iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.61178

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Hand Glide: Gesture-Controlled Virtual Mouse withVoice Assistant

Kavyasree K[1], Sai Shivani K[2], Sai Balaji T[3], Sahana Yele[4], Vishal B[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*Student, Department of Artificial Intelligence and Machine Learning, Ballari Institute of Technology and Management, India*

Abstract: ***Hand Glide presents a unified system that uses the MediaPipe library for accurate hand landmark recognition and categorization, enabling smooth integration of gesture-based control (GBC) and voice-controlled assistant (VCA) features. The system allows users to effortlessly translate hand movements into actions such as mouse manipulation, clicking, scrolling, and system parameter adjustments by providing distinct classes for hand gesture recognition (HandRecog) and execution (Controller), as well as a main class orchestrating camera input and gesture control. In addition, Hand Glide also has VCA functionality, which uses speech recognition and synthesis libraries and natural language processing (NLP) techniques. With these tools, Hand Glide can reply to greetings, perform tasks like Wikipedia searches and website openings, and converse with users through the GPT-2 language model. By merging these features, Hand Glide providesusers with an interactive and intuitive means of computer control, with broad applications in user interfaces, accessibility solutions, and immersive experiences.***
Keywords: GCB (gesture-based control), VCA (voice-controlled assistant), NLP (natural language processing), GPT-2 (Generative Pre-Trained Transformer 2).

I. INTRODUCTION

An inventive solution is offered by Hand Glide, which combines voice-controlled assistant (VCA) and gesture-based control (GBC) features fluidly. (1) Gesture-based control has become a popular interface technique that provides consumers with a naturalapproach to interacting with their gadgets. Yet, the accuracy and adaptability of current systems are frequently lacking. By utilising the powerful capabilities of the MediaPipe library for accurate hand landmark detection and classification, Hand Glide overcomes these drawbacks. Because of its cutting-edge technology, users can easily accomplish tasks like adjusting system parameters, clicking, scrolling, and manipulating the mouse. It also ensures precise interpretation of hand movements.

Furthermore, Hand Glide has three (3) voice-controlled assistant (VCA) features, giving consumers an extra degree of engagement. Voice commands improve accessibility and convenience by providing a hands-free option. Hand Glide allows users to do activities ranging from basic commands to sophisticated queries by utilising natural language processing (NLP) techniques and speech recognition and synthesis libraries. For example, users can visit webpages, search Wikipedia, and have conversations with others using the GPT-2 language model. Hand Glide was developed in response to the growing demand for more intuitive and efficient computer interface methods. Users with disabilities or those looking for a more natural interface may find it difficult to use traditional input devices like keyboards and mice. Furthermore, as technology advances, there is a need for interfaces that can accommodate a range of user contexts and preferences. To meet these demands, Hand Glide provides a flexible and easy-to-use interface that combines the advantages of voice interaction and gesture-based control. Hand Glide enables people to interact with their computers more naturally and fluidly, opening up new possibilities for improving user experiences and accessibility across multiple areas.

A. GPT-2 Language Model

The code improves the voice-controlled assistant's capacity to interact sensibly and contextually with users by utilising the GPT-2 language model to generate responses to unrecognised instructions.

II. LITERATURE REVIEW

1) Paper [1]: Kavitha, Janasruthi, Lokitha, and Tharani, "Hand Gesture-Controlled Virtual Mouse using Artificial Intelligence." The literature review investigates several methods for hand gesture detection and control systems, with a focus on applications such as screen control, robotic manipulation, and computer interaction. Wireless control is achieved using Arduino and Python, gesture recognition is done with MEMS acceleration sensors, and hand movements are detected with ultrasonic sensors.

©IJRASET: All Rights are Reserved

Studies on gesture-based interaction with technology, such as robots and computers, emphasise the possibility of natural and intuitive interfaces. In addition, the survey looks into gesture detection algorithms based on accelerometers, gyroscopes, and vision systems. Overall, the study illustrates the feasibility and efficacy of hand gesture control for a variety of applications, paving the way for improved human-computer interaction and robotic manipulation. Hand-Gesture-Controlled Virtual Mouse with Artificial Intelligence.

