Virtual Mouse

Shanu Thakur Dept. Computer Science MIT ADT University,
Pune, India Shanu27thakur@gmail.com[1]

Radha Krishna Dept. Computer Science MIT ADT University,
Pune, India rkrishna0019@gmail.com

Yuvraj Nikam Dept. Computer Science MIT ADT University.
Pune, India Yuvraj.nikam@mituniversity.edu.in

***Abstract*— Now a days the use of hand gesture recognition became popular due to the advance technology of artificial Intelligence (AI). In recent years, there has been a growing interest in developing alternative methods for human-computer interaction (HCI) that go beyond traditional input devices like keyboards and mouse. Hand gesture recognition has emerged as a promising approach, offering intuitive and natural ways for users to interact with digital interfaces.**

The proposed system utilizes computer vision techniques to detect and track hand gestures in real-time. By mapping specific gestures to mouse control commands, users can navigate and interact with graphical user interfaces (GUIs) without the need for physical input devices.

In this paper A hand gesture- controlled virtual mouse system utilizes the AI and Machine learning algorithms to recognize the proper hand gestures and translate them into the mouse movements. The system we are designed that the people who have problems or difficulty using a traditional mouse or keyboard it will be appropriate for them. For this we are using camera that capture images of the user's hand, which are processed by an AI and ML algorithm to recognize the gesture of hands being made. We trained the system by using a dataset of hand gestures.

***Keywords* — Hand gesture recognition, Machine Learning, Artificial Intelligence, Virtual Mouse, Python, Media Pipe**

I. Introduction

In our modern world, technology pervades every aspect of our lives, with advancements in computer technology progressing rapidly. Computers now undertake tasks that were once impossible for humans, exerting a profound influence on our daily routines.

Within the realm of computer science, progress has been exponential, fundamentally altering how we engage with machines. Initially, wired mice facilitated graphical user interface (GUI) navigation, evolving seamlessly into wireless functionality. These peripherals, such as mice, are integral to human-computer interaction (HCI), enabling actions like scrolling, pointing, and movement within GUI environments.

Recent breakthroughs have introduced speech recognition and eye-tracking techniques, promising to redefine HCI boundaries. Speech recognition allows hands-free control via voice commands, but may suffer from delays in executing mouse functions. Similarly, while eye-tracking provides precise cursor control, challenges arise with individuals wearing contact lenses or possessing distinctive eyelash features.

Over time, mouse technology has evolved from wired to wireless, enhancing functionality for smoother movements. Concurrently, speech recognition and eye-tracking technologies have emerged, offering alternative HCI methods. Speech recognition aids voice-based tasks like search and translation, albeit with recognition delays. Conversely, eye-tracking controls the cursor, albeit with limitations for some users.

Hand gesture-controlled virtual mice represent another innovative HCI approach, leveraging camera vision to track hand movements for on-screen actions. This technology eliminates the need for physical mice, enhancing accessibility and providing a more intuitive user experience, particularly when traditional input devices are impractical.
The latest advancements offer various advantages, including enhanced accessibility and natural interaction with computers. Hand gesture-controlled virtual mice find applications in diverse fields like virtual reality, gaming, and accessibility, offering a user-friendly alternative to traditional input methods.

II. Literature Review

Previous attempts at creating a virtual mouse involved cumbersome methods such as glove-based systems or using colored pieces of paper attached to hands for gesture recognition. However, these approaches were found to be impractical and inaccurate, with issues such as discomfort from wearing gloves for extended periods. The IEEE paper "Virtual Mouse using Hand Gestures" authored by Roshnee

2

Matlani and Roshan Dadlani introduces the Leap Motion project, which focuses on gesture-controlled laptops and computers. Hand gesture technology has gained popularity across various sectors, including medical, industrial, IT, and banking.

Another paper by E. Sankar Chavali, titled "Virtual Mouse Using Hand Gesture," highlights the rapid advancement of computer vision technology, enabling innovative approaches like hand recognition for controlling mouse functions. However, traditional static hand recognition systems, which assign actions to predefined hand shapes, are limited in their capabilities and can lead to confusion.

