International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

# Comprehensive Gestures and Voice Driven Input System

**Sarika[1], Usha M[2]**
[1] *Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India*
[2]*Assistant Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India*

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** In modern times, technology is an integral part of our life. The goal of this Comprehensive Gestures and voice Driven Input System is to create a real-time program that gathers through movements of the hands and turns them into digital art. The project seeks to create a user-friendly and accessible interface for computer-human interaction using OpenCV, Mediapipe, and Python with all of its many packages. To accomplish its goals, the project's main focus is on palm identification and color tracing algorithms. The device follows the user's finger movements using computer vision, allowing users to compose themselves or draw in the air. The innovation holds enormous potential for many applications, including facilitating system communication and reducing the requirement for traditional devices like laptops.
The project also investigates the creation of a voice-controlled assistant, a modern tool that is frequently utilized to improve human-computer connection. Many products have voice assistants built in. By merging all of these technologies, the effort hopes to develop a flexible system that can translate motions into text and recognize them, enabling unique forms of interaction and involvement.

***Key Words***: Machine learning, OpenCV, Mediapipe, hand gesture recognition, air writing, voice assistant, human-computer interaction

## I. INTRODUCTION

Human-computer interaction has undergone a dramatic transformation due to technological breakthroughs, bringing both new opportunities and obstacles. Interactive displays and keypad are two typical classic input modalities that new and older users frequently find difficult to use. Our concept uses gesture recognition and finger motion tracking to provide a more user-friendly solution to these problems. With the growth of typing devices from keyboards to touch screens, the newest technology finger motion tracking simplifies typing by identifying and interpreting hand movements. By using complex mathematical algorithms, gesture recognition systems improve user engagement by offering a practical replacement for mechanical input devices. The "touch-less typing" method improves the creativity and intuitiveness of using computers. Simultaneously, technology for speech detection presents a substitute for conventional keyboard input by converting audible words into printed text that is readable by machines. While speech dialing and speech-to-text applications are becoming more and more popular, they still present a challenge regardless of the intricacies of individual mannerisms, accents, and dialects.
Color tracking and palm detection are two of the primary features that let users draw over the screen like they would with a paintbrush. This Innovation opens up novel prospects for communication and artistic expression, especially for those who could find it difficult to use traditional input devices. Furthermore, the system's adaptability is increased by combining recognition of movements with speech recognition, enabling users to carry out a variety of tasks using basic hand gestures and voice instructions. The overall goal of this project is to create a simple to use, gesture-based painting program that will push the limits of interaction between.

## II. LITERATURE SURVEY

Niharika et al. talks about a novel paint technique that was created to keep kids out of the dust in the classroom during the COVID-19 outbreak. The program detects hand joints and recognizes hand movements using MediaPipe and OpenCV to provide natural Human-Computer interaction (HCI). Improving human-computer interaction is the main objective of this HCI strategy. The tool provides learners with an interesting and interactive way to engage in online learning through the use of hand gestures [1].
R Vasvil had initiated the uses of hand gesture recognition in sign language interpretation, industrial automation, and upper limb rehabilitation, among other applications. It is useful for human-computer interaction because it tracks hand movements using algorithms that are augmented by color markers on the fingertips. Smart surveillance, medical systems, and virtual environments are some examples of applications. This work also includes a painting application that tracks the movements of the fingertip and an identification recognition system that uses hand gestures to identify numbers 0 through 9 [2].
Ismail Khan et al. combined voice assistant technology and hand gesture-controlled mouse capability into a single platform known as GCVA (Gesture regulated Voice Assistant). Through natural motions and voice instructions, the system mimics traditional mouse and keyboard functions by utilizing technologies for gesture recognition, and AI/NLP for speech commands. This configuration offers reliability, interoperability with current computer installations, and practicality by doing away with the requirement for a real mouse. High accuracy gesture detection and responsive voice command execution were demonstrated in user testing, indicating that these systems may eventually replace conventional input devices [3].
M Raja presented a Gesture Controlled Virtual Mouse system that allows computer interaction using voice instructions and hand gestures. The system consists of two primary modules: one that allows for direct hand contact using MediaPipe's Hand detection, and another that improves recognition of gestures and control by using gloves of any uniform color. This strategy seeks to lessen the need on physical stimulation devices in order to simplify digital participation and availability [4]. Prihvi J et al. proposed a system that allows users to connect with computers without making physical touch via hand

gestures and vocal commands is introduced. Advanced ML and This novel method offers a comfortable and natural way to operate computers, with potential uses in the fields of education and healthcare. The paper's thorough analysis and system implementation will be helpful to academics and professionals in the field of human-computer interaction as they continue to advance interactive technology [5].