2) Paper [2]: ISmail Khan, Vidhyut Kanchan, Sakshi Bharambe, Ayush Thada and Rohini Patil, "Gesture-Controlled Virtual Mouse with Voice Assistant" This literature survey focuses on the Gesture-Controlled Virtual Mouse with Voice Assistant (GCVA) technology, which combines hand gesture recognition and voice instructions to operate devices with the goal of replacing traditional input methods. GCVA, which uses OpenCV and Media-Pipe for gesture recognition and AI/NLP for voice commands, allows users to perform mouse operations and execute commands without the use of physical peripherals. Tests found a 98% accuracy rate for gestures, ensuring consistent interaction. Additionally, response times for voice commands were assessed, indicating the system's efficiency in carrying out user requests. When compared to existing systems, GCVA demonstrated higher accuracy and functionality, highlighting its potential to transform human-computer interaction. By integrating natural gestures andvoice control, GCVA provides a unified user experience, removing the need for physical mouse and keyboard installations. This breakthrough aligns with the emerging landscape of AI-driven technologies and demonstrates the viability of implementing gesture and voice-based interfaces in common computer operations.

3) Paper [3]: E. Sankar CHAVALI "Virtual Mouse Using Hand Gesture" The literature overview investigates current advances in gesture detection and hand tracking, with a specific emphasis on the creation of a virtual mouse utilising hand gestures. It addresses the opportunities and challenges presented by these technologies, emphasising their potential to reduce connections between people and technology, particularly in light of the COVID-19 pandemic. The survey covers several approaches and frameworks used to create gesture-based interfaces, such as OpenCV and Media Pipe, with a focus on the integration of machine learning algorithms for accurate hand gesture recognition. It also discusses the practical implementation of the suggested system, its benefits in improving user-computer interface, and opportunities for future enhancement, such as refining fingertip detection algorithms for improved performance. . Overall, the review gives a thorough overview of cutting-edge techniques in gesture-basedvirtual mouse control and its implications for human-computer interaction.

4) Paper [4]: Sahil, Sameer, "Virtual Mouse Using Hand Gestures The literature review discusses earlier work in virtual mouse systems, such as glove-based recognition, coloured marker-based systems, and hand gesture detection using cameras. Several studies have addressed issues such as sluggish processing times, complex gestures, and system requirements. The proposed system extends these earlier studies by providing a device-free, real-time hand gesture recognition solution written in Python utilising OpenCV and MediaPipe.

5) Paper [5]: Mr. E. Sankar, B.Nitish Bharadwaj, and A.V. Vignesh, "Virtual Mouse Using Hand Gesture" This literature survey highlights The proposed virtual mouse system provides a substantial development in Human-Computer Interaction (HCI) by combining gesture recognition, computer vision, and speech recognition technology. Gesture recognition has long been a focus in computer vision, providing expressive interaction methods. Computer vision algorithms allow the system to understand hand gestures filmed by a webcam, eliminating the requirement for traditional mouse devices. Speech recognition improves accessibility by allowing users to give voice commands for mouse functions. User-centred design concepts are critical for ensuring that the system meets the different needs of its users. . The system's hands-free operation also overcomes hygienic problems, making it appropriate for use in situations such as the COVID-19 epidemic. Overall, the virtual mouse system demonstrates forward-thinking HCI paradigms that promise intuitive, accessible, and safe interaction experiences.

III. MOTIVATION

Developing a hand-gestured virtual mouse with voice control is an exciting and innovative project that requires problem solving in gesture and speech detection, sharpening abilities in computer vision and human-computer interface, and contributing to future technical developments.

This study has real-world applications in accessibility and gaming, which could improve user experiences across multiple domains. Key phases include developing natural hand movements for specific actions, developing a hierarchical structure for voice commands, and combining these components into a unified user interface.

Participating in the community and staying current on industry trends in natural user interfaces and upcoming technologies will increase the project's influence and relevance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

The acquired skills are closely aligned with current industry demands in human-computer interface, artificial intelligence, and software development, providing potential for future career advancement in these quickly changing fields. Overall, this project combines technological obstacles, creative expression, and societal influence, making it both academically interesting and personally rewarding.

IV.OBJECTIVE

1) *Innovative Interaction:* Provide a system that allows users to communicate with computers and other devices in an inventive and simple way by combining voice commands and hand gestures to control a virtual mouse.
2) *Enhanced Accessibility:* Provide hands-free or gesture-based control to users with mobility disabilities or those who prefer alternative input techniques.
3) *Seamless Integration:* Combine gesture detection and voice control technologies to produce a unified user experience that prioritises smooth interaction and accurate command execution.
4) *User-Centric Design:* Prioritize natural and efficient interactions, minimum learning curves, and customizable settings to meet the needs of a wide range of users.

