Hand Gesture Based Virtual Mouse Using Machine Learning

Neha Bhagwat[1] , Mansi Dusane[2], Moinoddin Inamdar[3]
Department of Computer Engineering[1,2,3],
Nutan Maharashtra Institute of Engineering and Technology, Maharashtra, India [1,2,3]

Abstract— This paper suggests a new method for controlling the cursor that uses hand movements that can be easily recorded by a regular webcam in place of the conventional mouse. This facilitates a hands-free user experience by doing away with the requirement for a physical device. The technology makes use of the webcam's capabilities to track hands accurately. It does more than just mimic hand gestures, though. Certain hand movements will be created for different functions, like dragging and left and right clicks. This allows users to use the cursor and its functionalities in a natural and simple way. For the system to be implemented as described, Python and OpenCV are required. These models, which mimic the swaying motion of human hand movements, have the potential to greatly improve the usability of such computer vision solutions. The world of technology nowadays is always changing due to the abundance of technologies. One such intriguing idea is the human-machine interface. The idea basic this is called signal acknowledgment, which is a technique for reproducing mouse functionalities on a screen without the requirement for any equipment.

Keywords—Gesture Control Virtual Mouse, Virtual Mouse, Hand Gestures, OpenCV

I. INTRODUCTION

The world is full of technology-based factors in our daily life. We have so many technologies, the computer technologies of the world are growing at the same time. They are used to perform various tasks that humans cannot do. In fact, they dominate human life because they have the potential to perform tasks that humans cannot. Human-computer interaction can be done with a printing device such as a mouse. A mouse is a device used to interact with a graphical user interface, which includes pointing, scrolling and scrolling, etc. Hardware mouse on PCs and touchpad on laptop needs a significant amount of time to perform complex tasks when the hardware mouse is involved. let's go, it would be damaged at some point. Through intricate algorithms and neural networks, computers have acquired the ability to discern not just the movements of hands but also the subtleties of facial expressions and body language, enriching the communication channel between users and machines. Moreover, the relentless march towards miniaturization in the realm of electronics is reshaping our conception of computing devices. From smartphones that fit snugly into our palms to wearable gadgets that seamlessly integrate into our attire, the physical manifestations of technology are becoming increasingly discreet. This trend towards invisibility is not merely cosmetic but reflective of a deeper shift towards a more seamless integration of technology into our daily lives. One notable innovation that exemplifies this trend is the virtual mouse system, a marvel of modern engineering that seamlessly integrates computer vision libraries like OpenCV with the versatility of Python programming. By harnessing the power of packages such as Pynput, Autopy, and PyAutoGUI, users can exert precise control over on-screen actions with mere flicks of their wrists. However, it is the incorporation of MediaPipe that truly elevates this system to new heights, enabling unparalleled accuracy in tracking hand movements and finger positions. This comprehensive framework empowers users to transcend the confines of traditional input devices, allowing them to interact with their computers in a manner that is both intuitive and immersive.
Whether navigating through digital landscapes or manipulating complex datasets, the boundaries between human and machine blur, giving rise to a symbiotic relationship where each augments the capabilities of the other.

II. EXISTING SYSTEM

In the current system, cursor control is facilitated through the use of a mouse, which may be either wired or wireless depending on user preference. An elective strategy that clients can investigate includes using hand signals to cooperate with the framework. This novel methodology is executed through a virtual mouse control framework that uses the feed from a webcam to recognize hued tips, like red, green, and blue, present on the fingers. These colored tips essentially serve as distinguishable objects that the webcam can easily track and interpret. By leveraging this technology, users can seamlessly execute standard mouse operations including dragging windows, minimizing applications, scrolling up and down, and performing left and right clicks entirely through hand gestures. To improve client experience further, it is arranged that the skin variety ID framework, inside which shaded fingers assume a key part, will be refined to kill the need of these hued finger markers.The objective of this enhancement is to advance the system's flexibility and adaptability so that skin color alone can be accurately identified to drive the cursor control functionality. In its current state, the system heavily relies on static hand recognition techniques, which involve processes like finger counting, fingertip identification, and hand shape analysis. While effective in optimizing the system's precision in interpreting user input, the complexity introduced by these techniques could potentially create challenges in terms of user-friendliness and ease of understanding. Thus, in future iterations, the focus will be on streamlining these processes to maintain efficiency without compromising user accessibility and usability[9].

