HAND GESTURE CONTROLLED VIRTUAL MOUSE USING ARTIFICAL INTELLIGENCE

Kavitha R[1], Janasruthi S U[2], Lokitha S[3], Tharani G[4]

[1]*Assitant Professor, Department of Computer Science and Engineering, Bannari Amman Institute of Technology, TamilNadu, India.*

[2 3 4]*Student, Department of Computer Science and Engineering, Bannari Amman Institute of Technology, TamilNadu, India.*

ABSTRACT

The use of hand gesture recognition in controlling virtual devices has become popular due to the advancement of artificial intelligence technology. A hand gesture-controlled virtual mouse system that utilizes AI algorithms to recognize hand gestures and translate them into mouse movements is proposed in this paper. The system is designed to provide an alternative interface for people who have difficulty using a traditional mouse or keyboard. The proposed system uses a camera to capture images of the user's hand, which are processed by an AI algorithm to recognize the gestures being made. The system is trained using a dataset of hand gestures to recognize different gestures. Once the gesture is recognized, it is translated into a corresponding mouse movement, which is then executed on the virtual screen. The system is designed to be scalable and adaptable to different types of environments and devices. All the input operations can be virtually controlled by using dynamic/static hand gestures along with a voice assistant. In our work we make use of ML and Computer Vision algorithms to recognize hand gestures and voice commands, which works without any additional hardware requirements. The model is implemented using CNN and mediapipe framework. This system has potential applications like enabling hand-free operation of devices in hazardous environments and providing an alternative interface for hardware mouse. Overall, the hand gesture-controlled virtual mouse system offers a promising approach to enhance user experience and improve accessibility through human-computer interaction.

Keyword: *Computer vision, hand gesture recognition, Media-pipe, virtual mouse.*

1. INTRODUCTION

The world is full of technology driven factors in our day to day life. We have so many technologies, throughout the world computer technologies are growing simultaneously. They are used to perform various tasks which cannot be performed by humans. In fact they are ruling the human lives because they have a potential to do the tasks which cannot be done by humans. The interaction between human and computer can be done with output device like mouse. The mouse is a device used for interacting with a GUI which includes pointing, scrolling and moving etc. The hardware mouse in computers and touchpads in laptops will require a huge amount of time to perform complex tasks, incase we are carrying hardware mouse wherever we go it would be damaged sometimes.

After decades the technology has made the mouse functionality from wired into the wireless to improve the functionality and for the easy movements in hassle free manner. As the technologies started growing there came the speech recognition technique. This recognition is mainly used for the voice recognition purpose for searching something with the help of their voice and for translation purposes but it can take time for recognition to perform mouse functions. Later the human computer interaction evolved with the eye tracking techniques for controlling the

cursor of the mouse. The major drawback of this technique is that some may wear contact lens or some may have long eyelashes so it may take some time to capture their eye movement.

Different types of attempts taken by many developers for developing the models for human gesture recognition. Those models require expensive gloves and sensors for capturing and color cap for marking the positions of the fingertips. The technologies are still emerging, one of the vast technologies artificial intelligence is playing a major role in every sector. Artificial intelligence makes human life fast and comfortable. To overcome the problems faced in the existing approaches we are going for the latest algorithms and tools in artificial intelligence.

Hand gesture controlled virtual mouse using artificial intelligence is a technology that allows users to control the movement of their computer mouse using hand gestures, without the advent of a physical mouse. This technology uses a camera vision based approach to track the movements of the user's hand and to perform mouse functions on the computer screen. The system works by capturing video input from a camera pointed at the user's hand. The computer vision algorithms then analyze the video feed to identify the user's hand and track its movement. This information is given to machine learning models which have been trained to recognize specific hand gestures, such as pointing or swiping, and translate them into corresponding mouse movements.

This latest super cool technology has various advantages, including its potential to improve accessibility for people and its ability to provide a more natural and intuitive user experience. It can also be useful in situations where a physical mouse or touchpad is not available or practical. The use of hand gestures as a control mechanism eliminates the need for a physical mouse and provides a more intuitive and natural way of interaction with computers. This technology has numerous applications in areas such as gaming, virtual reality and accessibility quite easy for people.

