Voice Assistant and Virtual Mouse using Different AI Techniques and ML

Kalipu Ravi Kumar[1], Shaik Kareem[2], Shaik Yaseen[3], S Syam Sundar Rao[4], T.A. Avadeesh [5]

Assistant Professor, Dept of CSE, Raghu Institute of Technology, Visakhapatnam, India[1]

Dept of CSE, Raghu Institute of Technology, Visakhapatnam, India[2,3,4,5]

Abstract: *This venture advocates a Human PC Cooperation strategy where cursor development can be controlled through ongoing camera by utilizing human hand stances acknowledgment. This strategy is an option in contrast to current procedures, which incorporate physically squeezing buttons or utilizing actual PC mouse. Instead, it manages various mouse events through camera and computer vision software and can perform any action that a standard computer mouse can. The Virtual Mouse color recognition software will continuously collect photos taken in real time and process them through a number of conversions and filters. Whenever everything is changed over, the application will utilize picture handling to extricate directions of determined variety position from changed over outlines. The process then looks at a list of color combinations, each of which corresponds to a specific set of mouse actions, and compares the current color schemes in the frames. On the off chance that the ongoing variety conspire matches, the application will play out the mouse order, which will be changed over into a genuine mouse order on the user\'s PC. Notwithstanding that creators have likewise fostered a voice right hand to further develop client efficiency by overseeing routine errands of the client and by giving data from online sources to the client since voice colleague is easy to utilize. The objective of this paper is to improve the virtual assistant's performance. In this way, what precisely is a remote helper. Software that understands natural language voice commands and can carry out tasks on your behalf is known as application software, which is also referred to as AI assistants or digital assistants. They are also referred to as virtual assistants. What is the role of a virtual assistant? The list of tasks that virtual assistants can carry out that you can carry out on your own smartphone or personal computer is constantly growing. Meeting scheduling, message delivery, and weather monitoring are just a few of the impressive tasks that virtual assistants typically carry out. Google Assistant and Cortana, two examples of previous virtual assistants, were limited in that they could only perform searches and were not entirely automated. To keep the song's control function intact, these engines, for instance, lack the ability to play and pause the song. They can only have the module to play and search for songs. In order to enhance the project's functionality, we are currently working on an automated Google, YouTube, and numerous other new tasks. Now, a virtual mouse that can only be used to control the cursor and click has been added to make the process easier.*

Keywords: Cricket Analytics; Detection, Gesture, Mediapipe, OpenCV, python

I. INTRODUCTION

The interest in the creation of systems for human-computer interaction has grown over the past few years. Students, engineers, and researchers are encouraged to concentrate on the development of natural, intuitive user interfaces by the availability of software packages that provide features for specific processing and make it simple to communicate between computers and computing devices. The use of them to control or interact with artificial systems is of interest to the research community. Non - verbal correspondence as motions is used to pass on a specific message. This message can be conveyed through a person's hands, face, or body movements. When interacting with other people, gestures can convey information. from extremely easy to extremely difficult hand movements. We don't need any additional computer hardware to use the current system, which allows us to control the mouse with a laptop or computer equipped with a web camera and microphone. A voice assistant is also used to complete additional tasks.

The development of technology has had a significant impact on the things we use every day. Because of advancements in technology, the way we live today and the way we lived in the past are very different. Computers were limited to a small number of tasks prior to the development of machine learning, artificial intelligence, and other technologies. As a direct consequence of this, a number of voice assistants, such as Cortana, Alexa, and Siri, have been developed. These voice assistants improve computers through speech recognition and eliminate the need for input devices. The AI virtual assistant includes numerous additional devices. like Amazon, Google, and Microsoft. The user can wake up or use the new model of Amazon's Echo virtual assistant by saying "Alexa." After that, a light appears on it, indicating that it is prepared to follow instructions and proceed. It can respond to a variety of commands, including "play music on Amazon Music or YouTube," "open Netflix," and "what is today's temperature." Remote helper innovations need a ton of information to work since they feed man-made reasoning (computer based intelligence) stages, including AI, discourse acknowledgment, and regular language handling. Programming for artificial intelligence makes use of sophisticated algorithms to learn from the user's data and improve its capacity to anticipate the needs of the user when they interact with a virtual assistant. The primary objective of the virtual mouse system is to substitute hand gestures for physical mouse control. The depicted framework might be carried out through a webcam or implicit camera that recognizes hand developments and hand tips and examinations these casings to perform explicit mouse exercises. The study reveals that the proposed artificial intelligence virtual mouse system has performed exceptionally well, is less delicate than the existing models, and successfully overcomes the majority of those models' drawbacks. When used regularly, the virtual mouse is useful. Because it eliminates direct touch, this project suggests a method for moving the cursor with just your hands, without the use of any devices, which will be useful in situations like COVID outbreaks. Operations like right and left clicks, double clicks, drag-and-drop, selection of multiple items, volume control, and other similar actions are supported by the proposed system. A webcam is the only necessary input device for the proposed system. To implement the suggested method, two pieces of software are required: Python and OpenCV. On the framework's screen, the camera's result will be shown so the client might change it further. NumPy, math, wx, and mouse are the python conditions that will be used to make this framework.