III. LITERATURE SURVEY

Earlier research explored using gloves and coloured markers for hand gesture recognition in virtual mouse control. However, these methods faced limitations. Gloves can be uncomfortable, impractical to wear for long periods, and cause skin irritation. Coloured markers might not always be reliable for precise control. Current approaches, like those utilizing Google's MediaPipe framework, offer a more promising solution. This technology allows for gesture control of the virtual mouse, enabling various actions like clicks, dragging, and adjustments with improved accuracy.

The difficulties of different lighting and backgrounds in hand gesture detection for human-computer interfaces (HCI) are addressed in this study. The authors suggest a method that combines a motion history image-based technique for categorising dynamic hand gestures with an adaptive skin colour model (based on face identification) for hand region segmentation. Four groups of learned Haar-like features are used by the system to categorise hand movements, including the up, down, left, right, fist, and waving hand. Most household appliances can be controlled using this combination. Five people tried the framework by making 250 signals at various distances; the framework showed its reasonability with a typical exactness of 94.1% and a handling season of 3.81 milliseconds per outline [1].

This work proposes a PC vision and AI based constant hand signal location framework for human-PC communication. The authors enhance S. Shriram's method for contactless cursor control, most likely by hand gestures. The technology makes use of convolutional neural network models that MediaPipe developed within an OpenCV and Python framework. A solitary camera fills in as the information gadget, and the framework's result is displayed for client change[2].

This exposition investigates PC vision and hand motion acknowledgment with regards to Coronavirus and their likely applications to VR and HCI, especially as a method for reducing dependence on actual gadgets. The proposed system may be able to perform gesture-based control, such as dragging and clicking elements, by utilising computer vision algorithms. Albeit not referenced unequivocally, hand following and motion following suggest these strategies. It utilises a webcam as the input device and is written in Python and OpenCV. Among the libraries it depends on are NumPy, math, and mouse (perhaps to simulate mouse clicks). The camera's output is shown for user calibration. This study raises the prospect that gesture recognition will become more important in future HCI and VR interfaces[3].

To further develop openness for human-PC communication (HCI), this exploration proposes a hand signal controlled virtual mouse framework that utilizes PC vision and AI (ML) procedures. The technology, which is intended for persons with physical restrictions, records hand gestures using a camera. On the virtual screen, gestures that have been recognised are subsequently converted into matching mouse movements. Scalability and adaptability across devices and contexts are the system's main goals. It gives total virtual control without the requirement for additional equipment by consolidating voice directions with static and dynamic hand signals [4].

This study describes a Human-Computer Interaction (HCI) method for controlling mouse cursor movement and clicks using hand gestures. To capture hand motions, the system employs a colour-detection-based camera, with the purpose of delivering a low-cost virtual interface using a standard webcam. This technique is compatible with the current emphasis on creating more natural HCI interfaces that serve as an alternative to existing technologies such as mice and touchscreens. It exploits touchscreens' ease of use by involving a webcam as an info gadget to give a more normal communication experience [5].

This paper proposes a novel Human-Computer Interaction (HCI) method that utilizes computer vision and deep learning models to replicate human hand gestures and control cursor functionalities. The system leverages the OpenCV and MediaPipe libraries within a Python framework, potentially including the PyAutoGUI package for automation tasks. With this approach, finger movements are recognized and translated into cursor movements, killing the requirement for actual equipment [6].

It proposes a hand gesture-based virtual mouse technology powered by intelligent machines (AI). PC vision and variety discovery are utilized to distinguish explicit hand locales that are shaded or veil like to control mouse exercises. The framework utilizes Python and contemporary PC vision calculations to accomplish this capacity. It is intended to work impeccably without the requirement for extra equipment and permits far off PC control through hand motions for left-and right-clicking as well as cursor development [7].

Using the help of computer vision, this project creates a virtual keyboard and mouse that can be operated using hand gestures. The device uses a webcam to record hand movements, and computer vision algorithms analyse the pictures to identify particular gestures (perhaps employing convex hull defect calculations). Following recognition, appropriate mouse and keyboard commands are assigned to these motions. The framework is being created on the Boa constrictor stage with Python fully intent on offering a virtual connection point that is sans equipment and remote [8].

