



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Special Issue 2, April 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# AI Based Approach to Control Cursor Using Hand Gesture

Boilla Vivek Vardhan Reddy<sup>1</sup>, Dumpala Yeshwanth<sup>2</sup>, Dr.E.SriVidhya<sup>3</sup>

U.G. Student, Department of Computer Science Engineering, Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai, Tamil Nadu, India<sup>1</sup>

U.G. Student, Department of Computer Science Engineering, Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai, Tamil Nadu, India<sup>2</sup>

Professor, Department of Computer Science Engineering, Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai, Tamil Nadu, India<sup>3</sup>

**ABSTRACT:** The mouse is one of the most practical gadgets to have resulted from advancements in Human-Computer Interaction (HCI) technology. Because it requires a battery for electricity and a dongle to connect to the computer, today's wireless or Bluetooth mice are still reliant on technology. The artificial intelligence virtual mouse technology can get over this restriction that has been developed by making use of a webcam or a camera that is built into the device itself for the purpose of capturing hand motions and hand tip identification using computer vision. The system takes use of the algorithm for machine learning as its primary means of processing data. The computer can be controlled digitally by employing hand movements to perform actions including scrolling, cursor, left- and right-clicking, and other actions. This eliminates the need for a real mouse to operate the computer. When it comes to identifying the hands, the system relies on deep learning. Furthermore, the proposed method will stop COVID-19 from spreading since it will do away with the necessity for human interaction and independence from computer-controlling gadgets.

**KEYWORDS:** Human-Computer Interaction, Hand Gesture, Mediapipe, Hand Tracking

## I. INTRODUCTION

The proposed method will stop COVID-19 from spreading since it will do away with the necessity for human interaction and independence from computer-controlling gadgets. The purpose of this research is to present Because it eliminates the requirement for direct human contact and independence from computer-controlling devices, The suggested approach will halt COVID-19's spread. Acknowledgement in order to execute articulations mouse limitations at ranges that are favorable to PC cheating adaptable to PC vision [1]. The smaller PC is stretched when a web camera or a standard camera is used to operating system pointer functions. works and material performs rather than making use of an antiquated mouse device is likely to result in the most objective application of the proposed system. The computer uses hand signals and the hand tip area as an HCI by abusing the useful PC vision. Using an intrinsic camera or a network camera, we will respond to the tip of the hand signal. Next, we'll practice the mouse pointer tasks and exploration work while moving the pointer at the same time. Our last phase will include the use of an AI virtual mouse. Nonetheless, the client controls the PC mouse operations throughout this article by using hand gestures and an innate camera or visual camera. This contrasts with utilizing a remote control or a Bluetooth mouse, which are a couple of gadgets made expressly for controlling a mouse, interacting with a computer, and also using a battery. A data interstate camera is embedded in the projected structure and blocks the cycle of edges that have been obtained in a way that makes it possible to Take note of the numerous hand motions and hand-tip developments that have an impact on a particular mouse task [2].

To enable the AI virtual mouse structure, Python is used as a programming language. Furthermore, the AI virtual mouse architecture makes variable degrees of use of Open CV, which is unavoidably the library for adaptive Computer vision. NumPy, AutoPy, and PyAuto GUI packs were employed for expanding the Computer screen to carry out verbalizations such as left click, right click, and evaluating boundaries [3].. The hands' trip and the chase of the hands'

tips are handled by the model using the Python Media-pipe bunch. There, the projected model shows, in an ideal scenario, a high level of accuracy, and in this way, the predicted Whether employing a cycle or alternatively a PC GPU, the model can perform effectively in simple applications.

### A. Problem Statement

The proposed artificial intelligence virtual mouse, which uses hand signals as its input method, model can perform well in simple applications when employing a cycle or alternatively a PC GPU., as well as for situations where it isn't possible to use a real mouse due to space constraints. Furthermore, because of the COVID issue, it is not safe to include devices by reaching them because the planned AI virtual mouse might also be altered. This is because the project aims to stop the disease from spreading by reaching devices. This is because the COVID scenario seeks to achieve what is already occurring, namely the spread of the disease through contact with equipment. overcame various challenges You can use a camera to manage the device mouse restrictions since hand sign and hand Hint disclosure is done making use of a camera or a unique camera, such as a webcam [4].In order to run a computer, you'll need a few pieces of equipment, most notably a mouse, utilising a camera or specialised camera, such as a webcam [4].As a result, the consumer uses either their natural camera or their visual camera throughout the entire procedure, controlling the mouse activity on their computer with hand movements.