Advancements in technology offer alternatives to traditional mouse usage, with gesture-controlled virtual mice combining voice commands and hand motions for seamless interaction. These systems leverage cutting-edge Computer Vision and Machine Learning algorithms, eliminating the need for additional equipment. Utilizing models developed by Media Pipe and tools like ESP32-CAM and OpenCV, gesture- controlled virtual mice offer intuitive and natural interaction with computers.

The paper "Hand Gesture Controlled Virtual Mouse Using Artificial Intelligence" by Kavitha R, Jana Sruthi S U, Lokitha S, and Tharani G emphasizes the importance of human- computer interaction, traditionally facilitated by output devices like mice. This technology enables users to control computer movement through hand gestures, replacing physical mice. By utilizing camera vision-based approaches, the system provides a more intuitive and natural interaction method, eliminating the need for physical peripherals and enhancing user experience across all computing devices.

III. Problem Statement

- The objective is to develop a virtual mouse capable of recognizing hand gestures and executing mouse functions solely through finger movements.
- Various combinations of fingers are utilized to perform different mouse operations based on the recognized finger combinations.
- Unlike previous systems, our proposed solution does not require users to color their fingers or utilize external devices or sensors.
- The goal of the proposed virtual mouse system is to offer a user-friendly and intuitive interface for interacting with digital environments, especially in situations where traditional input devices are impractical or inaccessible.
- By utilizing hand gestures as input commands, the system provides users with increased flexibility and mobility, reducing the dependence on physical peripherals.
- Furthermore, the absence of external devices and sensors simplifies the setup process and enhances the system's adaptability and accessibility across diverse computing platforms.

IV. Algorithms and Tools used

In our project Hand gesture controlled virtual mouse for hand and finger detection purpose we are using the algorithm Media Pipe which is a type of framework based on the cross platform developed by google and Open cv to perform some CV related task.

1. Media Pipe:

The Media Pipe framework, developed by Google, stands as an open-source platform enabling the development of cross- platform, real-time computer vision applications. It equips developers with a suite of pre-trained models and tools essential for constructing real-time applications geared towards tasks like hand tracking, pose estimation, face detection, and gesture recognition.

A cornerstone feature of MediaPipe is its modular architecture, empowering developers to forge custom pipelines by amalgamating different components known as "MediaPipe graphs." These graphs specialize in distinct tasks, such as feature extraction or model inference, and can be interconnected to fashion intricate processing pipelines.

Boasting support for both desktop and mobile platforms, MediaPipe proves itself versatile for a diverse array of applications across various devices. Renowned for its user-friendly interface, scalability, and performance optimizations, MediaPipe has garnered popularity among developers for crafting real-time multimedia applications encompassing augmented reality (AR) experiences, gesture-controlled interfaces, interactive games, and more. Its open-source nature and comprehensive documentation further enhance its accessibility to developers of all proficiency levels, fostering swift prototyping and experimentation in the realm of multimedia processing and computer vision.

MediaPipe embodies a multimodal approach, enabling its application across a myriad of video and audio formats. Developers leverage its algorithms for system construction and analysis through graphs, catering to a wide spectrum of application needs. Additionally, the framework is utilized for developing systems for various application purposes.The Media Pipe framework has interfaces which helps users to interact with other machine learning libraries like TensorFlow and Py-Torch and it also supports several programming Languages, like C++, Python, Java.

Media Pipe offers numerous features like:

1. Facial Recognition
2. Image Processing
3. Video and Audio Processing
4. Hand Tracking
5. Object Detection
6. Pose Estimation.

2. Open CV:

OpenCV, abbreviated for Open-Source Computer Vision Library, stands as an open-source software library dedicated to computer vision and machine learning. It boasts a wide array of functions geared towards image processing, object detection, feature extraction, and more. Written primarily in C++, OpenCV also offers interfaces for Python, Java, and MATLAB, ensuring accessibility for developers across various programming languages.

A standout feature of OpenCV lies in its extensive repository of pre-trained models and algorithms, facilitating seamless integration into projects for diverse computer vision tasks. These encompass techniques like facial recognition, object tracking, gesture recognition, and even deep learning-based approaches for tasks such as image classification and object detection.