2. LITERATURE REVIEW

Some work which is related to the AI virtual mouse had been performed previously in that glove were used by the user to recognize and collect data from the system. Later another system used colored pieces of paper which are attached on hands for gesture recognition. But these systems are not very feasible for performing mouse operations accurately. In a glove based approach recognizing the gloves is not viable and it might be allergic for users who have sensitive skin type. Also wearing gloves for a long time is difficult. It might sweat and result in skin rashes and allergic reactions. In the case of colored tips for gesture recognition and detection will not always give best results. Now some others have made contributions that use Google's work with the mediapipe framework. The current gesture controlled virtual mouse uses hand gestures to perform mouse functions, in which we have control over the mouse cursor and perform certain mouse operations like left click, right click, drag and drop, volume control and brightness control etc. Efforts have been made for hand gesture recognition with camera-based detection of the hand gesture interface.

[1] This review is about how a hardware-based system is developed. Although this model produces incredibly accurate results, many movements are challenging to execute while wearing a glove that severely limits the user's hand's range of motion, speed, and agility. Also wearing gloves for a long time will result in skin diseases and is not best suited for the users with sensitive skin type. [2] They created a machine-user interface that uses straightforward computer vision and multimedia techniques to accomplish hand gesture detection. However, a significant disadvantage is that skin pixel identification and hand segmentation from stored frames must be completed before working with gesture comparison techniques. [3] They described a system in this study for recognizing hand movements that relies on a mobile phone's camera and a connected mobile projector as a visual feedback medium. Other mobile applications can easily link to their framework to learn gesture recognition. The suggested architecture enables the quick and simple creation of research prototypes that support gestures, diverting the user's focus away from the device and towards the content.[4] A method for performing mouse functions without any electrical equipment like sensors. It requires a webcam alone. And mouse functions like clicking and dragging files are carried out through hand gestures. The suggested model performance is low with accuracy and lacks more mouse functionality. [5] This study focuses on the advanced study of robots with gesture controls. The first section gives an

idea of the art for hand gesture identification as it relates to how they are seen and captured by common video cameras. Based on estimations of the smoothed optical flow, we extract a collection of motion features. Face detection is used to produce a user-centric representation of this data, and an effective classifier is trained to differentiate. [7] In this model the hand's center is determined, and the hand's calculated radius is discovered. And using the convex hull technique, fingertip points have been determined. The hand gesture is used to control every mouse movement. And the problem of this approach is the frame must first be saved before being processed for detection, which takes longer than what is needed in real-time. [9] The vision based technique has been tried out in this system. Utilized a webcam for gesture recognition and detection. And no external devices like sensors and gloves were used. Completely focuses on leveraging the YOLOv5 algorithm and Artificial Intelligence (AI) to recognize hand gestures and improve HCI. [10] The system can create coloured masks utilizing techniques for color variation. Later mouse functions are carried out using hand gestures. This approach is difficult in its implementation.

3. ALGORITHMS AND TOOLS USED

For the purpose of hand and finger detection we are using the one of the effective open source library mediapipe, it is one type of the framework based on the cross platform features which was developed by google and Opencv to perform some CV related tasks. This algorithm uses machine learning related concepts for detecting the hand gesture and to track their movements.

3.1 Mediapipe

Google created the open-source MediaPipe framework to enable the development of cross-platform, real-time computer vision applications. For processing and analyzing video and audio streams, it offers a number of pre-made tools and components, such as object detection, pose estimation, hand tracking, facial recognition, and more.

Developers can quickly construct intricate pipelines using MediaPipe that combine numerous algorithms and processes and execute in real-time on a variety of h/w platforms, like CPUs, GPUs, and specialized accelerators like Google's Edge TPU. Additionally, the framework has interfaces helps us interacting with other well-liked machine learning libraries, including TensorFlow and PyTorch, and it supports several programming languages, like C++, Python, and Java.