### B. The purpose

To finish as many jobs on the board as is humanly possible is the main the synchronised AI virtual Mouse steamiest objective impulse alternative for the traditional and antique mouse construction. This can be accomplished with the aid of an internal network camera that records hand movements and the tip of the hand before processing the data to carry out particular mouse actions like left click, right click, and investigative actions. By doing away with the requirement for user input and the reliance on outside devices to operate the computer, the suggested structure will stop COVID-19 from spreading. This vision will be utilised to its maximum potential throughout COVID in order to maintain social segregation and close off communication.

Fraud using the Virtual Mouse Clients are granted permission to using hand gestures to manipulate the mouse if they have the hand movement affirmation [5]. The automated web camera of the system is utilised to keep tabs on the user's hand motions. The motion confirmation process made use of pc vision systems. The Open-CV Python Group contains a video get function, which can be used to extract data from a live video. This is an essential issue for us, as we want to understand the applications that the model will create in order to track the progress of the mouse without physically manipulating it.

### C. Work Connected to It

Below are a few related works that have been approved on a virtual mouse that still use shading tips among the hands and wear gloves among the hands to locate hand motions, but the mouse's accuracy has improved.The wearing of gloves has no bearing on the acknowledgment in any way; nevertheless, some customers' hands don't seem to fit the gloves; and occasionally, the acknowledgment is incorrect due to a mistaken placement of tips that are meant to be concealed. During the entirety of this strategy, the customer was given the opportunity to wear a data glove. This strategy was initially introduced by Quam in 1991.

The extended structure that was developed by Quamtho' yields more precise results; using the framework, it is possible to successfully carry out a wide variety of signal control operations. Poop "the most important limitation imposed by this concept is the need for additional unassuming hand signals [6]. A focus on "Cursor system abuse Hand Gesture Recognition" was extended by Neha, Parul Gandhi, and AshwiniM.2014.

## II. METHODOLOGY

The primary objective of the artificial intelligence virtual mouse system that has been presented is to provide a successor to the dated mouse system that is currently in use in order to be able to perform and exert control over

mouse-related operations. the usage of an internal network camera to capture hand motions and fingertip movements, then manipulate those frames to perform the necessary mouse activities, such as the left click, the right click, and the scroll function, is one possible method for accomplishing this goal.

The proposed method will, as a result, put an end to the propagation of COVID-19 by doing away with the need for human intervention and removing the reliance on external devices that are necessary for computer operation. This vision will be extremely useful during the course of COVID for the purpose of maintaining social distance and untouchability [7].

The suggested AI virtual mouse system's major objective is to provide a replacement for the existing, antiquated mouse system in order to carry out and control the mouse functions. This might be accomplished by using an internal network camera that captures hand gestures and hand tips and then processes those frames to perform the required mouse movements, such as the left click, right click, and scroll function. In doing so, the suggested approach will halt the spread of COVID-19 by doing away with the need for human interaction and dependency on outside gadgets to run the computer. This vision will be extremely useful during the course of COVID for the purpose of maintaining social distance and untouchability [7].