The other usefulness presented by this undertaking is voice associate. Text searches have lost ground to voice searches. The experts are as of now foreseeing that half of searches will be by means of voice by 2024. Voice aides are ending up being more intelligent than any time in recent memory. Voice Partner is useful to facilitate our everyday undertakings, for example, showing date and time, performing look through on google, finding any area on google maps, opening any application and so on. Voice or text-based commands are both accepted by the voice assistants. Before an order can be given by voice-based intelligent assistants, they need an activating word, also known as a wake word. Quantum is the wake word for this project.

With the advancement of technologies in augmented reality and devices that we use every day, Bluetooth or wireless technologies are making these devices smaller. A computer vision-based AI virtual mouse system based on hand gestures and tip detection is proposed in this paper for carrying out mouse functions in a computer. Instead of using a traditional mouse, the proposed system's primary goal is to use a web camera or a computer's built-in camera to perform scrolling and cursor functions. As a HCI, the computer uses computer vision to detect hand gestures and hand tips [1]. With the utilization of the computer based intelligence virtual mouse framework, we can follow the fingertip of the hand signal by utilizing an inherent camera or web camera and play out the mouse cursor tasks and looking over capability and move the cursor with it.

Some devices, like the mouse, the dongle that connects to the PC, and a battery that powers the mouse, are used when using a wireless or Bluetooth mouse; however, in this paper, the user controls the computer mouse with hand gestures and either their built-in camera or a webcam. In the proposed framework, the web camera catches and afterward processes the casings that have been caught and afterward perceives the different hand signals and hand tip motions and afterward plays out the mouse capability.

The AI virtual mouse system is developed using the Python programming language, and OpenCV, a computer vision library, is also utilized in the AI virtual mouse system. The model in the AI virtual mouse system that is being proposed uses the MediaPipe package to track the hands and the tip of the hands. Additionally, the Pynput, Autopy, and PyAutoGUI packages were used to move around the computer's window screen and perform functions like left click, right click, and scrolling. The proposed model's results showed a very high level of accuracy, and the model can work

very well in real-world applications with a CPU and no GPU. Computer technology has grown a lot in the last ten years and is now a necessary part of everyday life. The mouse is the most important computer accessory for Human Computer Interaction (HCI). In some real-world situations, such as with Human-Robot Interaction (HRI), the mouse is not suitable for HCI. There have been many explores on elective strategies to thecomputer mouse for HCI. Utilizing hand gestures is the most natural and intuitive method for HCI and a viable alternative to the computer mouse. Therefore, the goal of this project is to investigate and develop a hand gesture-based Computer Control (CC) system. Webcams, which are now included in the majority of laptops, have recently been utilized in face recognition-based security applications. To outfit the maximum capacity of a webcam, it tends to be utilized for vision based CC, which would successfully take out the requirement for a PC mouseor mouse cushion. Other human-computer interaction (HCI) applications, such as a sign language database or motion controller, can benefit greatly from a webcam's versatility. The Microsoft Kinect and Nintendo Wii are two examples of HCI technologies that have made significant progress over the past few decades for use in gaming. Playing video games more naturally and interactively is made possible by these gaming technologies. Movement controls is the fate of gaming and it have massively boostedthe deals of computer games, for example, the Nintendo Wii which sold more than 50 million control center inside ayear of its release. HCI utilizing hand signals is extremely instinctive and powerful for one toone collaboration with PCs and it gives a Characteristic UI (NUI). The use of hand gestures to control a cursor has been the subject of extensive research. Other than HCI, hand motion acknowledgment is likewise utilized in communication through signing acknowledgment, which makes hand motion acknowledgment much more critical

II. SYSTEM ANALYSIS

2.1 Existing Analysis
The mouse that controls the cursor in the current system can be wireless or wired, and we can use hand gestures to monitor the system. Colored fingers act as an object that the webcam senses color like red, green, or blue to monitor the system. On the other hand, the skin color recognition system can perform basic mouse operations like minimize, drag, scroll up, scroll down, and left-click right-click using hand gestures without the use of colored fingers. This is because the current virtual mouse control system uses colored tips for detection. The current system uses static hand recognition, such as fingertip identification, hand shape, and the number of fingers, to explicitly define an action, making the system more difficult to use and more difficult to comprehend.