## III. EXISTING SYSTEM

The accuracy of this procedure is 90.125%. The complicated convex hull approach is utilized to set the virtual pointer and computer technologies.
Recognition of many hand gestures in the backdrop may have limitations. Another intricate technique employed in the development of the hand gesture-based mouse is the color segmentation procedure.
Only a limited number of languages may be translated into texts using speech to text, which is a technology restriction. Teaching mid-air and multi-touch gestures is more challenging than teaching single-touch movements.
With the latter, users simply need to accurately follow a path in order to execute a command; hand position is not important. However, when multi-touch and mid-air motions are added, the position and motion of many fingers, or even the entire hand, become significant.
Teaching systems often focus on commands that can also be executed using a single-touch input device, such as a mouse or pen, and train the user about the essential hand movement and path for an action rather than the pose and form of contact.
One of the methods' shortcomings is that you can't always count on the voice identification software to accurately translate your words into the screen. Programs frequently make mistakes that are the result of miscommunication because Individuals are not capable of doing so to distinguish between homophones and cannot comprehend language context inside the identical thing way that people can.

Some of the disadvantages are:

- The convex hull algorithm's complexity.
- The cost of a Leap motion controller is high.
- It takes longer to train.

## IV. PROPOSED SYSTEM

With the Gesture-Controlled Vr Mouse System, users can utilize hand movements that are recorded by a webcam to communicate with their computers. It tracks and interprets movements of the hand in real-time, converting them into mouse clicks and pointer movements by utilizing computer vision techniques.

❖ **Palm tracking and gesture identification**: uses computer vision algorithms to track and detect hand motions made by the user, such as Media pipe and OpenCV use identifies particular cursor control gestures, such as moving the cursor and clicking with the left and right buttons.

❖ **Cursor movement and interaction**: Converts hand motions that are sensed into mouse movements that are simulated on a computer screen.improves the experience of digital content by enabling clicking actions using recognized gestures.

❖ **Emulation of a Virtual Keyboard**: Provides a hand gesture-based virtual keyboard interface for users to manipulate. Produces inputs from the keyboard based on detected motions, enabling the use of the keyboard for composing and other purposes.
❖ **Combining External Device Integration**: Integrates optionally with additional hardware, such as the Leap Motion controller, to increase the precision of gesture recognition.The AI Virtual Art System does away with the necessity for conventional input devices like drawing tablets by enabling users to generate digital artwork with hand motions.

❖ **Hand Gesture-Based Drawing**: Strategies based on computer vision are used in this strategy to track the user's hand movements. Transforms hand movements into digital canvas shapes and strokes. supports specialized motions for functions including sketching, erasing, and tool selection.

❖ **Real-Time Canvas Interaction**: This feature offers a responsive canvas on which users may manipulate and paint with simple hand gestures.uses gesture commands to provide color selection and dynamic brush control.

❖ **Interaction with Speech Identification**: This feature allows you to annotate artwork or speak manage the painting process. It also features speech-to-text functionality.

Some of the key advantages are:

- It requires little to no training time.
- Although, leap motion controller is not used, the system requires less pricey.
- The interface is easy to use.