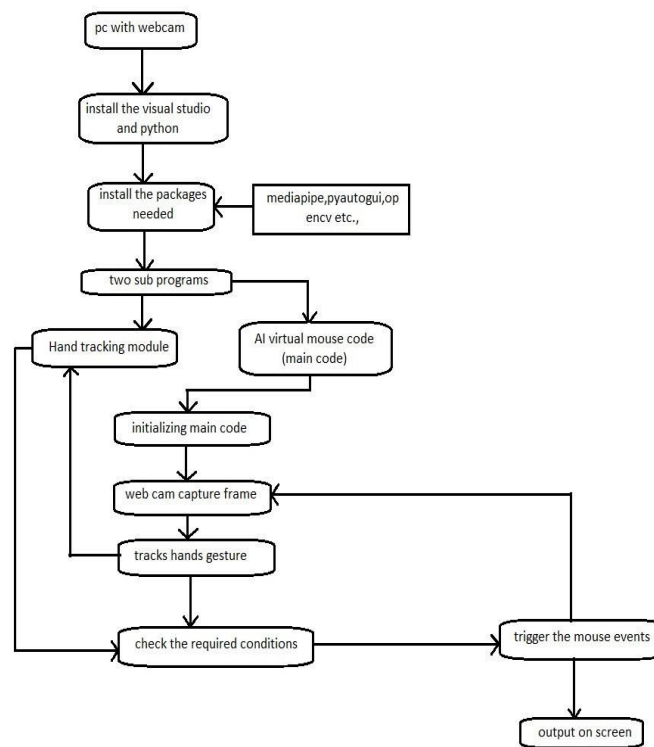


Fig 2.1 system architecture diagram

Exploiting virtual mice Users can operate a mouse with the aid of hand gestures thanks to hand gesture recognition. Hand movements are followed using the system's digital web camera. Gesture recognition was accomplished using computer vision techniques. The major goal is to identify the models' applications so that they can build mouse movements without touching or using a mouse [8]. The OpenCV Python library has a video capture feature that is used to extract information from a live video.

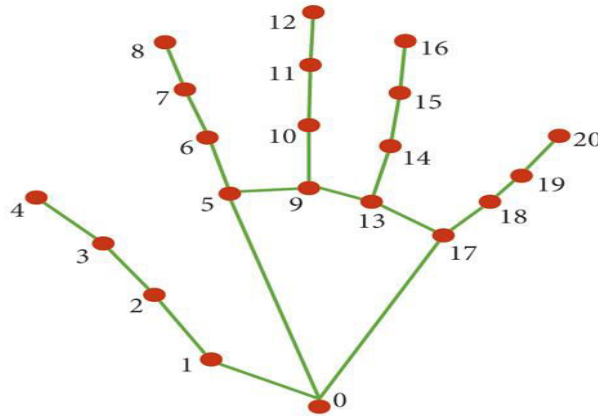


Fig 2.2 pointing the skeleton of the hand for hand tracking

The Media Pipe framework is employed in order to accomplish the goal of recognizing different types of hand motions. Applications involving computer machine vision make use of the OpenCV library. [9] To track and detect hand gestures and hand tips, the rule makes use of machine learning methods.

**D. SELECTED METHODOLOGY OR PROCESS MODEL:**

This system was built in such a way that the entire display can be represented by a small screen box; this box, in turn, depicts the entire screen.

Coordinates are used in the calculation process to produce translations.

- Camera Settings The digital camera that is attached to the desktop or laptop computer serves as the operations unit manager during runtime. Making a Video Capture object is required in order to successfully record a video. It is possible for its argument to be the index of a device or the name of a video file [10]. Each camera should be utilized according to the device index, which is only a number. As we almost never use more than one camera at a time, we almost always record it as a "0" in our logs. We are able to integrate any number of cameras into the system and report the results as one, two, and so on. After that, you will take still images of it frame by frame. However, in the very end, you must ensure that the capture is released. We are also able to apply color detection techniques to any image by making a few basic tweaks to the code [11].

The act of capturing frames The use of the infinite loop ensures that the camera will take pictures at each instance and will remain open for the entirety of the programme so that it can record the frames. We take a frame-by-frame recording of the live broadcast stream. The next step is for us to convert from RGB (the default) colour space to HSV colour space for each and every frame that was just taken. OpenCV provides access to more than 150 different methods for converting between colour spaces [12]. However, we are going to focus on the BGR to grey and BGR to HSV conversions because these are the ones that are utilized the most.

- Masking Technique The mask is produced by employing a specific image and applying restrictions that vary from region to region. The resultant mask is an object that is coloured red. Following this, we perform a bitwise and operation calculation between the Input image and the Threshold image. The end result focuses mostly on highlighting objects of the colour red [13]. The output of the AND operation is saved in the variable res, and then the frames and mask are displayed by utilizing the imshow() method with three different windows.

Showcase the Picture in Its Frame Because it is a high-level GUI procedure, the inShow() method has to call waitkey somewhat frequently. The processing of the imshow() function will be finished once the waitkey function is called. The waitkey() method awaits a key event for the "delay" parameter (here, 5 milliseconds). Because High GUI performs events like as redraw, resizing, and so on, we still call the waitkey method even when there is a 1ms delay [14].