For computer vision and ML tasks, MediaPipe is a comprehensive library that offers a many of features. Here are a few of the library's main attributes and features:

1. Video and Audio Processing: MediaPipe provides tools for processing and analyzing video and audio streams in real-time. This includes functionalities such as video decoding, filtering, segmentation, and synchronization.

2. Facial Recognition: MediaPipe can detect and track facial landmarks, including eyes, nose, mouth, and eyebrows, in real-time. This functionality is useful for applications such as facial recognition, emotion detection, and augmented reality.

3. Hand Tracking: MediaPipe can track hand movements in real-time, allowing for hand gesture recognition and interaction with virtual objects.

4. Object Detection: MediaPipe can detect and track objects in real-time using machine learning models. This functionality is useful for applications such as augmented reality, robotics, and surveillance.

5. Pose Estimation: MediaPipe can estimate the poses of human bodies in real-time, allowing for applications such as fitness tracking, sports analysis, and augmented reality.

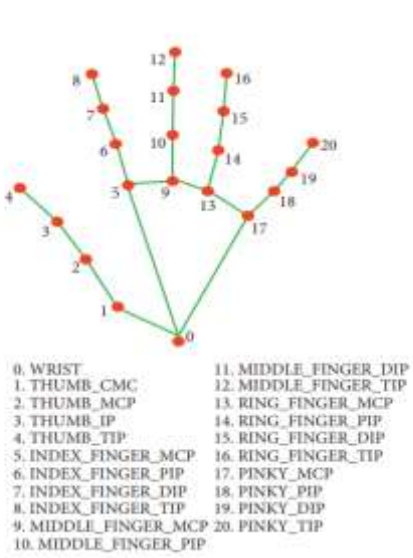

Fig -1: Hand Coordinates or Landmarks

For a variety of tasks, such as object detection, position estimation, facial recognition, and more, MediaPipe offers tools for training and deploying machine learning models. All in all, MediaPipe is a potent tool kit that gives programmers the ability to easily create sophisticated real-time computer vision and ML applications.

3.2 Opencv

A computer vision and ML software library called OpenCV is available for free download. Its objective is to aid programmers in the development of computer vision applications. Filtering, feature identification, object recognition, tracking, and other processing operations for images and videos are all available through OpenCV. Python, Java, and MATLAB are just a few of the numerous programming languages that it has bindings for. It is written in C++.Robotics, self-driving cars, AR, medical image analysis, and other fields are just a few of the fields where OpenCV can be employed. A wide range of algorithms and tools are included in the library, making it simple for programmers to build sophisticated computer vision applications.

The steps listed below can be used to broadly classify OpenCV's operation:

1. Loading and Preprocessing the Image/Video: OpenCV can load images or videos from a variety of sources such as files, cameras, or network streams. Once the image or video is loaded, it can be preprocessed by applying filters or transforming the image to a different color space, such as converting a color image to grayscale.

2. Feature Detection and Description: OpenCV can detect and extract features from an image or video, such as edges, corners, and blobs. These features can be used to identify objects or track their motion over time. OpenCV also provides algorithms for describing these features, which can be used to match them across multiple frames or images.

3. Object Detection and Recognition: OpenCV can be used to detect and recognize objects in an image or video. This can be done using a variety of techniques, such as template matching, Haar cascades, or deep learning-based methods.

4. Tracking: OpenCV can track objects in a video stream by estimating their position and motion over time. This can be done using a variety of algorithms, such as optical flow, mean-shift, or Kalman filtering.

5. Image and Video Output: Finally, OpenCV can be used to display or save the processed images or videos. This can be done by showing the images in a window, writing the video frames to a file, or streaming the video over a network.

In general, OpenCV offers a large variety of tools and techniques for working with image and video data, making it a potent library for computer vision applications.

4. PROPOSED METHODOLOGY

4.1. The Camera Used in the AI Virtual Mouse System

The proposed system uses web camera for capturing images or video based on the frames. For capturing we are using CV library Opencv which is belongs to python web camera will start capturing the video and Opencv will create a object of video capture. To AI based virtual system the frames are passed from the captured web camera.

