# ”GESTURE RECOGNITION BASED VIRTUAL MOUSE AND KEYBORD”

Prof. Jayshri Mankar[1], Mr. Shantanu Zade[2], Mr. Vishwajit Kamble[3], Mr. Tushar Bammani[4] , Ms. Thanuja Gaddam[5]

*Mr. Shantannu Zade[1] Students, Dept of Computer Engineering, GSMCOE , Maharashtra, India*
*Mr. Vishwajit Kamble[2] Students, Dept of Computer Engineering, GSMCOE , Maharashtra, India*
*Mr. Tushar Bammani[3] Students, Dept of Computer Engineering, GSMCOE , Maharashtra, India*
*Ms. Thanuja Gaddam[4] Students, Dept of Computer Engineering, GSMCOE , Maharashtra, India*

## ABSTRACT

*Researchers around the globe nowadays are concentrating on making our devices more interactive and trying to make the devices more operational with minimal physical contact. Thai research proposes an interactive computer system that uses computer vision to create a virtual keyboard and mouse using only hand gestures. We can utilize a built-in camera or an external camera to track the image of different gestures performed by the person's hand and according to different gestures conduct mouse cursor operations such as right and left clicks, as well as move the cursor. In addition to that, we can control the console using different gestures such as two finger to select the particular alphabet and a eye motion gesture to perform swipe operation in either the left or right direction. This proposed system with no wire or any other external device will act as virtual keyboard and mouse. The only hardware aspect of this system is the webcam which is used to capture images whereas the coding part will be done in python programming language*.

**Keywords:** *Hand Motion, Webcam, Vision, Finger Recognition, And Gesture Based*.

## INTRODUCTION

Gesture recognition and shadowing is a kind of image processing process. In recent times, a number of gesture recognition ways have been proposed. Hand tracking has several operations including motion capture, human-computer interaction and human behaviour analysis. Several types of detectors and detection gloves are used for hand motion detection and tracking. Instead of using more precious detectors simple webcams identify the gesture and track the motion.. The proposed system which is webcam grounded, might be suitable to exclude the need of a mouse and keyboard. The process of interaction with a computer using hand gesture is a veritably intriguing & effective approach to HCI (Human- Computer- Interaction). The Virtual Mouse works as a medium of the user and the machine only using a camera. It helps the user to interact with a machine without any mechanical or physical bias and control mouse functions. Generally, we use a mouse, keyboard or other interacting bias which is substantially compact with the computer machine. The wireless bias also need a power source and connecting technologies, but in this paper, the user's bare hand is the only input option using a webcam. So, it's a veritably interactive way to control the mouse cursor and keyboard. This system has the implicit to replace the typical mouse and also the remote regulator of machines. The only hedge is the lighting condition. That's why the system still can't be enough to replace the traditional mouse as utmost of the computers are used in poor lighting conditions. In particular, people with severe movement disabilities may have physical impairments which significantly limit their capability to control the fine motor. thus, they may not be suitable to class and communicate with a normal keyboard and mouse.

## LITERATURE SURVEY

Paper 1 Because of the rapid advancement of computer vision, there is an increasing demand for humanmachine interaction. Hand gesture recognition is frequently employed in robot control, intelligent furniture, and other aspects because hand gestures can represent enhanced information. Paper 2 Patients with Amyotrophic lateral sclerosis (ALS) or stroke are unable to communicate their basic wants and demands. Because they can still utilise their eyes and move their heads, they can communicate via eye trackers. This research provides new approaches for improving eye tracker software's speed and ease of use.

Paper 3 Text production is one of the most common computer tasks, a seemingly insignificant task that can be limiting for people with severe neuromotor illnesses like Amyotrophic Lateral Sclerosis, which can result in Locked-in syndrome. Because they may only be able to communicate and engage with the outside world through eye movements, these people require augmentative and alternative communication technologies. This project looks into eye movement tracking as a way of interaction and displays a virtual keyboard that uses gaze detection as a text input. Paper 4 One of the research in human-computer interaction is a virtual mouse with fingertip identification and hand motion tracking based on image in a live video. The use of fingertip identification and hand motion recognition to control a virtual mouse is proposed in this research. Two approaches for tracking the fingers are used in this investigation. One method is to use coloured caps, while the other is to detect hand gestures. Finger recognition utilising colour identification, hand gesture tracking, and implementation on the on-screen cursor are the three primary processes.