● Mouse Movement We have to begin by computing the Centre of each identified red object that we can easily manage by taking the average of the highest and lowest points in the enclosing box. presently, we have obtained a pair of coordinates from the Centre of the two objects, and we can, we'll, or we will receive the red purpose depicted in the image [15]. Whenever possible, we like to square measure by shifting the detected point from camera resolution to the specific screen display. As a result, we frequently decide where to put things based on where the mouse is. On the other hand, navigating the mouse pointer will require some time. So, we are required to wait till the mouse pointer reaches that moment. So, we've begun a loop, and while we're technically "doing something" in that space, all we're really doing is just waiting to see if the current mouse location is the same as the one that has been assigned. That's meant to be the open hand gesture [16].

**a. Bringing the Essential Library in Install and import all necessary libraries into your preferred IDE at the outset:**

The onency, numpy, mediapipe, and autopy framework packages are the ones that need to be set up. After successfully installing all of the libraries in pycharm, The next thing we need to do is import each of those libraries into the code section, which will allow us to use this library in our system [17].

**b. Beginning the process of initializing the capturing device:**

The next thing we need to do is initialize the picture capture device that is connected. We are employing the principal camera of the system. if you have any other picture capture devices attached to your computer that this programme does not make use of.

**c. Taking the pictures and doing the processing:**

The artificial intelligence virtual mouse system proposed here makes use of the primary camera in order to capture the frames. In order to locate the hand within the video frame, the video frames first need to be processed, after whereby the frames' BGR to RGB format will be altered.

**d. Initializing the mediapipe and the capture window:**

Following this, we initialize the mediapipe, which is the component that is in charge of tracking the hand gesture as well as other activities. The next step is to determine the minimum level of confidence required for hand detection as well as hand tracking. This indicates that it has an 80% chance of succeeding based on the action that was carried out by the main hand. After that, make an axis line with mediapipe that runs across the palm and the centre of the fingers. It establishes a connection with the next-door finger and monitors the finger's motion [18]. Next, we define the values for the x- and y-axes to specify the hand's axis in the 2D format. We are able to maintain track of the finger with the assistance of this axis, which allows us to determine which one is up and which one is down. Moreover, we store the preceding frame's x- and y-axis values as 0, 1, and so forth.

**e. Defining hand landmarks:**

Moreover, we store the preceding frame's x- and y-axis values as 0, 1, and so forth. The designated locating location of the mediapipe 21 is capturing the hand land mark. (Figure 2). The mediapipe is going to follow these 21 joints or locating points as they move along. If there is not a landmark detected, the default value will be printed, which is nothing. It indicates that there is no need to carry out any actions. In all other circumstances, it will give back the coordinates' X, Y, and Z values. The index finger and thumb are not regarded as being a section of the hand's landmark area. The graph line connecting the 21 points is being used to build a loop that returns the value of the tip point. The lines travel downward when we fold the finger or move it downward, and the graph's lines do the same. It does not matter how much space there is between the fingers at this point. It is possible for there to be a gap of 5 centimetres between the fingers. In that case, it will reset to the place it was in when it was first found. On the basis of each unique image, we translated the decimal coordinate relative to the index after establishing the height, width, and centre of the hand landmark. A green rectangle encircling the hand appeared once the landmarks of the entire hand were successfully captured. We can use the mouse inside this box to navigate.

The graph line connecting the 21 points is being used to build a loop that returns the value of the tip point. Examples of these situations include being at a maximum distance from the camera, dim lighting, tracking against a dark or crowded

background, and more. One person's hand can be recognized at a time by the system that has been developed. The system employed in this study, in contrast to other systems, only has a single CPU and does not require any additional GPU systems or other specialized hardware.

### III. RESULT AND DISCUSSION

#### A. ALGORITHMS FOR HAND GESTURE TRACKING:

**MediaPipe:** Google's open-source MediaPipe framework is a cross-platform tool for creating multimodal ML pipelines. An innovative machine learning model called MediaPipe Hands was developed by Google to detect and track human hands in real-time from a single image or video source. The model identifies 21 unique landmarks or important spots on the hand, including the fingertips, knuckles, and wrist, that can be utilised to decipher different hand stances and motions. A pipeline of ML models is used by MediaPipe Hands.

