Virtual Mouse Using Hand Gesture and Voice Assistant

Khushi Patel[1], Snehal Solaunde[2], ShivaniBhong[3], Prof. Sairabanu Pansare[4]
[1,2,3,4] *Nutan College of Engineering and Research*

Abstract- **This paper aims to improve the recognition of human hand postures in a Human Computer Interaction application, reduce the time spent computing and improve user comfort related to human hand postures. The authors developed an application for computer mouse control. Based on the proposed algorithm and selected hand feature, the application has good time-based performance. The user finds it easier to operate the system due to the proposed hand postures in combination with the voice assistant.**

Index Terms**—Detection, Gesture, Mediapipe, OpenCV, python**

I. INTRODUCTION

Over the past few years, the interests regarding the development of human-computer interaction systems have increased. The availability of software packages that provide features for particular processing and make it simple to communicate between computers, computing devices encourages students, engineers, and researchers to concentrate on the creation of natural intuitive user interfaces. Designed a multi-modal interface for persons who use wheelchairs and have impairments. A user interface like this one can improve the quality of life for persons with disabilities by giving them a few alternatives for controlling their wheelchair. Because human hand postures and gestures are still a potent inter human communication mode, the research community is interested in using them to control or interact with artificial systems. The use of hand gestures as a control or communication modality continues to stimulate the scientific community's curiosity.
with the goal that we can likewise make work areas as easy to use as cell phones.

II. SCOPE OF PROJECT

Our goal is to implement additional gestures so that in the future, users will be able to perform more tasks efficiently. This project proposes a system that only uses the appropriate hand to perform gestures. It will therefore be possible in the future for improvement of the currently implemented technique to utilize both hands for different gestures. Many applications have been substantially improved through the rapid development of hand gesture recognition systems. Among other things, it allows Doctors and surgeons to see the scans of each patient more clearly, enables online learning and distance training, and can also control televisions. For hearing impaired individuals, signing language interpretation is one of the most effective applications. Projects should serve the objective for which they were developed, thereby illustrating the success of their execution. And through the voice assistant feature, we've automated tasks such as finding locations on Google Maps, navigating documents, launching and stopping gesture recognition, performing Google searches, and sleeping / waking up the voice assistant. This feature will help the end user save both time and effort, as well as help the handicapped and blind use the computer.

III. LITERATURE SURVEY

To accomplish each module in this project we went through a lot of previous works. One of the major parts of this project is the face detection part.

1]Journal of Healthcare Engineering | Hindawi in 2021,"Deep Learning-Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19 Spread". The concept of advancing the human-computer interaction using computer vision is given.

2]V. Geetha, C.K.Gomathy, KottamasuManasa Sri Vardhan,NukalaPavan Kumar, "The Voice Enabled Personal Assistantfor Pc using Python April 2021" , International Journal of Engineering and Advanced Technology 10(4):162-165.

3]Chaithanya C, Lisho Thomas, Naveen Wilson, and Abhilash SS in 2018 proposed "Virtual Mouse Using Hand Gesture" where the model detection is based on colors.

4]L. Thomas, "Virtual mouse using hand gesture,"

International Research Journal of Engineering and Technology (IRJET, vol. 5, no. 4, 2018.View at: Google Scholar

5]Vinay Kr. Pasi, Saurabh Singh, and PoojaKumari in 2016 proposed "Cursor Control using Hand Gestures" in the IJCA Journal. The system proposes the different bands to perform different functions of the mouse. The limitation is it depends on various colors to perform mouse functions.

6]AbhayDekate, ChaitanyaKulkarni, Rohan Killedar, "Study of Voice Controlled Personal Assistant Device", International Journal of Computer Trends and Technology (IJCTT) – Volume 42 Number 1 – December 2016.

7]Dung-HuaLiou, ChenChiung Hsieh, and David Lee in 2010. proposed a study on "A Real-Time Hand Gesture Recognition System Using Motion History Image." The main limitation of this model is more complicated hand gestures.

IV. TECHNICAL OVERVIEW

The purpose of this project is to eliminate touch by using a Hand Gesture Recognition and Voice Assistant. It consists of two main sections: Gesture Recognition and Voice Assistant (Albus).

A. Virtual Mouse using Hand Gesture :

With the help of a web camera and colour detection technique, we have manipulated the mouse cursor movement and various click events. It is possible to virtually control all i/o operations using static and dynamic hand gestures, along with the assistance of a voice assistant. Performing mouse actions in Python requires the use of OpenCV module, which is used for mouse actions. A webcam captures the hands in real-time. A process is performed to extract only the colored fingertips from the video. Following the calculation of the center of gravity of the palm based on the relative positions of the fingertips, various operations are performed to keep track of the cursor. The project recognizes hand gestures and voice commands automatically without additional hardware, using state-of-the-art Machine Learning and Computer Vision algorithms.