2.2 Proposed System
The system works by determining the position of the cursor based on the color of the hand. However, a variety of conditions and scenarios make it difficult for the algorithm to work in the real world, as shown in Fig. 1. The environment's sounds. Various skin textures; environmental lighting conditions. A skin-colored background object A gesture-controlled mouse

Step-1: Camera Initialization: Since the user's finger moves through the camera window, an image is used by the camera, so it's important to match the resolution of the system screen to that of the camera window. As a result, determining the number of pixels moved on the actual screen when the user moves one pixel in the camera window becomes critical. This is required so that the user's finger can completely cover the screen while the camera is taking a picture. The screen's resolution can be estimated by using the commands.

Step-2: All hand gestures are mapped and extracted: The acquisition toolkit will use all of the various gestures to capture the user's moving finger. The camera's capture rate must be at least 15 frames per second.

Step-3: Correlating the Resolution of the Screen and Camera Window: Since the cursor will get across the whole screen are apportioned parallel numbers for planning. Mediapipe bundle is utilized for extraction of various milestones of the hand which is utilized to perceive signals. The subtraction technique, which is simple and effective for locating the hand region in an image, is used for hand detection. The twofold numbers given are valuable in fingers and palm division. With the help of the Mediapipe module's landmark detection, the palm point is defined as the palm's center, which can also be found using the distance transform method. The distance transform image keeps track of how far away each pixel is from the closest border pixel. Double qualities given to various motions are:

Copyright to IJARSCT

www.ijarsct.co.in

DOI: 10.48175/IJARSCT-9351

ISSN
2581-9429
IJARSCT

243

FIST = 0 | PINKY = 1 | RING = 2 | MID = 4 | LAST3 = 7 | INDEX = 8 | FIRST2 = 12 | LAST4 = 15 | THUMB = 16 | PALM = 31 | V_GEST = 33 | TWO_FINGER_CLOSED = 3 | PINCH_MAJOR = 35 PINCH_MINOR = 36

Step-4: Hand gestures can be recognized and used: Finally, meaningful hand gestures are recognized and performed following successful hand landmark mapping and extraction. The OpenCV and Mediapipe modules are utilized for this. Hand gestures can be anticipated based on the number and types of fingers present. The content of the fingers determines which fingers are detected. The recognition of five fingers, for instance, is categorized as a neutral gesture; the recognition of two fingers is categorized as a move cursor gesture; the right finger is categorized as a right click; and so on.

Voice Assistant The process began with a microphone-based analysis of the user's audio commands. This could be doing anything from getting information to using the computer's internal files to anything else.

The voice assistant is demonstrated by the algorithm that follows.

Step-1: Speech to Speech: To convert speech input into text, the system makes use of Google's online speech recognition system. The corpora organized on the computer network server at the information center store the user's speech input temporarily so that they can access it. discourse acknowledgment framework, which is hence moved to find out about cloud. The equivalent text is then provided to the central processor, which receives it.

Step-2: Python Endpoint: After receiving the speech recognition module's output, the Python backend determines whether the command or speech output is an API Call, Context Extraction, or System Call. After that, the output is sent back to the Python backend so that the user can get the output they need.

Step-3: Extracting the Context: It is the process of automatically extracting structured information from machine-readable documents that are either unstructured or semi-structured. The majority of this activity involves processing documents written in human languages with the help of natural language processing. Late improvements in the handling of sight and sound records, for example, programmed explanation and content extraction from photographs, sound, and video, can be viewed as setting extraction test results.

Step-4: System messages: A system call is a method that a computer program uses to programmatically request a service from the kernel of the operating system it is running on. Starting and running new processes, interacting with hardware-related services like accessing a hard drive, and communicating with essential kernel services like process scheduling are all examples of this. System calls provide a crucial interface between a process and the operating system.