Both a Hand Landmark Model and a Palm Detection Model utilise the complete image and provide an orientated hand bounding box that is used for the hand detection procedure. Scalability on desktops and mobile devices is possible thanks to the built pipeline's capacity to execute on several platforms.

#### B. OUTPUT SCREENSHOTS:

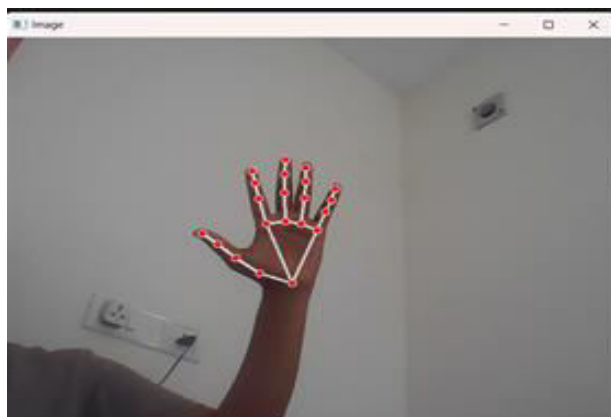


Fig 3.1. obtaining 21 land marks on hand

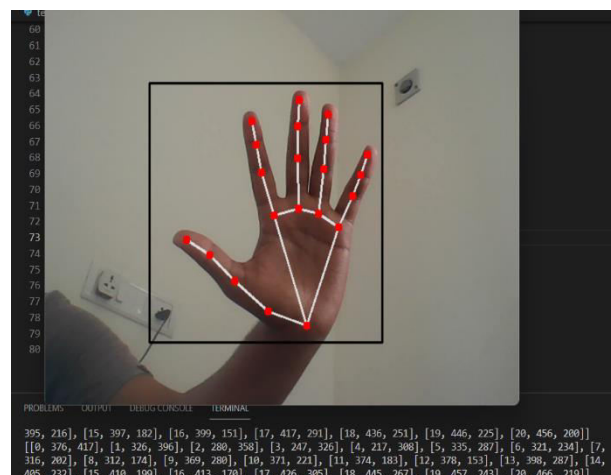


Fig 3.2. Drawing Rectangular Box Around Palm And Obtaining Coordinates





When both index and thumb finger is 1 and remaining 0 based on index finger tip.

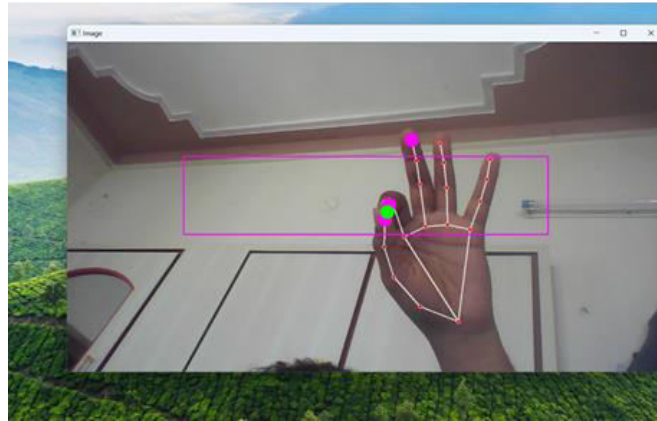


Fig 3.6 left clicks

When distance between index and thumb finger enough closer based on condition and remaining fingers up or 1.

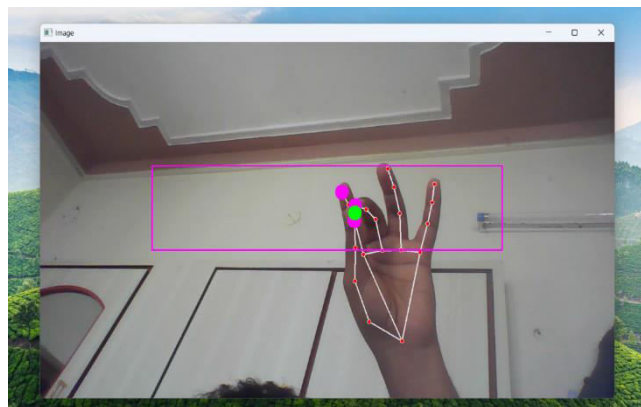


Fig 3.6 right clicks

When distance between center and thumb finger enough closer based on condition and remaining fingers up or 1.