This study suggests a hand gesture-based contactless cursor control system. The only input device for the system is a webcam, and Python and OpenCV are used in its creation. The writers probably use computer vision algorithms which aren't stated specifically to identify different hand movements that correspond to cursor movement. Certain developments can likewise be utilized to control the clicking and hauling capabilities. NumPy, math and mouse (maybe to mimic mouse clicks) are among the Python prerequisites utilized by the framework, which likewise shows the webcam yield for client adjustment [9].

IV. ARCHITECTURE

Figure 1 displays a functional block diagram of the proposed system, illustrating how it operates. We need to reach up and touch the webcam. The webcam starts up and starts capturing frames for the video. The provided image underwent pre-processing. The primary function of the image pre-primary processor is to standardise the image. The technique of scaling and pre-processing an image to achieve uniform heights and widths is known as standardisation[2].

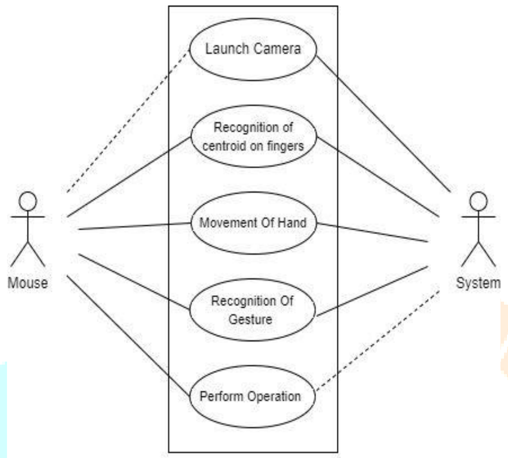

Fig-1. Use case Diagram

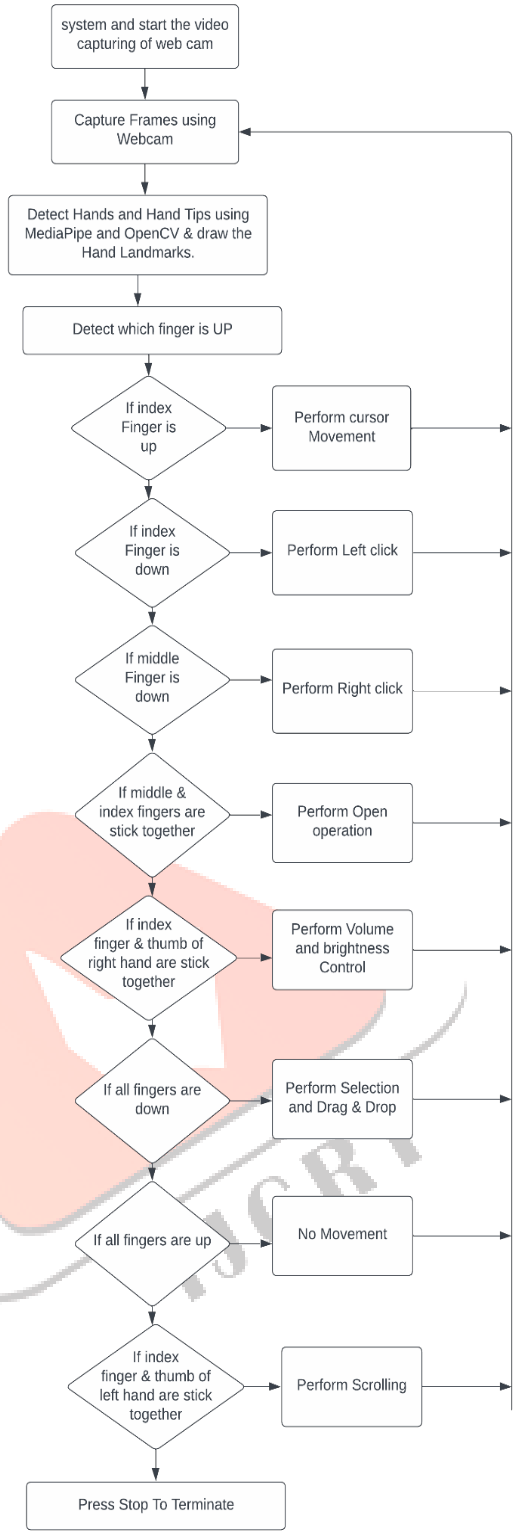

Fig-2. Flowchart

These days, an image processing technique is used to improve the quality of the average image. Following image processing, MediaPipe and OpenCV are used to identify the finger tips when the camera moves. After determining the hand and fingertip placements, it starts to outline. There are hand tourist spots and a crate enclosing the hand on the screen. On the Windows PC, draw a rectangle box big enough for the mouse. The fingers that are up and down on you will be recognized by it. In view of the finger discoveries, the mouse activity is performed. Then, the program gets back to the edges to play out the following activity. This is the overall system operation.

V. TOOLS USED

We used the open-source Mediapipe library to provide efficient hand and finger recognition. This structure, made by Google and based on cross-stage highlights, involves OpenCV for an assortment of PC vision errands. The Mediapipe program makes use of machine learning principles to identify and follow hand gestures.

A. Mediapipe