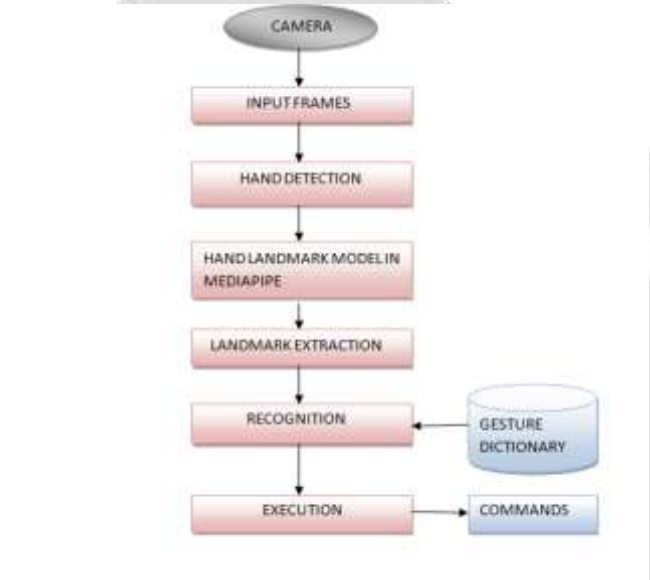

Fig -2: Flow graph of Hand Gesture Recognition

4.2. Capturing the Video and Processing

The capturing of the frame was done with the AI virtual mouse system until the program termination.Then the video captured has to be processed to find the hands in the frame in each set. The processing takes places is it converts the BRG images into RGB images,which can be performed with the below code,

```
image = cv2.cvtColor(cv2.flip(image, 1), cv2.COLOR_BGR2RGB)
image.flags.writeable = False
results = hands.process(image)
```

This code is used to flip the image in the horizontal direction then the resultant image is converted from the BRG scale to RGB scaled image.

4.3. Rectangular Region for Moving through the Window

The windows display is marked with the rectangular region for capturing the hand gesture to perform mouse action based on the gesture. when the hands are find under those rectangular area the detection begins to detect the action based on that the mouse cursor functions will be performed. The rectangular region is drawn for the purpose of capturing the hand gestures through the web camera which are used for mouse cursor operations.

Mouse Functions Depending on the Hand Gestures and Hand Tip Detection Using Computer Vision:

- For the Mouse Cursor Moving around the Computer Window.