Contribute to the improvement of human-computer interaction (HCI) by using computer vision, speech recognition, machine learning, and software engineering approaches to create a dependable and responsive control system.

V. METHODOLOGY

1) Understanding Voice Control with Text-to-Speech (TTS)

Use the pyttsx3 library to convert text to speech. Initialize the TTS engine and define a function (say) for speaking out text.
Voice Recognition: Use the speech_recognition library to recognize voice input. Define a function (listen_continuous) that listens for voice commands using a microphone and the Google Speech Recognition API.

2) The GPT-2 language model

Tokenization: Set up the GPT-2 tokenizer (GPT2Tokenizer) to encode text inputs for the GPT-2 model.
Generation: To create text responses depending on user inputs, utilise the pre-trained GPT-2 model (GPT2LMHeadModel).

3) Integrating Voice Commands

Define a function (process_command) to handle voice commands, including greetings, Wikipedia searches, website openings, app launches, and system exits.
Wikipedia Search: Use the Wikipedia library to search for and obtain information from Wikipedia based on user input.

4) Understanding Gesture Control and Hand Gesture Recognition:

Use the cv2, mediapipe, and pyautogui libraries for hand gesture recognition and system control.
Mediapipe: Use Mediapipe's mp_hands module to recognize and track hand landmarks in real-time video frames.
Gesture Encoding: Create enums (Gest and HLabel) that map hand movements to binary values and handle multi-handedness.

5) Gesture Recognition Logic

Use the HandRecog class to translate Mediapipe landmarks into identifiable motions. This covers computing distances, ratios, and dealing with noise-induced perturbations.
Controller Class: Use the Controller class to implement logic that executes commands based on observed gestures. Create functions for mouse control, system settings change, scrolling, and program start.

6) Integrated Voice and Gesture Control

Main class (gesture controller): Integrate voice commands and gesture control functionality into the GestureController class. Begin continuous listening for voice commands and handle them with the process_command function. Capture video frames, detect hand motions, and use gesture recognition to control system functions.

©IJRASET: All Rights are Reserved

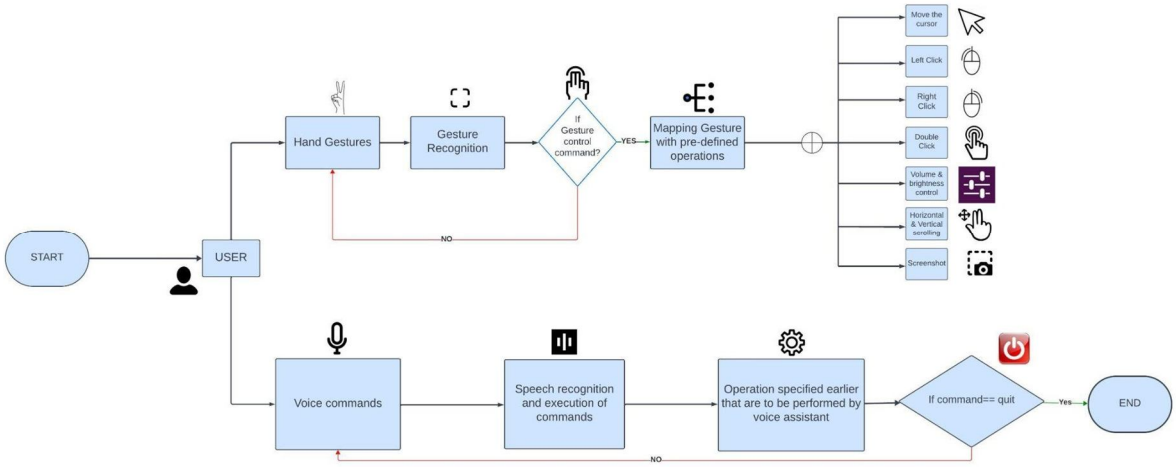


Fig. 5.1 System Flowchart

A gesture recognition technique is used by the system to identify hand motions used as input.
Hand gestures are then used by the user to manage the system after the recognized gestures are mapped to predefined operations.
Speech recognition or voice commands are not used by the system to carry out commands.
Moving the cursor, clicking with the left, right, or double click, adjusting the brightness and volume, scrolling both horizontally and vertically, and taking screenshots are among the actions that may be performed with hand gestures.
The system will react appropriately if the user enters the command "quit".

VI. RESULTS

A. V Gesture (Two-Finger Peace Sign)

Operation: Moves the mouse cursor to the current hand position on the screen.