Fig.1. Flowchart of Virtual Mouse

B. Voice Assistant:

A voice assistant – Albus is a digital assistant that uses voice recognition, language processing algorithms, and voice synthesis to listen to specific voice commands and return relevant information or perform specific functions as requested by the user. Based on specific commands, sometimes called intents, spoken by the user, voice assistants can return relevant information by listening for specific keywords and filtering out the ambient noise. While voice assistants can be completely software based and able to integrate into most devices, some assistants are designed specifically for single device applications, such as the Amazon Alexa Wall Clock. Today, voice assistants are integrated into many of

the devices we use on a daily basis, such as cell phones, computers, and smart speakers. Because of their wide array of integrations. There are several voice assistants who offer a very specific feature set, while some choose to be open ended to help with almost any situation at hand.

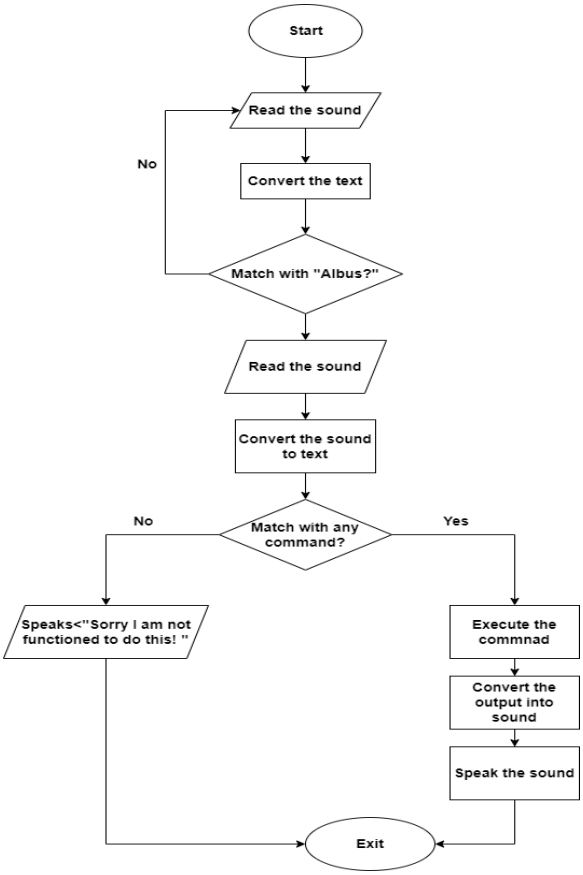


Fig.2. Flowchart of Voice Assistant

V. METHODOLOGY

Firstly we need to import the required libraries for the program like Speech Recognition, Gesture_Controller, pynput, pyautogui, Wikipedia, opencv-python, mediapipe, ctypes, comtypes, pycaw, screen-brightness-control, eel, webbrowser, datetime, math, os, etc.

CAMERA:
Based on the frames captured by a laptop or PC's webcam, the proposed AI virtual mouse system works. As a result of using the Python computer vision library OpenCV, the video capture object is created and the webcam will begin to capture video. The webcam then captures the frames and passes the ones to the AI virtual machine.

VIDEO CAPTURE: A webcam is used for the AI virtual mouse system, which captures every frame until the program is closed. Following is the code that finds the hands in the video frame by frame by converting BGR frame to RGB frame of the video. Defines a function called mp_hands that will help us track the position of our hands.
Detecting Which Finger is up and our job at this point is to detect which finger movements and the respective coordinates of the fingers that are up based on the information found in the library, and perform the mouse function accordingly. This stage involves maintaining which finger is up based on the tip Id we found using the MediaPipe as well as their respective coordinates and then performing the appropriate mouse function accordingly. It sets up a loop that will track the position of our hands every 0.1 seconds.
Mouse Features Using Computer Vision to Determine Hand Gestures and Hand Tip DetectionCursor Moving:The mousepointer is made to travel across the computer's window using the Python AutoPy package when the index and middle fingers are up for the Moving Function.The HandRecog class is initialized with the HandLabel of the person's right hand. The initialize_hand_result method is called to initialize the hand result with the first frame's landmark points. This method is called when the optical flow is complete. The update_hand_result function checks to see if the hand is moving, and if so, returns a None with the Gesture.MOVING.
If the hand is not moving, then the landmark points arechecked for a known gesture. The function then checks to see if the gesture has changed. If it has, it updates the current gesture and increments the gesture frame count. Theget_signed_dist method returns a signed distance in pixels between two landmark points. The get_dist method returns the normal Euclidean distance between two landmark points.