Fig -3: Computer Window with Mouse Cursor

- To Perform Left Button Click operation

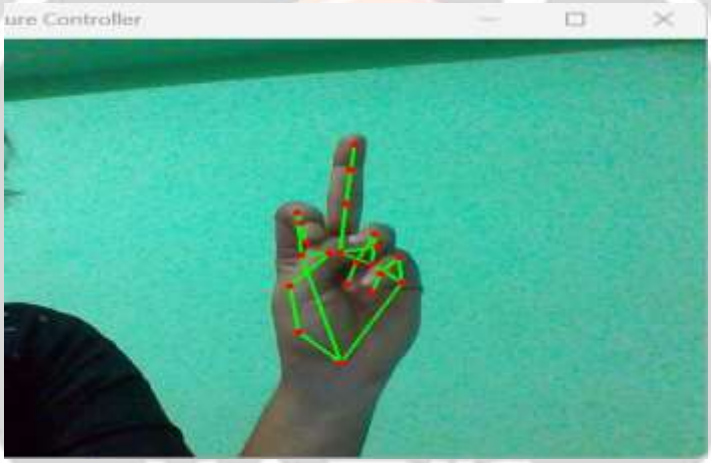

Fig -4: Mouse Operation-Left Click

- To Perform Right Button Click operation

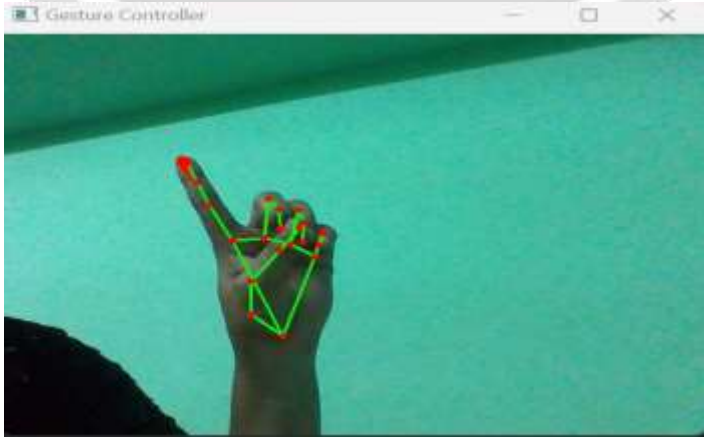

Fig -5: Mouse Operation-Right Click

- To perform a double click operation

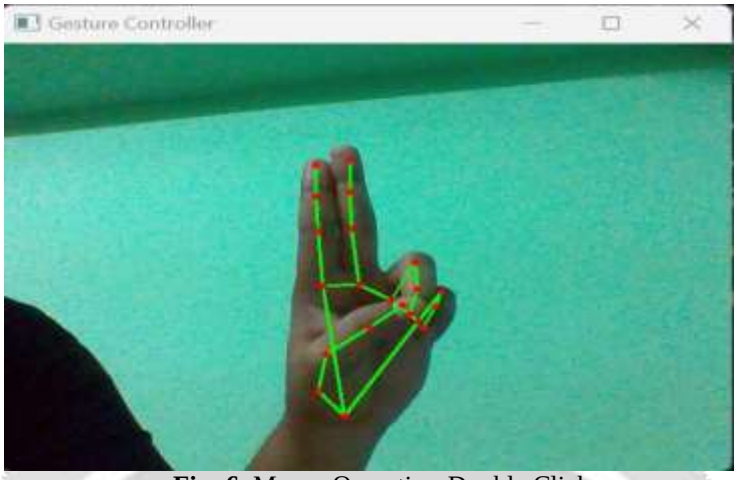

Fig -6: Mouse Operation-Double Click

- To perform scrolling operation

Fig -7: Mouse Operation-Scrolling

- To perform drag and drop operation

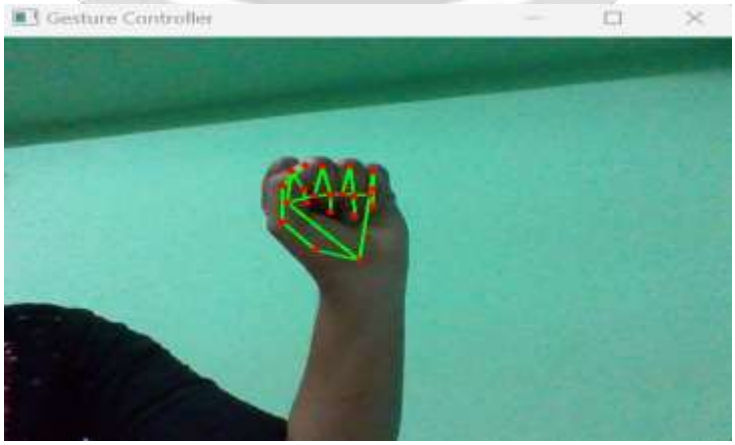

Fig -8: Mouse Operation- Drag and Drop

- To perform multiple item selection

Fig -9: Mouse Operation-Multiple Item Selection

- To perform volume controlling

Fig -10: Mouse Operation-Volume Control

- To perform brightness controlling

Fig -11: Mouse Operation-Brightness Control

- For No Action / neutral gesture to be Performed on the Screen

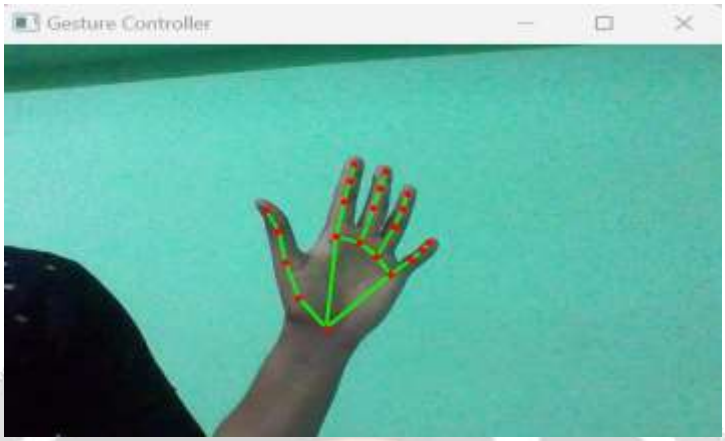

Fig -12: Neutral Gesture

4.4 Voice Assistant Features

The voice assistant feature has been included to launch gesture recognition through voice commands. And added certain features to improve the user engagement and they can assess whatever they need with less amount of effort and in hassle free manner. The voice assistant features which can be performed through the voice commands are:

- To launch and end the gesture recognition

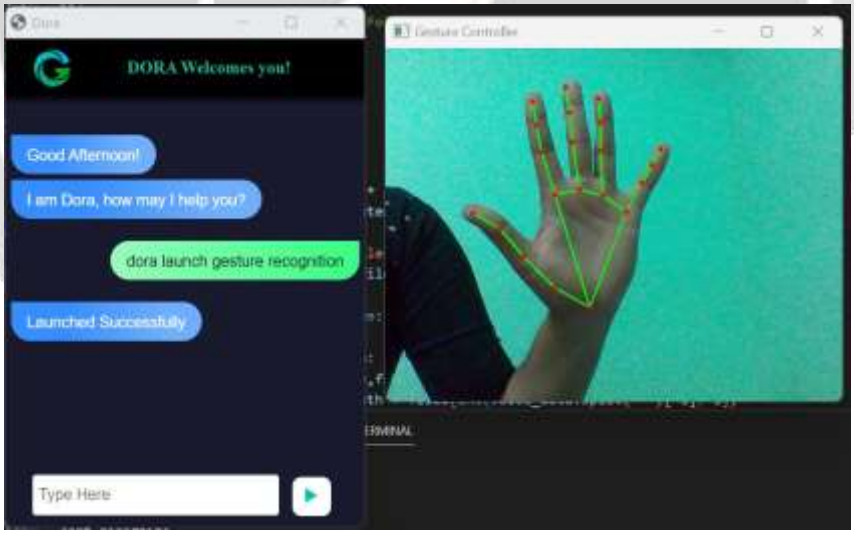

Fig -13: Launch and End

- To search for something over internet

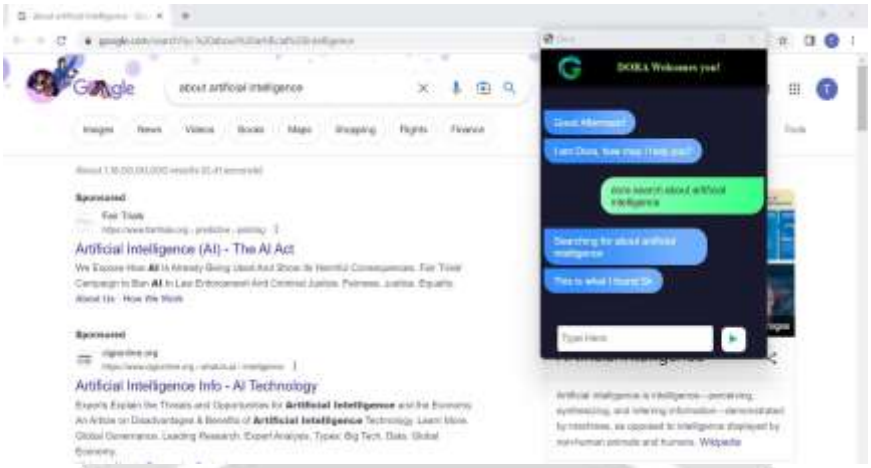

Fig -14: Voice Assistant Search

- To find a location what we are looking for

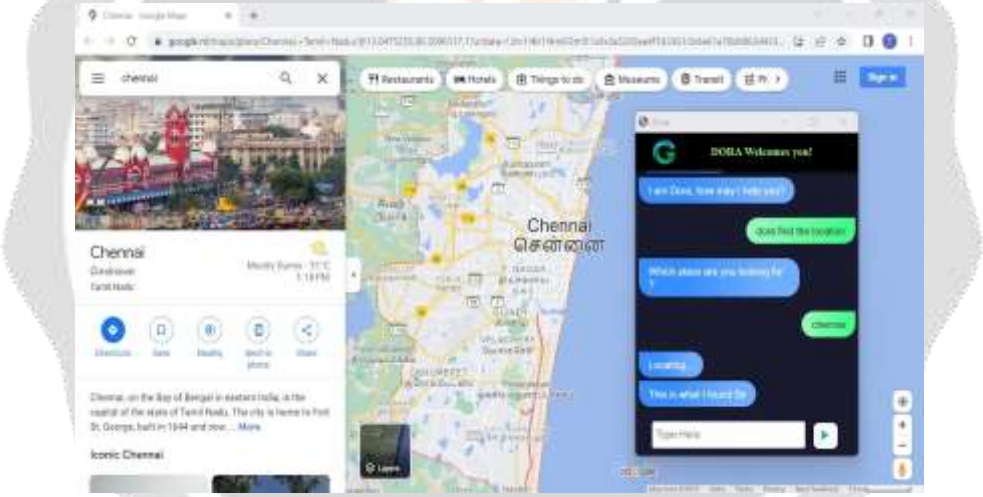

Fig -15: Find a Location

- To get an idea about Date and time

Fig -16: Date and Time

- To copy and paste contents

Fig -17: Copy and Paste

- Sleep / wakeup the voice assistant

Fig -18: Sleep or Wakeup

- To exit voice assistant

Fig -19: Exit Voice Assistant

5. RESULT AND INFERENCES

A hand-gesture- controlled virtual mouse could provide an alternative method for people with disabilities who may have difficulty using a traditional mouse or keyboard. This technology can make it easier for them to interact with computers and other devices. A hand gesture-controlled virtual mouse could also be useful for people who prefer to work or play games without being tethered to a physical mouse touchpad. This model would allow them to control their devices without the need for a physical interface.