This computer vision library is a versatile tool for developing real- time applications, supporting various programming languages like C, C++, Python, and Java, and compatible with platforms like Windows, Linux, macOS, iOS, and Android.

OpenCV shines in its comprehensive array of algorithms and tools tailored for real-time computer vision applications. These encompass methods ranging from facial recognition and gesture detection to motion tracking and stereo vision. Furthermore, OpenCV seamlessly integrates with deep learning frameworks like TensorFlow and PyTorch, enabling the utilization of cutting- edge neural network models for tasks like image classification and object detection.

Efficiency is a hallmark of OpenCV, with its algorithms optimized for performance across diverse hardware platforms, including CPUs, GPUs, and specialized accelerators such as Google's Edge TPU. This versatility renders it suitable for a broad spectrum of applications, spanning from simple image processing tasks to complex real-time systems, both on desktop and embedded platforms.

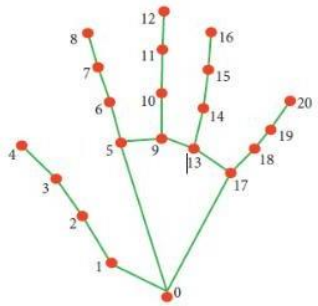
V. System Development

The development of a hand gesture-controlled virtual mouse system encompasses several crucial stages, each contributing to its robustness and efficiency. Initially, research and development efforts focus on algorithm design, leveraging computer vision techniques to detect and track hand gesture accurately.

In this project we are using various functions and conditions are used. The functions and conditions are explained as below:

1. Camera used in the AI Virtual Mouse System: The AI virtual mouse system is based on the frames that have been captured by webcam in laptop or PC. By using the OpenCV library, the video capture object is created, and the web camera will start capturing the video. The web camera captures and passes the frames to the System.

2. **Capturing the video and Processing**:
The virtual system uses the webcam where each frame is captured till the termination of the program. The video is processed from BGR to RGB colour space to find the hands in video frame by frame.



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

3. Rectangular Region for Moving through the window:

The virtual mouse makes use of the transformational algorithm, and it converts the coordinates of fingertip from the webcam screen to the computer window full screen for controlling the mouse.

4. Detecting which finger is up and Performing the Particular Mouse Function:

In this scenario, we identify the raised finger by referencing the tip ID assigned to each finger, a feature provided by MediaPipe. By examining the coordinates of the raised fingers, we determine which specific mouse function to execute based on predefined criteria or user-defined mappings.

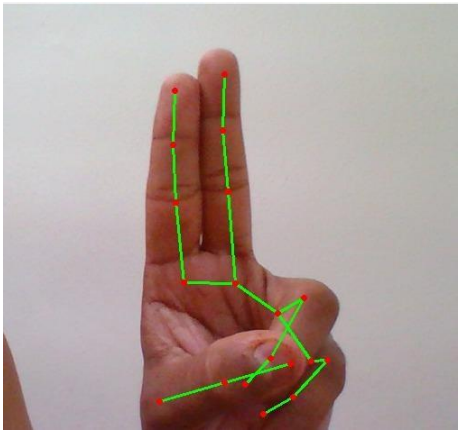
For instance, if the index finger is raised, it might correspond to a left-click action, while raising the thumb could signify a right-click action. By associating each finger with a specific mouse function, users can control the mouse cursor without needing a physical mouse.

This approach capitalizes on MediaPipe's hand tracking and finger detection capabilities, which provide precise information about each finger's position and movement. The system then interprets this data to execute the corresponding mouse function, offering users an intuitive and natural way to interact with digital environments.

5. The Mouse to perform Left Button Click:

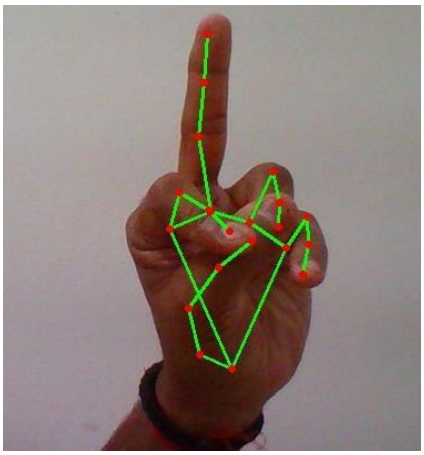
If the index finger of both the hands with the tip ID = 1 & the thumb finger with tip ID = 0 are up. The distance between the fingers is lesser than 30px, and the computer is made to perform the left mouse button click using the pynput.