Left Button Click: Left button click is made possible by wide-opening the index finger and dragging it onto the file or folder.

Right Button Click: Wide-opening the middle finger and dragging it onto the file or folder to explore more activities on it enables right button click.

Scroll Up/Down: The computer is designed to accomplish the mouse operation of scrolling up and down by merely pinching, which is done by connecting the tips of the index and middle fingers to the screen's scroll bar.

*Voice Assistant:*It takes input from the user, then converts the speech into text & analyze the text if it matches with the conditions it will give an output or else it will respond “cant recognize”.

VI. RESULT ANALYSIS

The idea of improving human-computer interaction through computer vision is presented in the suggested AI virtual mouse system. Because there are so few datasets available, it is challenging to compare testing of the AI virtual mouse system. For tracking of the hand gestures and hand tip detection, the webcam has been placed at various distances from the user in order to evaluate the hand gestures and finger tip detection in various lighting situations. The model is tested ,the AI virtual mouse system several times in bright light settings, in dim light situations, from close proximity to the webcam, and from a distance of four feet or more from the screen. The accuracy of the AI virtual mouse system was around 99 percent. We may conclude that the proposed AI virtual mouse system has performed successfully based on its 99 percent accuracy. Right click accuracy is poor because it is the most challenging gesture for computers to comprehend. Because the gesture required to carry out the specific mouse action is more difficult, the right click's precision is poor. Additionally, all other gestures have excellent and high precision. Our virtual mouse model performed extremely well, with 99 percent accuracy, as compared to earlier methods. When compared to existing virtual mouse models, it is clear that the suggested AI virtual mouse has outperformed them all in terms of accuracy. The new aspect of the suggested model is that it can virtually control a computer in a manner similar to a physical mouse, including performing most mouse actions including left and right clicks, scrolling up and down, and mouse pointer movement utilisingfinger tip recognition.