Depending on the technology used, a hand gesture-controlled virtual mouse may offer a higher degree of accuracy and speed than traditional mice or video editing. The success of this technology will depend on the user experience it provides. If the technology is easy to use, reliable and provides an intuitive interface, likely to be well-received. However, if the technology is difficult to use, unreliable, or unintuitive , users may quickly abandon it.

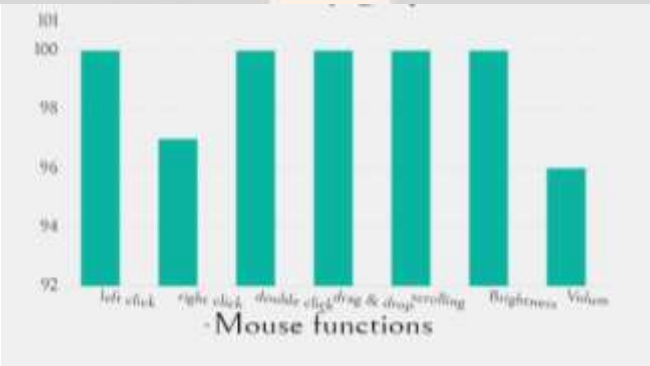

Chart -1: Accuracy For Mouse Function

Overall, the hand gesture-controlled virtual mouse using artificial intelligence has the potential to improve the accessibility of and convenience of computer interaction for users with physical disabilities or for users who prefer an alternative to traditional input devices.

6. CONCLUSION

AI virtual mouse using hand gestures is an innovative and exciting technology that has the potential to revolutionize the way we interact with computers. Here with the aid of a real-time camera, we have created a system to manage the mouse pointer and carry out its function. It offers users a more natural, intuitive, and accessible way to control the cursor on the screen, without the need for a traditional input device, a mouse.