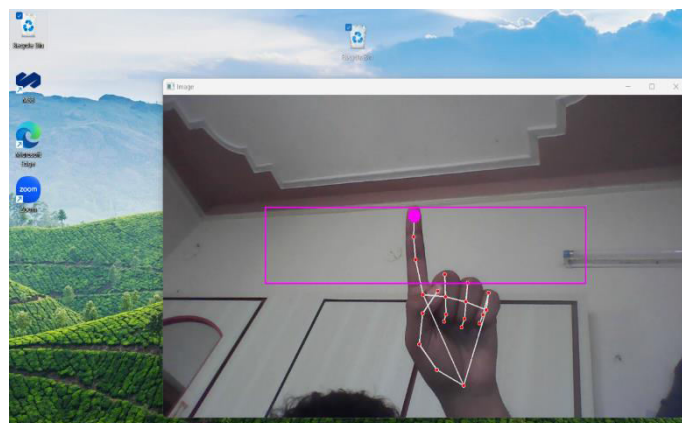


Fig 3.7 drag

When index finger is up and remaining are down.

#### IV. CONCLUSION

An accuracy rate of 80% to 95% has been achieved by the proposed technique. The system is able to read an image and find a hand in it. The hand tracking module performs the same standard mouse actions, such as moving the left and right mouse buttons and dragging an object, well. Nevertheless, there are some limitations that require high resolution cameras for greater tracking and precision.

#### REFERENCES

- [1] 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.
- [2] D. L. Quam, "Gesture recognition with a DataGlove," *IEEE Conference on Aerospace and Electronics*, vol. 2, pp. 755-760, 1990.
- [3] D.-H. Liou, D. Lee, and C.-C. Hsieh, "A real time hand gesture recognition system using motion history image," in *Proceedings of the 2010 2nd International Conference on Signal Processing Systems, IEEE, Dalian, China, July 2010*.
- [4] S. U. Dudhane, "Cursor control system using hand gesture recognition," *IJARCCCE*, vol. 2, no. 5, 2013.
- [5] K. P. Vinay, "Cursor control using hand gestures," *International Journal of Critical Accounting*, vol. 0975-8887, 2016.
- [6] L. Thomas, "Virtual mouse using hand gesture," *International Research Journal of Engineering and Technology (IRJET)*, vol. 5, no. 4, 2018.
- [7] P. Nandhini, J. Jaya, and J. George, "Computer vision system for food quality evaluation—a review," in *Proceedings of the 2013 International Conference on Current Trends in Engineering and Technology (ICCTET)*, pp. 85-87, Coimbatore, India, July 2013.
- [8] J. Jaya and K. Thanushkodi, "Implementation of certain system for medical image diagnosis," *European Journal of Scientific Research*, vol. 53, no. 4, pp. 561-567, 2011.
- [9] P. Nandhini and J. Jaya, "Image segmentation for food quality evaluation using computer vision system," *International Journal of Engineering Research and Applications*, vol. 4, no. 2, pp. 1-3, 2014.
- [10] J. Jaya and K. Thanushkodi, "Implementation of classification system for medical images," *European Journal of Scientific Research*, vol. 53, no. 4, pp. 561-569, 2011.
- [11] J.T.CamilloLugaresi, "MediaPipe: A Framework for Building Perception Pipelines," 2019, <https://arxiv.org/abs/1906.08172>.
- [12] Google,MP,<https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html>.
- [13] V.Bazarevsky and G. R. Fan Zhang, On-Device, MediaPipe for Real-Time Hand Tracking.
- [14] K. Pulli, A. Baksheev, K. Korniyakov, and V. Eruhimov, "Realtime computer vision with openCV," *Queue*, vol. 10, no. 4, pp. 40-56, 2012.
- [15] <https://www.tutorialspoint.com>opencv>.
- [16] D.-S. Tran, N.-H. Ho, H.-J. Yang, S.-H. Kim, and G. S. Lee, "Real-time virtual mouse system using RGB-D images and fingertip detection," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 10473-10490, 2021.
- [17] 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.
- [18] K. H. Shibly, S. Kumar Dey, M. A. Islam, and S. Iftexhar 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.



**SJIF: 8.423**



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INTERNATIONAL CENTRE



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details