A. Gesture Controller Output:

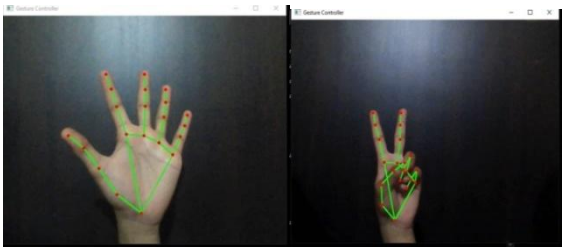


Fig.3. Neutral Gesture Fig.4. Moving Gesture

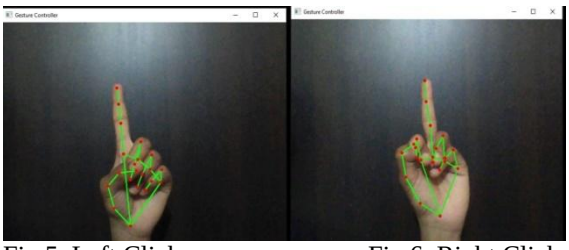


Fig.5. Left Click Fig.6. Right Click

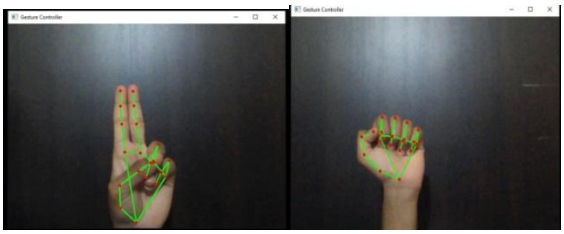


Fig.7. Double Click Fig.8.Drag & Drop

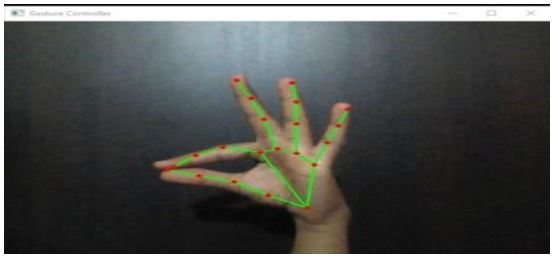
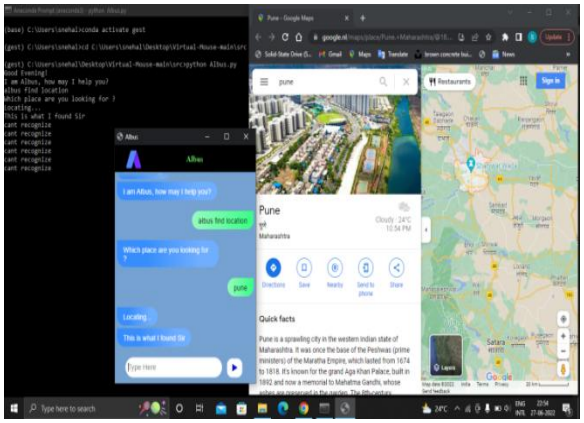


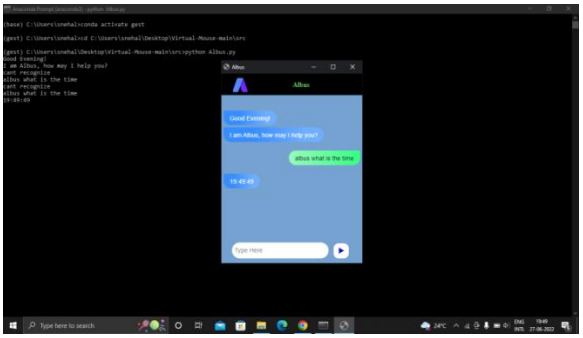
Fig.9. Volume Brightness Control

B. Voice Assistant Output:

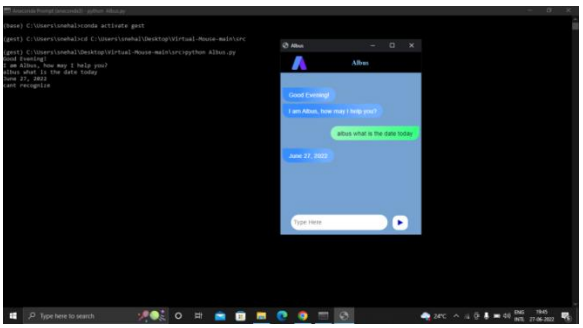


Command-"Albus Find a Location"

A location will be searched based on the input by the user for example: Hyderabad, Mumbai, etc. Afterwards, it will open Chrome/Firefox and search for the location on Google Maps.



Command- "Albus what is the time"



Command- "Albus what is the today date" it will return “Month Date, Year” eg: June 27, 2022

VII. CONCLUSION

In this project, we are working on a system to control the mouse cursor using a real- time camera.The system is based on computer vision algorithms and can do all Mouse tasks. However, it is difficult to get stable results because of the variety of lightning and skin colours of human races. Presentations would be easier with this system, and work space would be saved by using it.It provides features such as enlarging and shrinking Windows, closing window, etc. by using the Palm and multiple fingers.In order to make physically challenged people use desktops and laptops as smart as normal people, and consider the use of new edge technologies, this software is designed to make them as comfortable as the normal people.Python 3.7 (64-bit) and open source modules are used in the development of this project, making it suitable for future updates.

REFERENCE

[1] Abhik Banerjee, AbhirupGhosh, Koustuvmoni Bharadwaj,” Mouse Control using a Web Camera based on Color Detection”, IJCTT, vol.9, Mar 2014

[2] Angel, Neethu.P.S, ”Real Time Static & Dynamic Hand Gesture Recognition”, International Journal of Scientific & Engineering Research Volume 4, Issue3, March-2013.

[3] Chen-Chiung Hsieh and Dung-HuaLiou,” A Real Time Hand Gesture Recognition System Using Motion History Image”icsps, 2010

[4] Hojoon Park, “A Method for Controlling the Mouse Movement using a Real Time Camera”, Brown University, Providence, RI, USA, Department of computer science, 2008.

[5] Erdem, E. ardimci, Y. Atalay, V. Cetin, A. E. “Computer vision based mouse”, Acoustics, Speech, and Signal Processing, 2002. Proceedings.(ICASS). IEEE International Conference

[6] AbhirupGhosh, Abhik Banerjee “Mouse Control using a Web Camera based on Color Detection”

[7] J.L.Raheja, A.Chaudhary, K.Singal “Proposed using HSV algorithm but this uses special sensor is used to capture image and processes it” User has to spend more money for the sensor.

[8] Baohuaqiang, yijiezhai, mingliangzhou.” SqueezeNet and Fusion Network-Based Accurate Fast Fully Convolutional Network for Hand Detection and Gesture Recognition”. IEEE.2021.