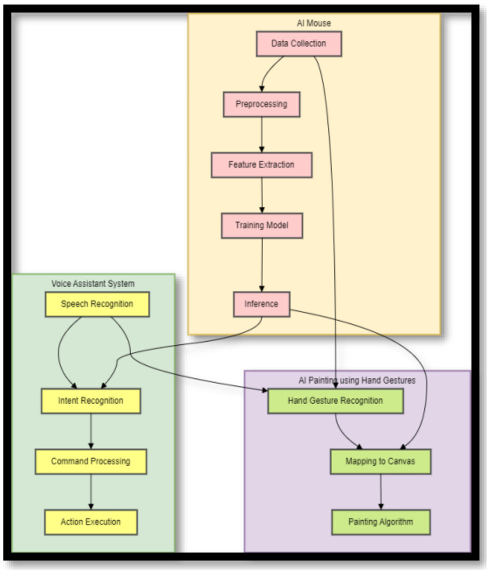


**Fig - 1:** System Architecture

## V. IMPLEMENTATION

**Hand Gesture Recognition Module**

The goal of the hand gesture recognition module is to record and analyze hand movements in order to identify particular gestures that are used to operate virtual painter and mouse apps. Real-time video feed of the user's hand captured by a webcam. OpenCV: Manages image processing jobs like color space conversion, frame capture, and gesture recognition preprocessing.
Mediapipe: Recognizes movements and detects hand landmarks.
Image Capture: The webcam constantly records frames from videos.
Preprocessing: To make the frames compatible with Mediapipe, they undergo conversion from the BGR to the color space of RGB.
Hand Signature Identification: 21 landmarks on the hand, including fingertips and knuckles, are identified by Mediapipe. Movement Identification: Certain gestures (such as pointing and pinching) are identified based on the locations and motions of these landmarks.

**Voice Recognition Module**

By enabling the system to comprehend spoken instructions, this module improves accessibility and functionality.
Microphone: Collector of user audio input. Text is generated from spoken words using the Speech Recognition Library.
Audio Gather: The microphone is what records spoken commands from the user.
Voice-to-Text Conversion: The audio input is processed, converted to text, and then understood as an instruction by the voice detection library.

**Virtual Mouse Control Module**

This module offers an alternative method of controlling the computer's cursor without a real mouse by translating recognized hand motions into mouse operations.
Recognition of Hand Gestures Output: Information on the kind of gesture made.
Cursor Control Library Autopy, enables cursor control via programming.
Interpretation of Gestures: Certain mouse activities are associated with recognized gestures.
Cursor Movement: The cursor moves when the hand moves.
Clicking: Certain motions, such as pinching, can be used to mimic both left and right mouse clicks.
Scrolling: It is possible to map vertical hand motions to scroll actions.

**Virtual Art Module**

The goal of this module is to enable users to create a digital canvas for writing or painting by utilizing their hand movements to draw on the screen.
Recognition of Hand Gestures Data representing fingertip gestures and positions is the output.
Sketching Interface: The drawing is done through a graphical interface.
Sketching Activation: The system switches to drawing mode when it recognizes a drawing gesture, such as pointing with one finger. Fingertip Tracking: For drawing lines on the screen, the fingertip's position is monitored.

## VI. METHODOLOGY

- Taking and processing videos: In order to recognize hand motions frame by frame, the webcam continually records frames, which are subsequently changed from BGR into RGB color space.
- Cursor Control with Hand Gestures: The mouse cursor in the computer window is moved by using transformational techniques to translate fingertip locations from the webcam screen. Cursor motions are made easier by the webcam feed's rectangular section that matches the entire computer screen.
- Mouse actions Based on motions: Certain mouse actions, such as clicking and scrolling, are activated by hand motions that are identified by the tip IDs of lifted fingers. As a result, people are able to communicate with the computer using their palms.
- The keyboard Emulation: The operating system displays simulated keyboard keys on the machine's the screen, allowing individuals to carry out mouse functions using gestures with their hands. By detecting finger movements, the interface minimizes the need for actual hardware by selecting and operating keys.
- Integration with Speech Recognition: Using Python's Speech Recognition package, another feature includes speech-to-text conversion. This function enables spoken instructions to be translated into text input, improving user involvement.
- Hand Tracking Module: This module uses the webcam to record and examine hand gestures. It then uses the the Mediapipe library's Glove Recognition and Palm Detection models to analyze each frame and identify the location of each hand.
- Cursor Control Module: This module uses the autopy library to provide mouse functions including scrolling, clicking, and movement, improving hand-only computer interface.