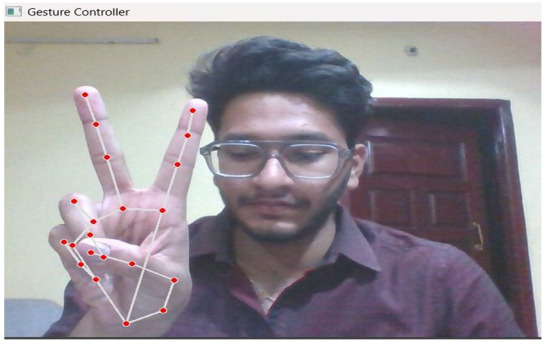


Fig. 6.1 V gesture

B. Fist Gesture

Operation: Holds down the left mouse button for continuous mouse clicks while moving the cursor with hand movement.

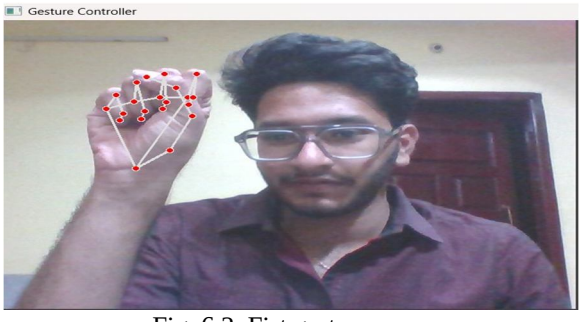


Fig. 6.2: Fist gesture

C. Index Finger Extended

Operation: Performs a right-click operation at the cursor's current location.

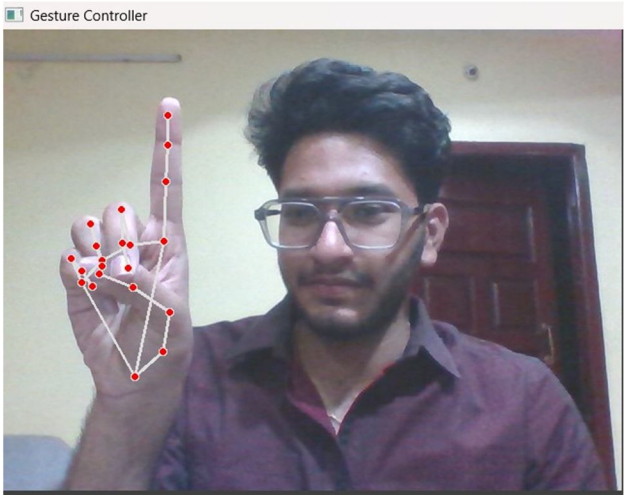


Fig. 6.3: Index finger extended;

D. Middle finger extended:

Operation: Performs a left-click operation at the cursor's current location.

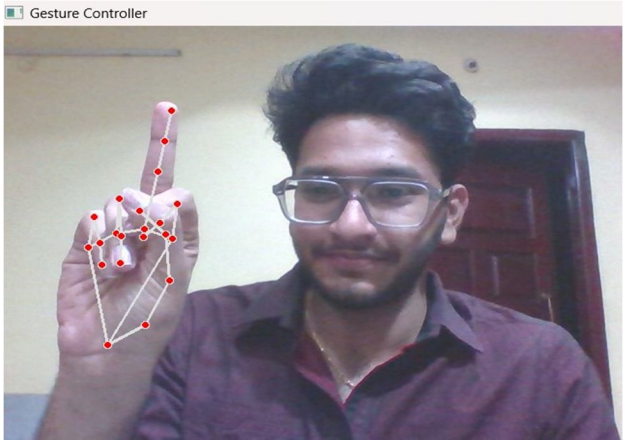


Fig. 6.4: Middlefinger extended

E. Two-Finger Closed Gesture

Operation: Executes a double-click action at the cursor's current position.

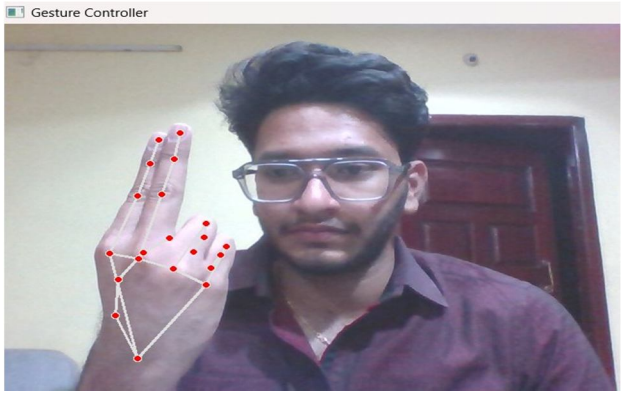


Fig. 6.5: Two-finger closed gesture

F. Pinch Gesture (Minor Hand):

Operation: Performs vertical or horizontal scrolling based on hand movement direction.