Google created the open-source MediaPipe framework to make it easier for developers to create real-time, cross-platform computer vision applications. Object detection, posture estimation, hand tracking, facial recognition, and a number of additional features are included in these technologies. Developers may quickly create complex pipelines that combine several algorithms and processes by using MediaPipe. Real-time execution of these pipelines is possible on a variety of hardware platforms, including as CPUs, GPUs, and specialised accelerators such as Google's Edge TPU. Moreover, the framework offers interfaces for smooth communication with other well-known machine learning libraries, such PyTorch and TensorFlow [6].

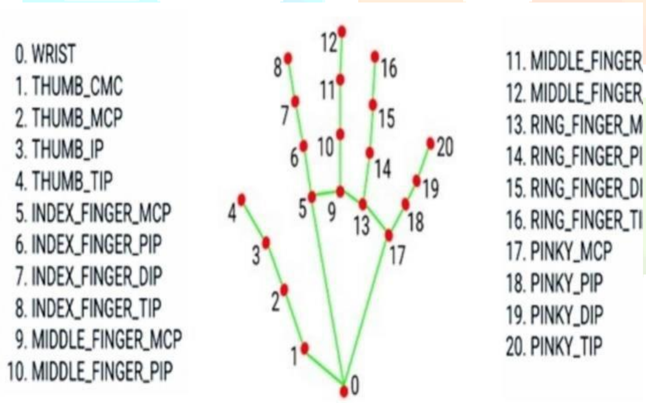

Fig -3: Hand Coordinates or Landmarks

B. OpenCV

An exhaustive programming library for PC vision and AI, OpenCV can be downloaded for nothing determined to assist developers with making PC vision applications. Filtering, feature identification, object detection, and tracking for both photos and movies are just a few of its many features. OpenCV can be used with other programming languages, such as Python, Java, and MATLAB, through its bindings. Numerous fields, such as robots, self-driving cars, augmented reality, medical image analysis, and more, use the C++ library. It is an effective tool for developers, offering a wide range of instruments and methods that simplify the creation of intricate computer vision applications across several fields.

VI. SCOPE OF THE PROJECT

The two most popular techniques for recognizing hand gestures are vision-based, which uses images processing methods that rely on camera input and hardware-based methods that need the user to wear a device. It is clear that the proposed system is vision-based, utilizing camera inputs and image processing techniques. Our goal is to provide more gestures so that in the future users may finish more jobs faster. This concept describes a system in which gestures are made exclusively with the appropriate hand. Consequently, advancements made to the current approach will make it possible to use both hands for specific movements in the future. This recommended method will save time and effort for the user, and it will make using computers easier for those who are blind or impaired.