Step-5: Text-To-Speech: The capability of computers to read text aloud is referred to as text-to-speech (TTS). A TTS Engine turns written text into a phonemic representation, which is then turned into waveforms that can be made into sound. Composed message is changed over by a TTS Motor into a phonemic portrayal, which is in this manner converted into waveforms that can be yield as sound.

- Camera: The video capture object is created using the Python computer vision library OpenCV, and the webcam will begin recording video. The webcam then, at that point, catches the edges and passes the ones to the simulated intelligence virtual machine.
- Video Catch: Identifying Which Finger is Up: At this point, it is our responsibility to perform the mouse function in accordance with the information in the library by identifying which finger movements and their respective coordinates are up. Monitoring which finger is up in light of the tip Id creators found through the MediaPipe and its related directions, as well as doing the essential activities, are expected as of now. The mouse behaves accordingly. A loop is set up to track where our hands are every 0.1 seconds.
- Hand gesture: Mouse Features Hand Gestures and Hand Tip Detection Using Computer Vision Cursor Moving: Using the Python AutoPy module, the mouse pointer moves across the window of the computer while the index and middle fingers are raised.
- Click the Left Button: By widening your index finger and dragging it onto the file or folder, you can click the left button.
- Click the right button: Wide opening the center finger and hauling it onto the record or envelope to investigate more exercises on it empowers the right button to click.
- Look Up/Down: By connecting the tips of the index and middle fingers to the scroll bar on the screen, the computer is designed to perform the mouse-like action of scrolling up and down with a single pinch.
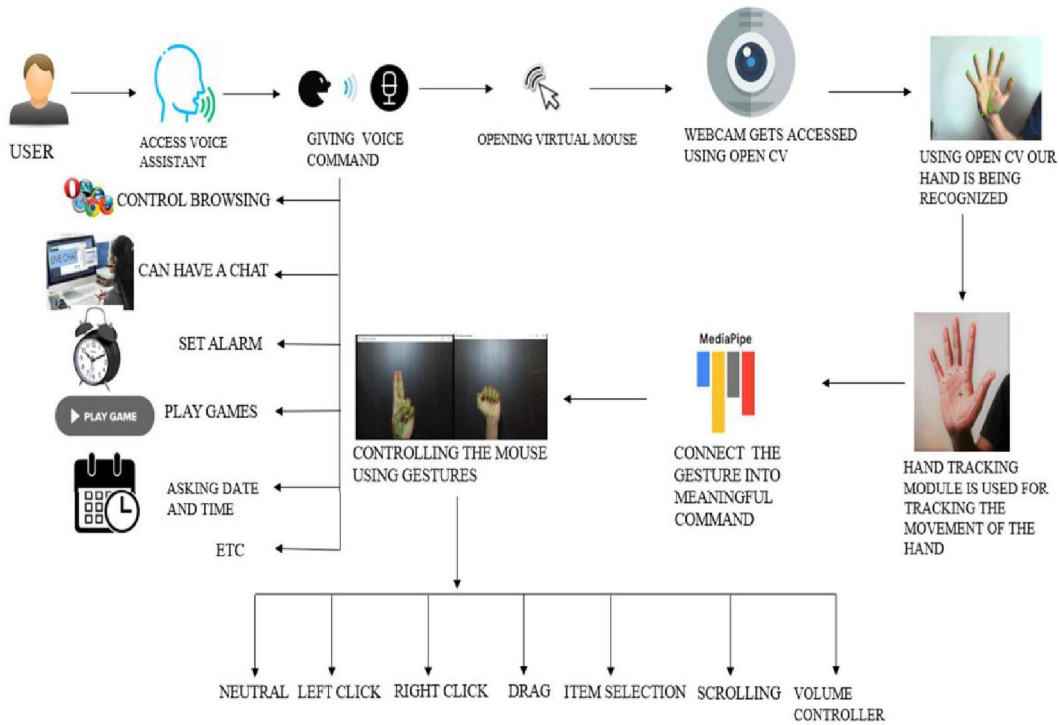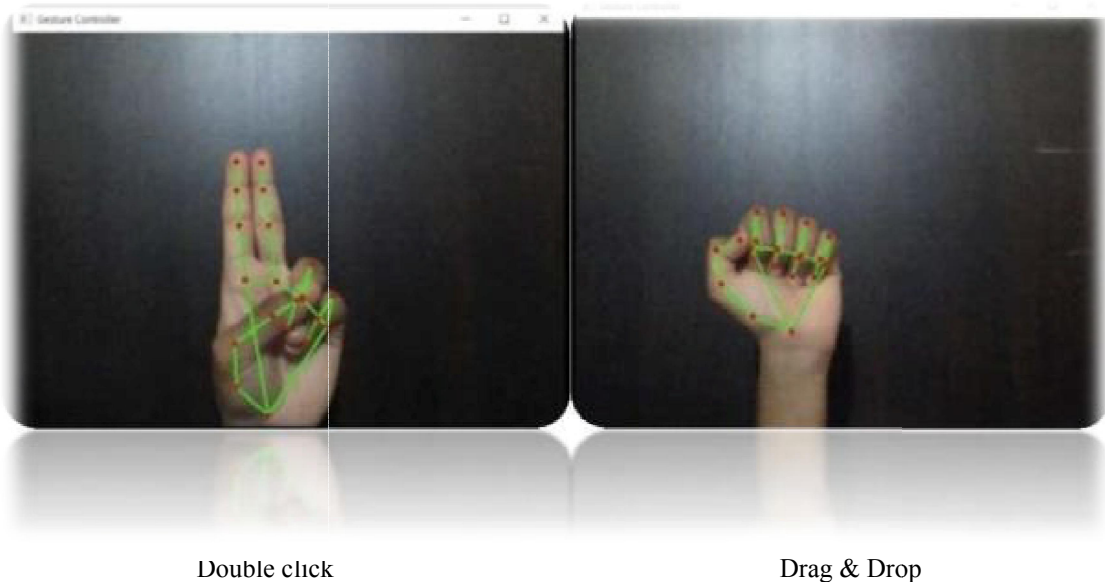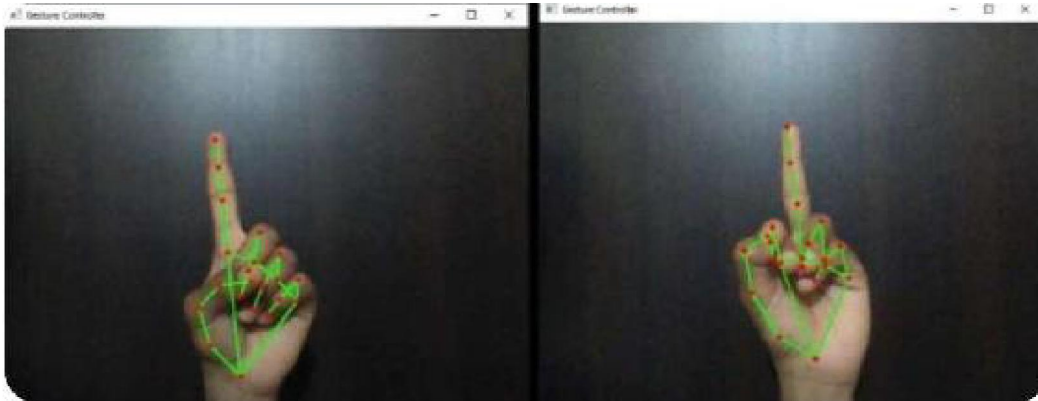
Fig: ARCHITECTURE

