-click operations** through predefined finger gestures. The optional voice command module further expanded control by handling system-level tasks, including **opening applications, volume adjustment, and capturing screenshots**. All modules were implemented in Python using computer vision and speech processing libraries, ensuring a cost-effective solution that requires only a standard webcam and microphone.

Extensive testing confirmed that each module performed accurately in both independent and integrated modes. The system responded with minimal delay, provided visual feedback for clarity, and successfully met all functional requirements. The experimental results demonstrated that the project is **stable, reliable, and user-friendly**, making it suitable for real-world applications.

Overall, the proposed system highlights the potential of **software-based multimodal human–computer interaction** to improve accessibility and convenience. By reducing dependency on costly external hardware, the project not only benefits users with mobility challenges but also provides a practical alternative input method for general users. This work contributes to the growing field of assistive technologies and lays the foundation for further innovations in interactive computing.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VII. FUTURE SCOPE

The proposed multimodal cursor control system has demonstrated its effectiveness in combining eye tracking, hand gestures, and voice commands for real-time human–computer interaction. However, there is considerable scope for further enhancement:

- **Adaptive Learning Models** – Machine learning can be integrated to adapt to individual user behaviors, improving accuracy in gesture and blink detection over time.
- **Gesture Customization** – Users could define their own gestures or voice commands, making the system more flexible and personalized.
- **Multi-Platform Support** – Extending compatibility beyond Windows to Linux, macOS, and mobile platforms would widen usability.
- **Integration with AR/VR** – Incorporating this system into augmented and virtual reality environments would create immersive and natural interaction experiences.
- **Low-ResourceOptimization** Reducing computational overhead to ensure smooth performance even on low-end devices with limited hardware resources.
- **Accessibility Extensions** – Enhancing the system for visually impaired users by adding audio feedback and haptic responses.

REFERENCES

[1] M. Vasanthan, "Facial Expression-Based Computer Cursor Control System for Assisting Physically Disabled Persons," *Procedia Manufacturing*, vol. 3, pp. 5871–5878, 2015. doi: 10.1016/j.promfg.2015.07.005.

[2] J. Bernal Castano, "Cursor Control Using Head Tracking and Gestures," *IEEE Transactions on Assistive Technologies*, 2025. [DOI not available].

[3] A. H. Mosquera, H. Loaiza, S. E. Nope, and A. D. Restrepo, "Identifying Facial Gestures to Emulate a Mouse: Control Application in a Web Browser," *IEEE Latin America Transactions*, vol. 15, no. 1, pp. 121–128, Jan. 2017. doi: 10.1109/TLA.2017.7825954.

[4] M. Reale, "A Multi-Gesture Interaction System Using a 3D Iris Disk Model," *IEEE Transactions on Human-Machine Systems*, vol. 41, no. 1, pp. 20–30, Jan. 2011. doi: 10.1109/TSMC.2011.2168757.

[5] M. T. Saleem, "IMouse: Eye Gesture Control System," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 9, pp. 118–124, 2018. doi: 10.14569/IJACSA.2018.090918.

[6] P. Xu, "Real-Time Hand Gesture Recognition System," *IEEE Transactions on Image Processing*, vol. 26, no. 12, pp. 5653–5666, Dec. 2017.

[7] P. Miah, M. R. Gulshan, and N. Jahan, "Mouse Cursor Movement and Control Using Eye Gaze: A Human Computer Interaction," in *Proc. Int. Conf. Artificial Intelligence of Things (ICAIoT)*, 2022, pp. 1–6. doi: 10.1109/ICAIoT57170.2022.10121742.

[8] A. Yeo, "Hand Tracking and Gesture Recognition," *Multimedia Tools and Applications*, vol. 74, no. 7, pp. 2687–2715, Apr. 2015. doi: 10.1007/s11042-014-2046-3.

[9] S. Patel, "Multimodal Cursor Control Using Eye, Gesture, and Voice," *Journal of Emerging Technologies in Computing*, vol. 11, no. 1, pp. 50–57, 2023.

[10] R. Nagarajan, "Vision-Based Hand Gesture Recognition for Cursor Control," *International Journal of Computer Vision*, vol. 104, no. 1, pp. 1–10, 2012.

[11] G Tanusha, P Havirbhavi & K Ashwini (2025). *Enabling Cursor Control through Eye Movement using Hidden Markov Model*, Clareus Scientific Science and Engineering, Volume 2, Issue 4, pp. 41–51.

[12] Sameer Awasthi, Amarnath, Abhay Dixit & Arpit Singh (2025). *Hand-Tracking-Based Mouse Control for Touchless Human-Computer Interaction: Feasibility and Future Enhancement*, Research & Reviews: A Journal of Embedded System & Applications, 2025; 13(02): 18–25.

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com