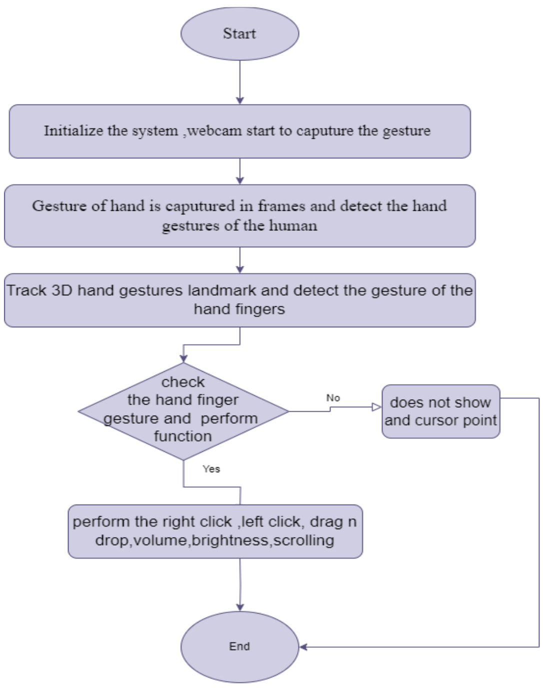


**Fig – 2:** Flowchart of Gesture Controller

## VII. RESULTS

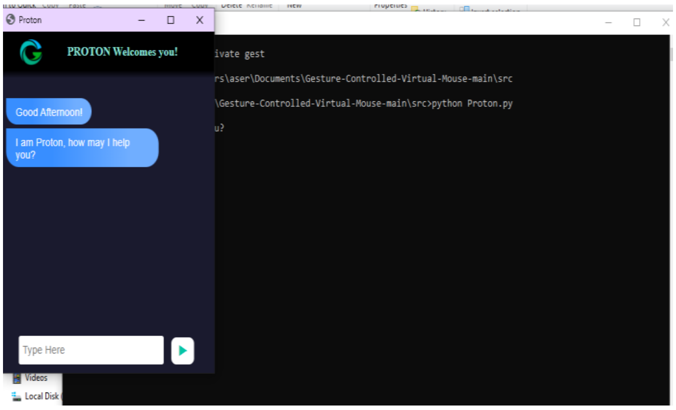

**Fig – 2:** Voice Assistant Started

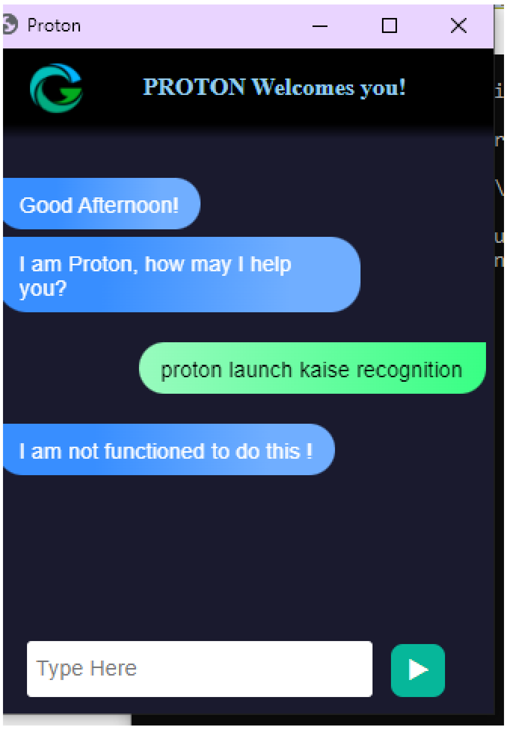


**Fig – 2:** Launch untrained module

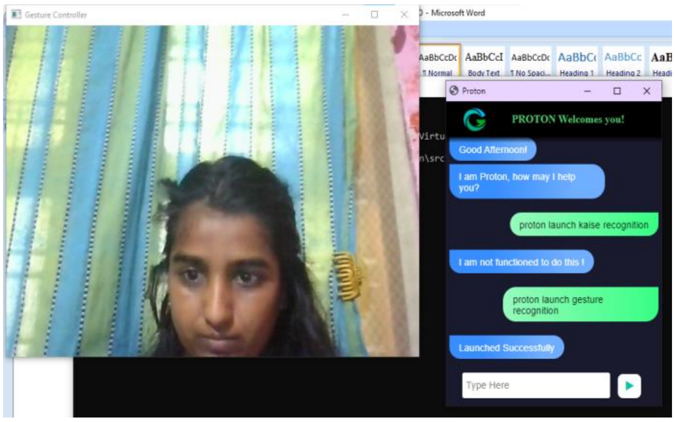

**Fig – 3:** Launch Gesture module

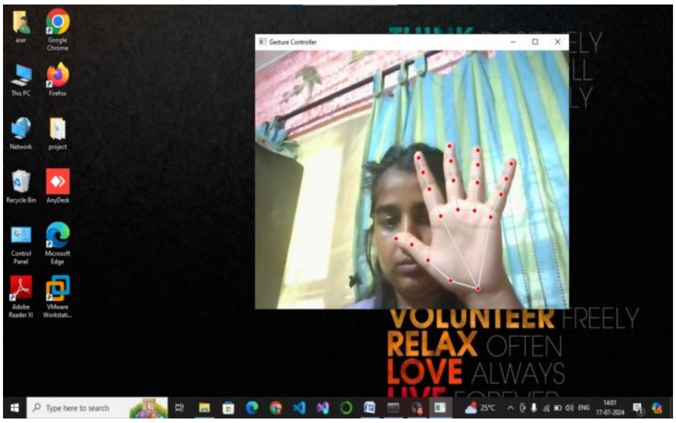

**Fig – 4:** Gesture interface

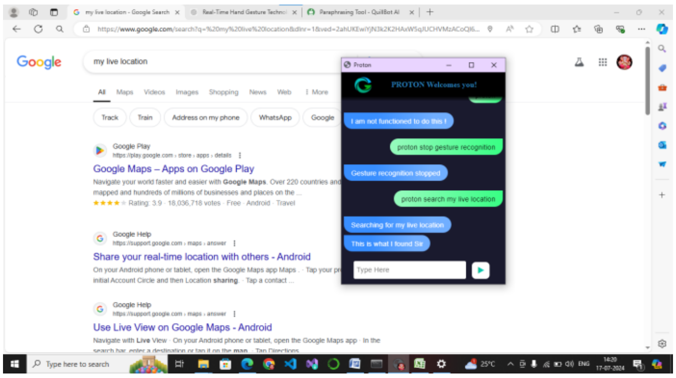

**Fig – 5**: Find Location


**Fig – 6:** Paint module

## VIII. CONCLUSION

For users who find standard input devices difficult to use, the proposed system's integration of gesture-based control with voice help provides a substantial leap in human-computer interaction. This system enables natural and intuitive interaction with digital environments using real-time hand gesture recognition using Mediapipe and OpenCV. Simple hand gestures can be used to manipulate the cursor, move the mouse, and draw onto a virtual canvas. Furthermore, adding a module for voice recognition expands the system's functionality by allowing users to speak instructions and enter text. Together, Comprehensive gesture and voice driven input system projects address a number of the drawbacks related to conventional input techniques. They give a novel approach that lessens the need for actual hardware, lowers accessibility hurdles, and offers an interface that is easier to use. Together with voice-to-text conversion, the ability of the system to convert hand gestures into mouse movements and drawing command creates new avenues for digital interaction and artistic expression