VII. RESULT ANALYSIS

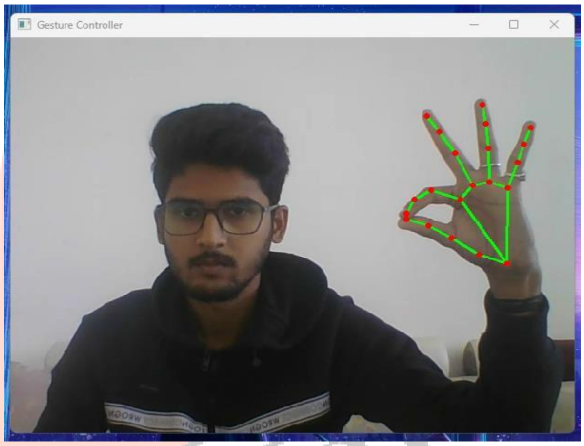

Fig- 3: Brightness and volume control

With help of this gesture, we can control brightness level and volume level by moving hand towards upward, downward, towards left and right.

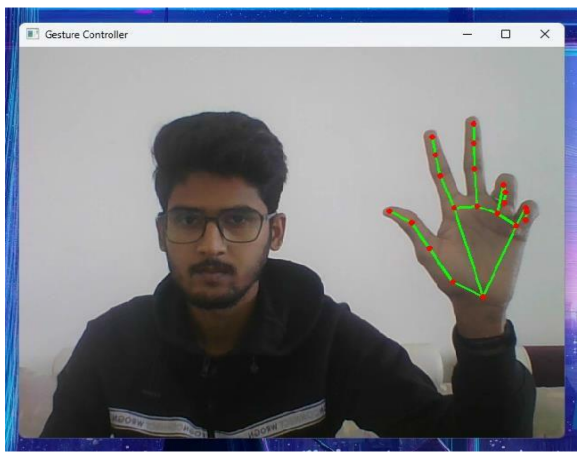

Fig-4: Cursor Movement

With help of this gesture, we can control cursor by moving hand in given window.

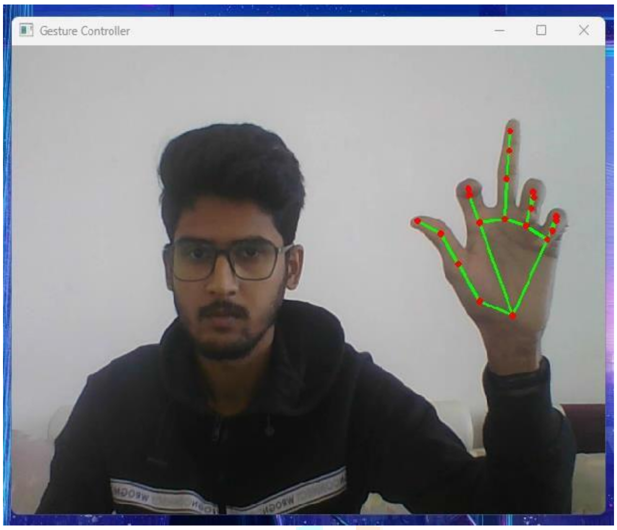

Fig – 6: Left Click

With help of this gesture, we can perform left click operation.

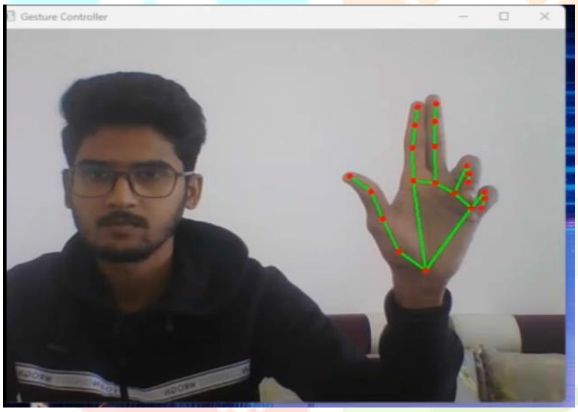

Fig-7: Double click

With help of this gesture, we can perform double click operation.

Fig-8: Right click

With help of this gesture, we can perform left click operation.