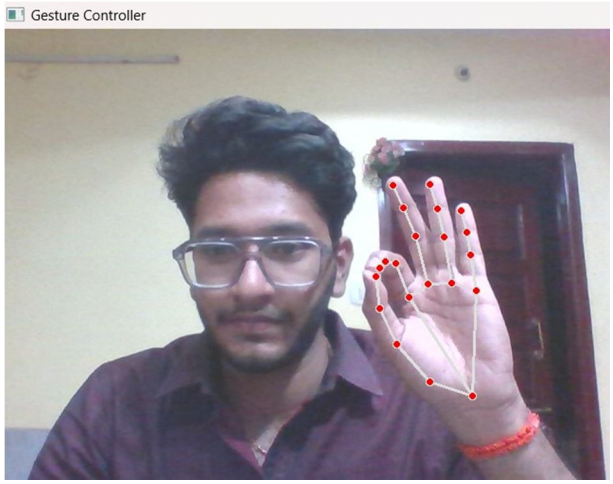


Fig. 6.6: Pinch gesture (minor hand)

G. Pinch Gesture (Major Hand)

Operation: Controls system brightness or volume based on pinch gesture direction and magnitude.

Fig. 6.7: Pinch gesture (major hand)

H. Four-Finger Closed Gesture

Operation: Takes a screenshot and saves it as "sc.png" in the working directory.