## IX. FUTURE ENHANCEMENTS

The virtual art and gesture-controlled mouse systems have a bright future ahead of them. Future updates will prioritize improving gesture detection, with the goal of applying more sophisticated models for machine learning and reliable training datasets to increase accuracy. Additionally, a wider variety of gestures will be included to facilitate intricate interactions. Users will be able to communicate with virtual things in their real world through integration of augmented reality (AR), which will improve immersion and provide real-time visual feedback. More sophisticated voice commands and inquiries will be possible with the expansion of voice command capabilities to include multilingual support and cutting-edge NLP (Natural Language Processing) methods.

## X. REFERENCES

[1] Niharika M1, Neha J2, Mamatha Rao3, Vidyashree K P4 VIRTUAL PAINT APPLICATION USING HAND GESTURES International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 09 Issue: 04 | Apr 2022

[2] R.Vasavi Nenavath Rahul2 , A Snigdha3 , K. Jeffery Moses4 , S.Vishal Simha5 Painting with Hand Gestures using MediaPipe International Journal of Innovative Science and Research Technology Volume 7, Issue 12, December – 2022

[3] Ismail Khan, Vidhyut Kanchan, Sakshi Bharambe*, Ayush Thada, Rohini Pat Gesture Controlled Virtual Mouse with Voice Assistant International Research Journal of Multidisciplinary Scope (IRJMS), 2024; 5(1 ) 02nd January 2024; Published 30th January 2024

[4] Voice Assistant and Gesture Controlled Virtual Mouse using Deep Learning Technique M Raja IEEE 25 April 2023

[5] Prithvi J, S Shree Lakshmi, Suraj Nair and Sohan R Kuma Gesture Controlled Virtual Mouse with Voice Automation International Journal of Engineering Research & Technology (IJERT)Vol. 12 Issue 04, April-2023