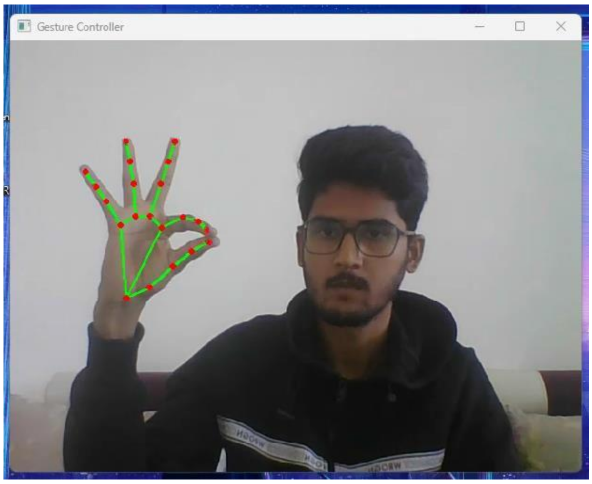

Fig-9: Scrolling

With help of this gesture, we can perform scrolling operation.

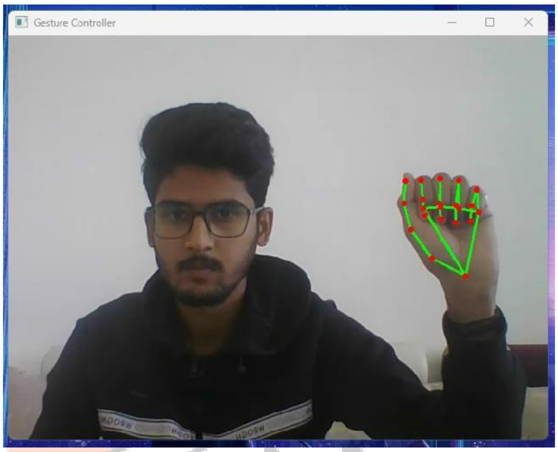

Fig-10: Selection, drag & drop

With help of this gesture we can perform selection, drag, drop operation.

VIII. FUTURE SCOPE

There is a tonne of promise for the future with hand motions and virtual mouse control. Future developments in gesture recognition should facilitate more intricate interactions by enabling users to effortlessly make multi-finger gestures, zoom, and scroll. This normal connection point will be additionally improved through voice order combination, giving a really natural method for communicating with PCs. Multimodal interaction will enable collaborative workspaces where users can easily integrate hand gestures with touch screens or keyboards, increasing efficiency, beyond individual demands. Adaptive interfaces that accommodate users with impairments will ensure that accessibility features are kept up to date and promote a more inclusive computer environment. Beyond conventional computers, 3D hand

tracking opens up even more possibilities by enabling natural interaction in virtual and augmented reality settings. Consider interacting with items in a virtual
the world or just a simple hand gesture to operate smart home appliances. You can make an exceptional and strong virtual mouse framework utilizing hand motions that impacts the bearing of human-PC communication by focusing on continuous execution to ensure liquid cursor development, UI plan that offers clear direction on upheld signals, areas of strength for and security measures to defend client information.

IX. APPLICATIONS

Virtual mouse control with hand gestures transcends the traditional desktop experience, offering a plethora of exciting applications. It empowers users with disabilities by enabling cursor control through head tracking or eye gaze, fostering greater independence. For creative professionals, imagine manipulating objects in 3D design software with intuitive hand movements or effortlessly zooming through images with a wave. Virtual and augmented reality experiences come alive with natural hand interaction, allowing users to navigate virtual worlds or interact with augmented objects seamlessly. The world of gaming can be revolutionized with gesture-controlled actions, adding a strategic and immersive layer to gameplay. Even smart homes and workplaces can benefit. Imagine controlling lights, adjusting thermostats, or manipulating media players with simple hand gestures. Presentations can be navigated effortlessly, and collaborative whiteboards can be manipulated with hand movements, fostering a more interactive workspace. By investigating these assorted applications, your virtual mouse project utilizing hand motions holds the possibility to change human-PC collaboration across different areas.