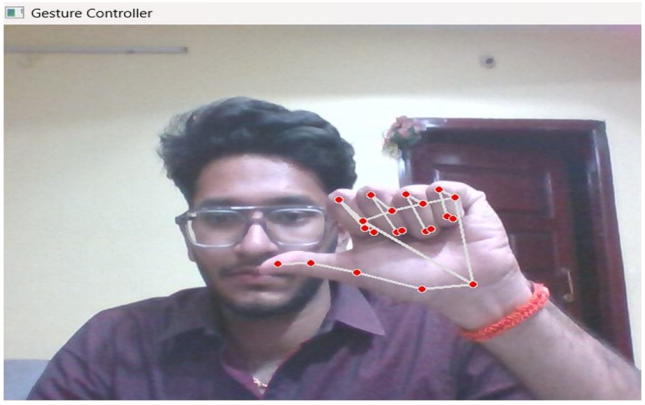


Fig. 6.8: Four-finger closed gesture

VII. FUTURE SCOPE

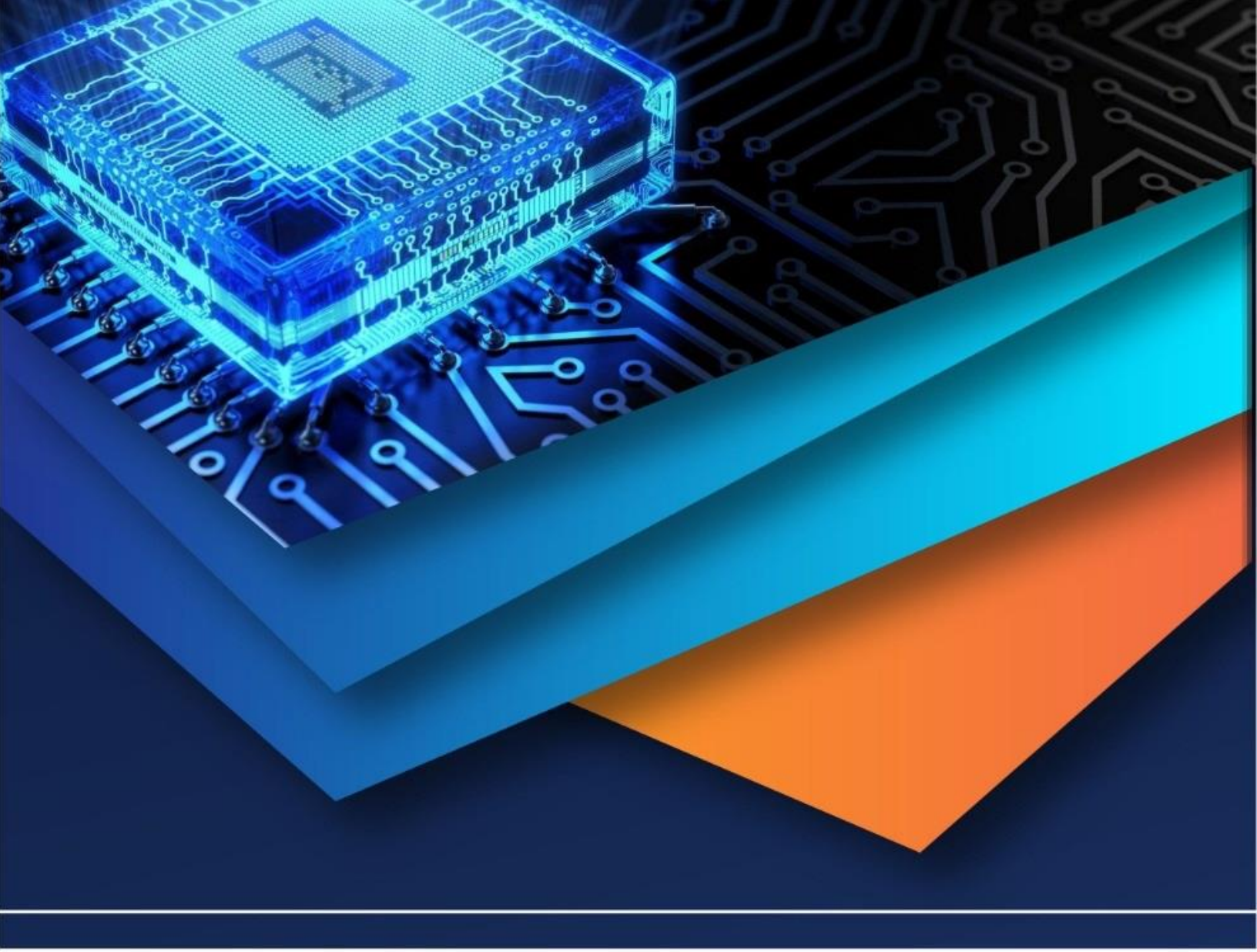
1) *Enhanced Gesture Recognition:* Improve algorithms to recognize a wider range of hand movements accurately, including 3D gestures.
2) *Intelligent Adaptation:* Use context data for dynamic system responses tailored to user environments and preferences.
3) *Advanced AI Assistive Technology:* Implement AI models for better natural language understanding and predictive assistance.
4) *Cross-Platform IoT Integration:* Extend control to mobile devices and IoT ecosystems for seamless interaction.
5) *Security and Privacy:* Incorporate biometric authentication and privacy measures to protect user data and ensure compliance.

VIII. CONCLUSION

Creating a gesture-controlled virtual mouse that works with a voice controller entails combining real-time gesture recognition and speech command processing to provide precise cursor control and task execution. Noise interference, recognition accuracy, and system robustness are all challenges that must be addressed on a continuing basis. Future revisions may investigate advanced features such as context-aware interactions and improved feedback mechanisms for a variety of domains, including gaming, smart homes, and accessibility tools. This integrated method demonstrates unique human-computer interaction paradigms that provide a natural and immersive computing experience.

REFERENCES

[1] Virtual Mouse Using Hand Gesture, Mr. E. Sankar, B. Nitish Bharadwaj, and A. V. Vignesh. ISSN: 2582-390, 2023 IJSREM.
[2] Gesture-Controlled Virtual Mouse with Voice Assistant Ismail Khan, Vidhyut Kanchan, Sakshi Bharambe, Ayush Thada, Rohini Patil, ISSN (O): 2582-631X, 2024 IRJMS
[3] Virtual Mouse Using Hand Gesture And Voice Assistant Khushi Patel, Snehal Solaunde, Shivani Bhong, and SairabanuPansare ISSN: 2349-6002 2024 IJIRT.
[4] GESTURE-CONTROL-VIRTUAL-MOUSE Bharath Kumar Reddy Sandra*1, Katakam Harsha Vardhan*2, Ch. Uday*3, V Sai Surya*4, Bala Raju*5, Dr. Vipin Kumar*6, ISSN: 2582-5208 IRJMETS
[5] GESTURE AND VOICE-CONTROLLED VIRTUAL MOUSE MUHAMMED AVADH SHAN S1, MOHAMED SHEFIN1, APARNA M2 ISSN-2349-5162, 2023 JETIR.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)