VI. Output Screen

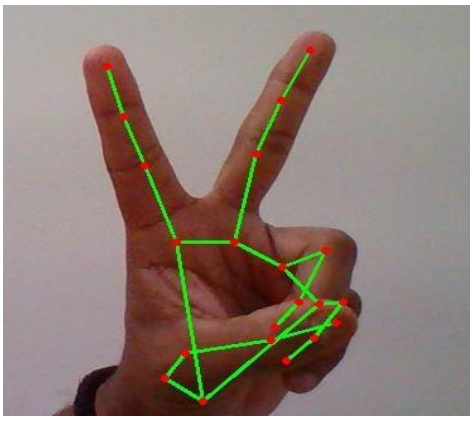


Double Click

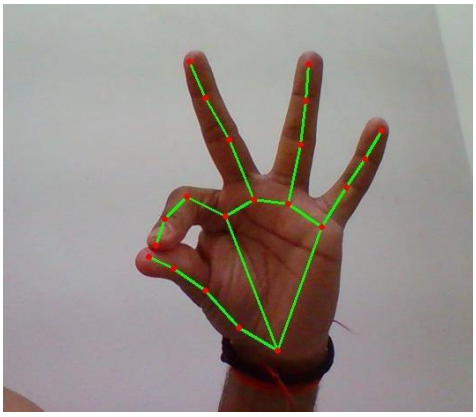
Increase Brightness and Voulme



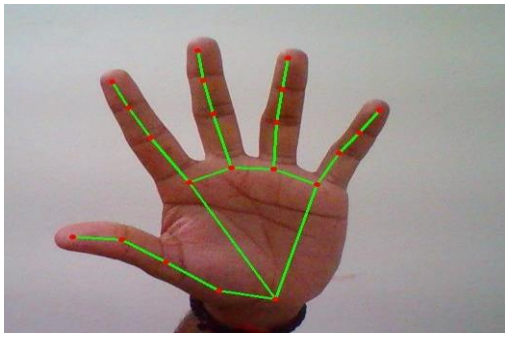
Left Click



Move Cursor



Neutral Gesture



VII. Conclusion

In conclusion, the document encapsulates with a holistic journey of ideation, research, development and implementation of a hand gesture-controlled virtual mouse system. By leveraging cutting- edge technologies and innovative approaches, the project not only addresses existing challenges but also opens doors to new possibilities in human-computer interaction. As technology continues to evolve, hand gesture recognition stands at the forefront of intuitive interfaces, paving the way for a more accessible and seamless computing experience for all.

The development of a virtual hand gesture-controlled mouse system offers users a natural and intuitive interface for interacting with computers. By leveraging computer vision and machine learning technologies, this system enables users to navigate digital interfaces and perform mouse functions using hand gestures, without the need for physical input devices. Through its modular architecture and advanced algorithms, the system provides a versatile and responsive interface, addressing accessibility barriers and enhancing user experience across various scenarios. This innovation has the potential to revolutionize human-computer interaction, offering new possibilities for intuitive computing in fields such as virtual reality, gaming, and accessibility.

VIII. References

1. OpenCV - Open Computer Vision Library
2. https://www.bing.com/search?q=MediaPipe&qs=n&form=QBRE&sp=-1&ghc=1&lq=0&pq=mediapip&sc=10-8&sk=&cvid=0E550E66823341A0856D8BE06E82144C&ghsh=0&ghacc=0&ghpl=
3. Virtual Mouse using Hand Gestures | IEEE Conference Publication | IEEE Xplore
4. (PDF) Virtual Mouse Using Hand Gesture (researchgate.net)
5. MediaPipe | Google for Developers
6. OpenCV - Wikipedia