Furthermore, with additional voice assistant support, AI virtual mouse using hand gestures can further enhance the user experience. Voice assistant which is integrated with the virtual mouse system will provide users with even more control over their devices. Users can given voice commands to do a range of tasks, such as opening applications, navigating through menus, and performing web searches, in addition to controlling the cursor on the screen using hand gestures. As technology continues to evolve, we can expect to see even more innovative solutions that enhance the user experience and improve accessibility for all.

7. REFERENCES

[1]. Quam, D.L., et.al. (1990). Gesture Recognition with a Dataglove. In IEEE conference on Aerospace and Electronics (pp. 755-760).

[2] Guoli Wang., et.al. (2015). Optical Mouse Sensor-Based Laser Spot Tracking for HCI input, Proceedings of the Chinese Intelligent Systems Conference (pp. 329-340).

[3] Baldauf, M., and Frohlich, p. (2013). Supporting Hand Gesture Manipulation of Projected Content with mobile phones. In the European conference on computer vision (pp. 381-390).

[4] Roshnee Matlani., Roshan Dadlani., Sharv Dumbre., Shruti Mishra., & Abha Tewari. (2021). Virtual Mouse Hand Gestures. In the International Conference on Technology Advancements and innovations (pp. 340-345).

[5] Mayur, Yeshi., Pradeep, Kale., Bhushan, Yeshi., & Vinod Sonawane. (2016). Hand Gesture Recognition for Human-Computer Interaction. In the international journal of scientific development and research (pp. 9-13).

[6] Shriram, S., Nagaraj, B., Jaya, J., Sankar, S., & Ajay, P. (2021). Deep Learning Based Real-Time AI Virtual Mouse System Using Computer Vision to Avoid COVID-19 Spread. In the Journal of Healthcare Engineering (pp. 3076-3083).

[7] Steven Raj, N., Veeresh Gobbur, S., Praveen., Rahul Patil., & Veerendra Naik. (2020). Implementing Hand Gesture Mouse Using OpenCV. In the International Research Journal of Engineering and Technology (pp. 4257-4261).

[8] Sneha, U., Monika, B., & Ashwini, M. (2013). Cursor Control System Using Hand Gesture Recognition. In the International Journal of Advanced Research in Computer and Communication Engineering (pp. 2278-1021).

[9] Krishnamoorthi, M., Gowtham, S., Sanjeevi, K., & Revanth Vishnu, R. (2022). Virtual mouse using YOLO. In the international conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (pp. 1-7).

[10] Varun, K.S., Puneeth, I., & Jacob, T.p. (2019). Virtual Mouse Implementation using OpenCV. In the International Conference on Trends in Electronics and Informatics (pp. 435-438).

[11] Quek, F., et.al. (1994). Towards a vision based hand gesture interface, in Proceedings of Virtual Reality Software and Technology (pp. 17-31).

[12] Tharsanee, R.M., Soundariya, R.s., Kumar, A.S., Karthiga, M., & Sountharrajan, S. (2021). Deep Convolutional neural network-based image classification for COVID-19 diagnosis. In Data Science for COVID-19 (pp. 117-145). Academic Press.

[13] Newell, A., Yang, K., & Deng, J. (2016, October). Stacked hourglass networks for human pose estimation. In the European conference on computer vision (pp. 483-499). Springer, Cham.

[14] Ramakrishna, V., Munoz, D., Hebert, M., Andrew Bagnell, J., & Sheikh, Y. (2014).Pose machines: Articulated pose estimation via inference machines. In the European Conference on Computer Vision (pp. 33-47). Springer, Cham.

[15] Tharani, G., Gopikasri, R., Hemapriya R., & Karthiga, M. (2022). Gym Posture Recognition and Feedback Generation Using Mediapipe and OpenCV. In International Journal of Advance Research and Innovative Ideas in Education (pp. 2053-2057).

[16] Shibly, K.H., Kumar, S., Islam, M.A., & Iftekhar Showrav, S. (2019). Design and Development of Hand Gesture Based Virtual Mouse. In the International Conference on Advances in Science, Engineering and Robotics Technology (pp. 1-5).