Paper 4 Dynamic and Personalized Keyboard for Eye Tracker Typing. Patients who suffer from Amyotrophic lateral sclerosis (ALS) or stroke cannot talk and express their everyday basic needs and requests. They can communicate using eye trackers since they can still use their eyes and sometimes move their heads. This study suggests new methods for improvements in both speed and ease of use for eye tracker softwares. The first one is letter prediction to improve the speed, and second one is a new design that obviates the need of blinking with eye trackers, thus providing more and longer sessions of writing.

Paper 5 In human-computer interaction, virtual mouse implemented with finger tip recognition and hand gesture tracking based on image in a live video is one of the studies. In this paper, virtual mouse control using finger tip identification and hand gesture recognition is proposed. This study consists of two methods for tracking the fingers, one is by using colored caps and other is by hand gesture detection. This includes three main steps that are finger detection using color identification, hand gesture tracking and implementation on on-screen cursor. In this study, hand gesture tracking is generated through the detection of the contour and formation of a convex hull around it. Features of hands are extracted with the area ratio of contour and hull formed. Detailed tests are performed to check this algorithm in real world scenarios.

## PROPOSED METHOD

Object detection may be a engineering that deals with finding occurrences of objects in photos and videos and is connected to computer vision, image processing, and deep learning. during this system, something called “Haar Cascades” is employed to detect objects. it is a machine-learning approach during which a cascade function is learned employing a sizable amount of positive and negative photos. After then, it's utilized to search out items in other photos. Haar cascades have variety of benefits, one amongst which is their speed. to coach the classifier, the tactic requires an outsized number of positive images (images of faces) and negative images (images without faces). After that, features are extracted from it. They're really like our convolutional kernel. Each feature may be a single value produced by subtracting the overall of pixels within the white and black rectangles. instead of computing at each pixel, it divides the screen into subrectangles and creates array references for every of them. The Haar features are then computed using them. it is vital to notice that while doing object detection, practically all of the Haar characteristics are meaningless because the sole features that matter are those of the article. However, Adaboost helps to settle on the best characteristics from many thousands of Haar features to represent an object. The system uses haar cascades with OpenCV. A repository of pre-trained Haar cascades is maintained by the OpenCV library. the bulk of those Haar cascades are utilised for one among two purposes: Face detection, Eye detection, Mouth detection, Full/partial body detection. To implement mouse events, mediapipe library is employed. MediaPipe may be a framework for creating machine learning pipelines for time-series data like video and audio. For keyboard implementation, CVzone library in python is employed. CVzone could be a computer vision application that produces image processing and AI operations simple to use. it's built round the OpenCV and Mediapipe libraries.

A. **Haar Cascade Algorithm**:
Haar Cascade is a machine learning object detection algorithm used to identify objects in an image or video. The algorithm has mainly four stages: 1. Haar Feature Selection 2. Creating Integral Images 3. Adaboost Training 4. Cascading Classifiers

B. **Data obtaining:**
The initial move is to capture the image from camera and to define a region of Interest in the frame, it is important as the image can contain a lot of variables and these variables can result in unwanted results and the data that needs to be processed is reduced to a large extent. To capture the image a web-camera is used that continuously captures frames and is used to get the raw data for processing. The input picture we have here is uint8. The Procured image is RGB and must be processed before the components are separated and acknowledgement is made.

C. **Data Pre-Processing:**
Pre-processing method can be completed 2-steps process: (a) Segmentation (b) Morphological filtering First Process is the Segmentation process. It is done to change over grey-scale picture into the binary picture. That is, one will be hand and another one is background. Algorithm can be Haar Cascade used for this process and Gray scale picture are converted into binary picture having area of interest as the hand and the background.