Virtual Mouse

Double click   Drag & Drop

Left Click   Right Click

Volume Control

Voice Assistant

III. CONCLUSION

We are utilizing the voice assistant and the virtual mouse for this project. Our goal is to create an effective voice assistant that completes tasks quickly. In this we added some computerization like Google Mechanization, YouTube Robotization to make completely menial helpers which clients utilized effectively and actually. Regarding the virtual mouse, we conclude that we designed one that operates solely with our fingers and does not require a touchpad or mouse ball. By utilizing it we complete all the necessary work with this like clicking, looking over, moving cursor, and so on. Python 3.7 and 3.9 and open-source modules are utilized in the advancement of this task, making it reasonable for future updates

Movement affirmation strategy gives the positive signs for performing different mouse capabilities. The usage of YCbCr concealing model in research eliminates the dependence on light power during test. This method, which can be used with cell phones for quite some time, is extremely useful for disabled people at the time of show by using IP cameras, controlling TV channels, and so forth. Better methods for using the mouse and reducing the amount of slack felt when moving the cursor will be included in future works.

The principal objective of the simulated intelligence virtual mouse framework is to control the mouse cursor capabilities by utilizing the hand signals as opposed to utilizing an actual mouse. Using a webcam or a built-in camera that recognizes hand gestures and the tip of the hand and processes these frames to carry out the specific mouse functions is one way to implement the proposed system.