X. ADVANTAGES

Hand motion based virtual mouse control has many advantages over ordinary procedures. First of all, it encourages accessibility by giving users with limited hand mobility a another means of input. For individuals who experience difficulty with customary mouse, envision having the option to control the cursor by simply gazing at the ideal spot or by utilizing head following, which would foster more prominent freedom. Second, utilizing hand motions to communicate with PCs is a more instinctive and regular way to deal with use innovation. Imagine using a simple wave of your hand to manipulate items in 3D design software or a flick of your fingers to easily browse through photographs. For designers and artists, this simple contact can improve workflow and creativity. Thirdly, with augmented and virtual reality, hand motions can open up new possibilities. Consider connecting with augmented reality or exploring virtual worlds. normal hand signals to control things, bringing about an amazingly enamoring and vivid experience. All things considered, hand gesture-based virtual mouse control makes a computer more accessible, improves user experience, and creates opportunities for creative interaction across a range of industries.

XI. CONCLUSION

The primary aim of the AI virtual mouse system is to seamlessly replace the conventional physical mouse with a more intuitive interface based on hand gestures. By harnessing the capabilities of webcams or integrated cameras to recognize and interpret hand movements and gestures, this system eliminates the need for a physical mouse altogether. Through real-time processing of frames captured by the camera, the system translates hand gestures into corresponding mouse functions, offering users a novel and efficient way to interact with their computers. One of the key advantages of this AI virtual mouse system lies in its superior accuracy and performance compared to previous models. By leveraging advanced algorithms and machine learning techniques, the system can accurately track hand movements and gestures with remarkable precision. This elevated degree of exactness improves the general client experience as well as addresses normal difficulties related with customary mouse input strategies, like cursor nervousness or slack. Furthermore, the transition to an electronic interface controlled by hand gestures offers a myriad of practical benefits. Past the comfort of killing the requirement for actual equipment, the computer based intelligence virtual mouse opens up additional opportunities for clients with portability disabilities or ergonomic worries. By allowing users to control mouse functions through natural hand movements, the system promotes accessibility and inclusivity in computing environments.

XII. REFERENCES

[1]. Chen-Chiung Hsieh and Dung-Hua Liou, "A Real Time Hand Gesture Recognition System Using Motion History Image ", ICSPS, 2010.

[2]. G N Srinivas, S Sanjay Pratap, V S Subrahmanyam , K G Nagapriya, A Venkata Srinivasa Rao, "Virtual Mouse Control Using Hand Gesture Recognition", IRJET ,Feb 2023, e-ISSN: 2395-0056 p-ISSN: 2395-0072.

[3]. Bharath Kumar Reddy Sandra, Katakam Harsha Vardhan, Ch. Uday, V Sai Surya, Bala Raju, Dr. Vipin Kumar, "GESTURE-CONTROL-VIRTUAL-MOUSE", IRJETS, April-2022, e-ISSN: 2582-5208.

[4]. Kavitha R, Janasruthi S U , Lokitha S , Tharani G." HAND GESTURE CONTROLLED VIRTUAL MOUSE USING ARTIFICAL INTELLIGENCE".

[5]. Abhik Banerjee , Abhirup Ghosh , Koustuvmoni Bharadwaj , Hemanta Saikia, "Mouse Control using a Web Camera based on Colour Detection", Mar 2014, IJCTT.

[6]. Nagaraj M. Lutimath, Niharika K, K Bhuvanesh, Mohan M and Muhammad Zidan K M," The Design of Hand Gesture Controlled Virtual Mouse Using Convolutional Neural Network Technique", May 2023, IJIRT, ISSN: 2349-6002.

[7]. Piyush Pilare, Coral Mahato, Chanchal Khergade, Shubham Agrawal, Prasheel Thakre," IMPLEMENTATION OF HAND GESTURECONTROLLED MOUSE USING ARTIFICIAL INTELLIGENCE ", August - December 2022, ISSN: 2254-4143.

[8]. Ajay Sarwade, Gautami Sirsul, Shruti Salaskar, Vrushali Chennur, " GESTURE RECOGNITION BASED VIRTUAL MOUSE AND KEYBOARD", IJCRT, 2023, ISSN: 2320-2882.

[9]. Vijay Kumar Sharma, Vimal Kumar, Md. Iqbal, Sachin Tawara and Vishal Jayaswal "Virtual Mouse Control Using Hand Class Gesture" , · December 2020, ISSN NO : 1869-9391.

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

[11]. Anadi Mishra, Sultan Faiji , Pragati Verma, Shyam Dwivedi, Rita Pal, "Virtual Mouse Using Hand Gesture".

[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).