D. **Background Subtraction**:
Background subtraction (BS) is a common and widely used technique for generating a foreground mask by using static cameras. As the name suggests, BS calculates the foreground mask performing a subtraction between the current frame and a background model, containing the static part of the scene or, more in general, everything that can be considered as background given the characteristics of the observed scene. Background modeling consists of two steps Background Initialization and Background Update. In the first step, an initial model of the background is computed, while in the second step that model is updated in order to adapt to possible changes in the scene.

**E. Morphological Filtering:**
After thresholding we have to make sure that there will be no noise is present in image, so we are using morphological filtering Techniques, These Techniques are divided into Dilation, Erosion, Opening and Closing If the division is not continuous, then there may have some _1s‘ in the background which is called as background noise, Also, there is a possibility that system generated an error in recognizing gesture this may be termed as gesture noise. If we want flawless contour detection of a gesture, then abovementioned errors should be nullified. A morphological separating (filtering) approach is employed utilizing grouping of dilation (enlargement) and erosion (disintegration) to accomplish a smooth, shut, and finish the contours of a hand motion.

F. **Extraction of Features**:
Pre-prepared or pre-processed picture is accessible to be utilized and different highlights of the resultant picture are removed. Features that can be extracted are Finding Contours, Finding and correcting convex hull and Action.

## SYSTEM ARCHITECTURE

A System Architecture for gesture recognition based on virtual mouse and keyboard using the Haar cascade algorithm typically consists of several components that work together to enable gesture recognition and interactions