We can conclude from the model's results that the proposed AI virtual mouse system performs exceptionally well, is more accurate than the existing models, and overcomes most of the existing systems' limitations. The AI virtual mouse can be used for real-world applications because the proposed model is more accurate. It can also be used to reduce the spread of COVID-19 because the proposed mouse system can be used virtually using hand gestures instead of a physical mouse.

The model suffers from a few drawbacks, such as a slight decrease in right-click accuracy and some difficulties when selecting text via drag and drop. As a result, the next step will be to work on improving the finger tip detection algorithm to produce more accurate results in order to overcome these limitations.

Voice-Controlled Gadgets utilizes Normal Language Handling to handle the language expressed by the human and figure out the inquiry and interaction the question and answer the human with the outcome. The comprehension of the gadget implies Man-made reasoning should be coordinated with the gadget so the gadget can work in a brilliant manner and can likewise control IoT applications and gadgets and can likewise answer question which will scan the web for results and cycle it. It is intended to control the device solely with human voice, minimizing human involvement. Other intelligent voice-controlled devices, such as Internet of Things (IoT) applications and devices, can interact with the device, as can sending an email to a client, adding events to the calendar, and so on. Using machine learning, queries can be grouped into specific result sets and used in subsequent queries to improve the accuracy of the devices. In the last ten years, the devices' accuracy has increased exponentially. Additionally, the devices can be made to respond to user commands in the user's native language and accept commands in either language.

IV. FUTURE SCOPE

This paper was written with the intention of making the machine more interactive and receptive to human behavior. The development of a technology that is affordable and adaptable to any standard operating system was the sole objective of this paper. The mouse pointer is controlled by detecting the human hand and moving the pointer in that direction, respectively, using the proposed system. the system controls basic mouse actions like dragging, left-clicking, and moving the cursor. When the angle between the fingers of a human hand is less than 15 degrees, the method performs the left-click function. The purpose of this work is to develop PC understanding projects that are able to deal with ongoing issues and achieve the goals of associations as well as individuals' everyday lives and daily activities. There is a degree in encouraging the machines games, language revelation, PC vision, ace systems, advanced mechanics, and so on. The more you know about AI sciences like science or material science, the better. Learn about mind science and the concrete framework for common management strategies for AI. Learn a few Machine expressions. Concentrating on a single important machine language is smart. Occupations are regularly to depend subsequent to getting the

programming vernaculars. Calling choices in simulated intelligence where understudy can land positions at Occupation will be offered like: Game Developer, Mechanical, Researcher, PC Researcher and information investigator.

REFERENCES

[1]. S.Shriram" Deep Learning-Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19 Spread", Applied Sciences, Volume 2021, Article ID 8133076, 1 Jul 2021, View at: Publisher Site |Hindawi.

[2]. Akshay L Chandra" Mouse Cursor Control Using Facial Movements — An HCI Application, View at: Publisher Site| data scienceAbdul Wahid Ali " Energy efficieny in routing protocol and data collection approaches for WSN: A survey, View at: Publisher Site|ieee explore AynurUnal " Smart Trends in Information Technology and Computer Communications, View at:Publisher Site|link.springer.com

[3]. S. U. Dudhane, "Cursor control system using hand gesture recognition," IJARCCE, vol. 2, no. 5, 2013. View at: Publisher Site| Google scholoar K. H. Shibly, S. Kumar Dey, M. A. Islam, andS. Iftekhar Showrav, "Design and development of hand gesture based virtual mouse," in Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1–5, Dhaka, Bangladesh, May 2019.View at: Publisher Site| Google Scholoar

[4]. A.Haria, A. Subramanian, N. Asokkumar, S. Poddar, and J. S. Nayak, "Hand gesture recognition for human computer interaction," Procedia Computer Science, vol. 115, pp. 367–374, 2017. View at: Publisher Site| Publish site |Scholoar

[5]. https://www.tutorialspoint.com› open CV

[6]. Google, MP, https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.

[7]. J. Katona, "A review of human–computer interaction and virtual reality research fields in cognitive InfoCommunications," Applied Sciences, vol. 11, no. 6, p. 2646, 2021. View at: Publish site|GoogleScholor

[8]. V. Bazarevsky and G. R. Fan Zhang. On-Device, MediaPipe for Real-Time Hand Tracking

Copyright to IJARSCT
www.ijarsct.co.in

DOI: 10.48175/IJARSCT-9351

ISSN
2581-9429
IJARSCT

248