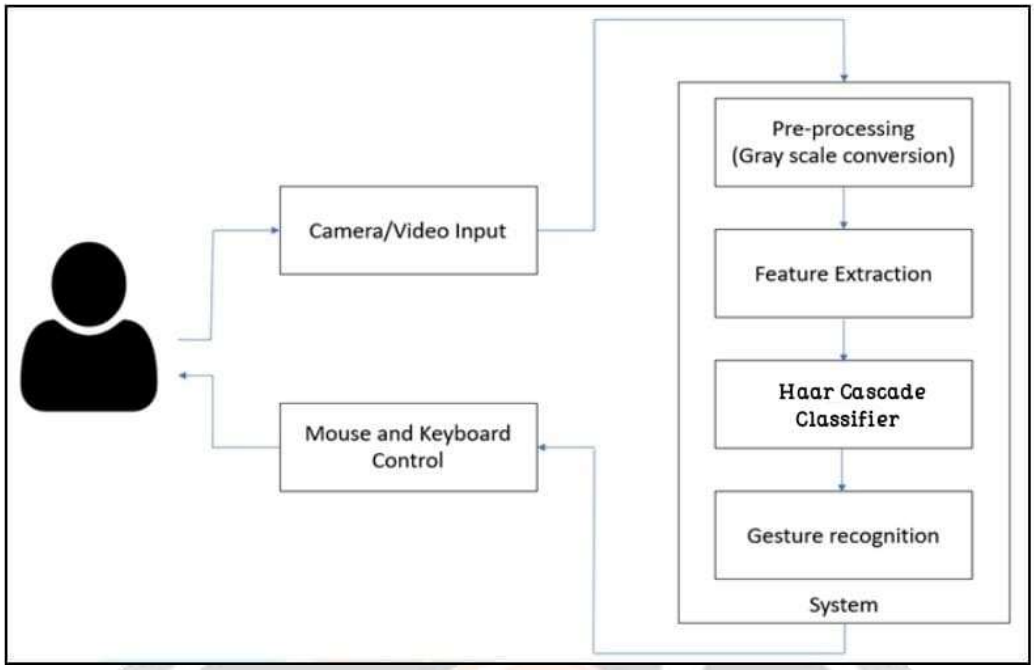


Fig: System Architecture

1. **Input Device:**
   The system utilizes an input device, such as a camera or eye tracker, to capture the user's hand or eye movements.
2. **Image/Video Capture**:
   The input device capture images or video frames of the user's hand or eye movements, providing the necessary data or gesture recognition.
3. **Preprocessing:**
   The capture images or video frames undergo preprocessing to enhance the quality of the input data. This maybe involve converting the images to grayscale, equalizing the histogram or applying other images enhancement techniques,
4. **Haar Cascade Classifier:**
   The preprocessed images or frames are fed into the Haar cascade classifier which is trained to detect the region of interest corresponding to the gesture being performed. The Haar cascade classifier consists of multiple stages, each containing a set of weak classifier that collectively provide accurate detection of the gesture.
5. **Gesture Tracking:**
   Once the Haar cascade classifier detects the gesture region the system tracks the movement of the detected region in subsequent frames. This tracking enables real-time gesture recognition and continuous interaction.
6. **Gesture Recognition**:
   Gesture recognition algorithm analyze the tracked hand or eye movements to recognize specific gestures. These algorithms may employ various techniques such as shape analysis, motion analysis, or machine learning-based approaches, to interpret the gesture and map them to predefined actions or commands.
7. **Action Mapping**:
   The recognition gestures are mapped to corresponding actions or commands. For example, eye gesture mapped to gaze-based cursor control, blinking for clicking, or other contextual actions.
8. **Virtual Mouse and Keyboard Actions:**
   The system performs the virtual mouse or keyboard actions. These actions can include can include moving the cursor on the screen, emulating mouse clicks or button presses, scrolling, typing, or any other interactions typically associated with a physical mouse or keyboard.
9. **Output Display:**

The system's output such as cursor movement or the generated keyboard inputs, is displayed on the output device, such as computer screen or virtual reality headset, providing visual feedback to the user.

10. **Interactions Loops:**
    The system continuously repeats the process the capturing, preprocessing, classifying, tracking, recognizing, mapping and performing actions in a loop. Enabling real-time and interactive gesture-based interactions.

## USECASE DIAGRAM

A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. A use case diagram can identify the different types of users of a system and the different use cases and will often be accompanied by other types of diagrams as well. The use cases are represented by either circles or ellipses.

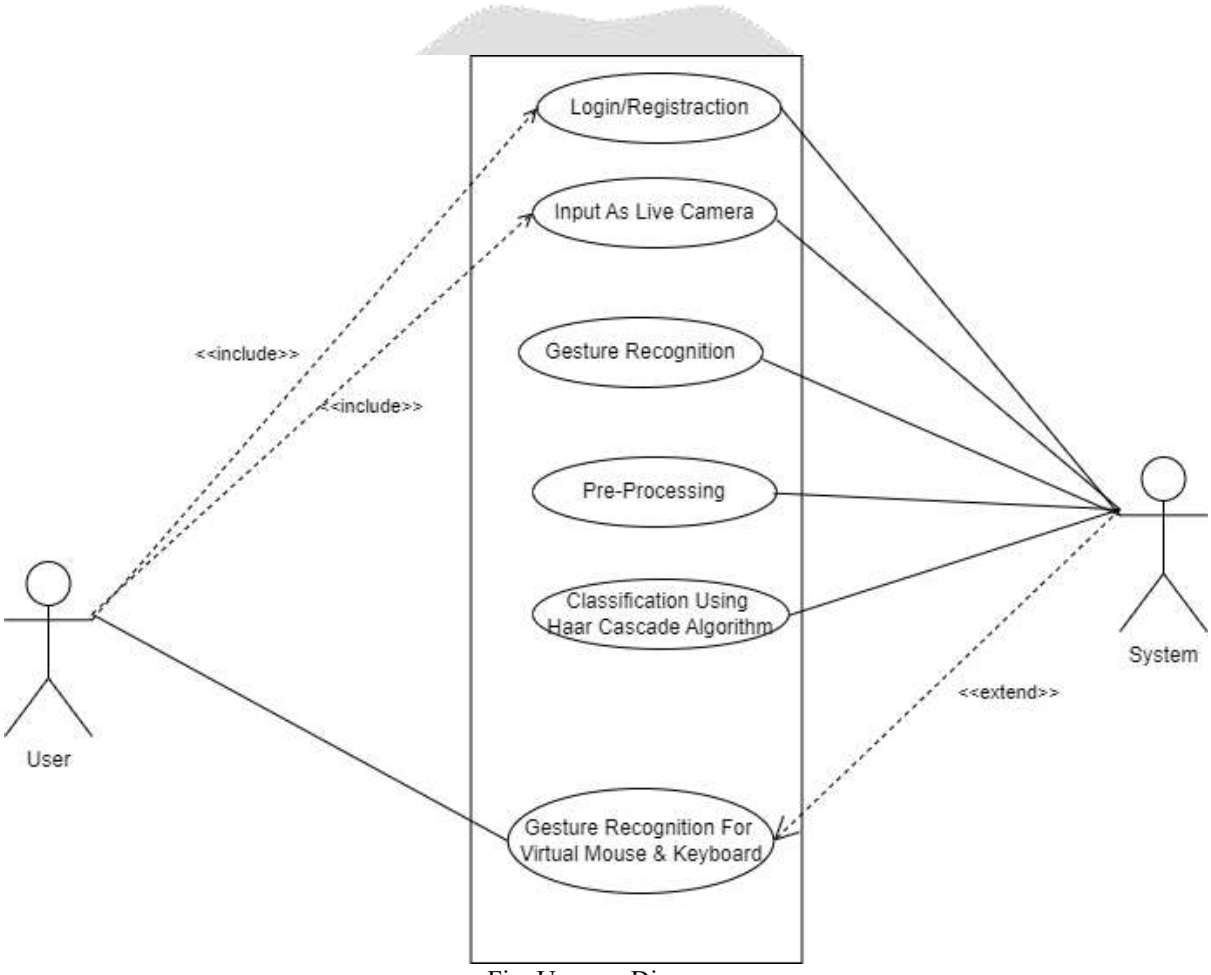


Fig: Usecase Diagram

## FUTURE SCOPE

The system can be designed for medical imagery during surgery without touching patients or any surgical instruments. iii. This mouse system can be applicable as sign language for the dumb & deaf people. It can help them to interact with computing machines.

## CONCLUSION

The main objective of the virtual mouse system is to control the mouse cursor functions by using the eye gestures instead of using a physical mouse and the virtual keyboard is controlled by tracking coloured object.

The proposed system can be achieved by using a webcam or a built-in camera which detects the eye gestures processes frames to perform the particular mouse functions and detect coloured object and perform keyboard function accordingly ,then display to text file. From the results of the model, can come to a conclusion at the proposed virtual mouse and keyboard system has performed very well and has a greater accuracy compared to the existing models and also the model overcomes most of the limitations of existing systems.
Also can use for handicapped persons if they can only move hands,so they can communicate through this proposed system. For the future work, swipe keypads which detect the gestures in air view can also be implemented. This would improve the results when the typing speed is quick.

## REFERENCES:

[1] Vantukala VishnuTeja Reddy1, Thumma Dhyanchand2 , Galla Vamsi Krishna3. "Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition" (2020 IEEE-HYDCON)
[2] Jing-Hao Sun,Ting-Ting Ji,Shu-Bin Zhang "Research on the Hand Gesture Recognition Based on Deep Learning" 2020 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE) [3] Rafael Augusto Da Silva and Antonio Cl "Algorithm for decoding visual gestures for an assistive virtual keyboard" IEEE Latin America Transactions ( Volume: 18, Issue: 11, November 2020)
[4] Kadir Akdeniz1 , Zehra C¸ ataltepe "Dynamic and Personalized Keyboard for Eye Tracker Typing" 2020 24th Signal Processing and Communication Application Conference (SIU)
[5] Ue-Hwan Kim, Sahng-Min Yoo, and Jong-Hwan Kim, "I-Keyboard: Fully Imaginary Keyboard on Touch Devices Empowered by Deep Neural Decoder", IEEE Transactions on Cybernetics ( Volume: 51, Issue: 9, September 2021)
[6] Sugnik Roy Chowdhury Sathyabama, Sumit Pathak Sathyabama, M.D. Anto Praveena Sathyabama, "Gesture Recognition Based Virtual Mouse and Keyboard", Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)