JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Survey Paper on the Development of Virtual Mouse Using Hand Gesture

Prof. S. U. Mali
Suyash Joshi
Omkar Bedake
Rutvik Bonde
Laksh Gujarati

Department of Computer Engineering NBN Sinhgad Technical Institutes Campus

Abstract:

The invention of the computer has been a turning point in history. It makes our day-to-day life much simpler. The way to communicate with a computer i.e., Human-Computer Interaction (HCI), is through input and output devices. The mouse is one of the input devices. Mouse is an important part of Human-Computer Interaction (HCI). With evolving technology, the cursor can be controlled without physical mouse hardware but using the tip of a finger, hand gestures, etc. These hand movements can be recorded through a webcam. The movement of the cursor can also be controlled by color detection or fingertip tracking using a webcam. This new era of technology is useful in many fields, in many situations. This can be accomplished using extra hardware components or developing a software system. This is achieved through various image processing techniques and gesture recognition techniques without extra hardware. In this paper, we have proposed a system that uses a hand gesture recognition-based Convolutional Neural Network (CNN) model for controlling the mouse.

Keywords: Virtual mouse, Hand Gesture Recognition, CNN model

I. INTRODUCTION :

A mouse is a pointing device that makes human-computer interaction easy. There has been much development in the types of mice. Beginning from handheld wired mouse to Bluetooth mouse, then touchpad and touch screen to glove- based detector mouse, recently there have been many newly developed image processing techniques. Image processing is a method of performing operations on an image to get useful information. Gesture recognition is a type of image-processing process. Many techniques have been explored till now like 3D convolutional neural network, contour, and convex hull area. Different pre-processing techniques have been invented and employed for accurate and efficient outcomes of the previously mentioned gesture recognition techniques like k-cosine and border-tracing, background subtraction, and computing four motion matrices.

In this paper, we have proposed an interactive way of controlling the movement of the mouse by hand gesture. The in-built webcam is used to capture the live feed. Using Python libraries hand detection in the image and background subtraction is done. The pre-processed image is then passed to the CNN model for recognizing the gesture and accordingly, the cursor is controlled. CNN model is trained by the dataset prepared. In the rapidly evolving landscape of human-computer interaction, the development of innovative and intuitive input devices has become a focal point. One such advancement is the creation of a virtual mouse controlled through hand gestures—a cutting-edge technology that promises to redefine the way we interact with digital interfaces. By harnessing the power of computer vision and machine learning, this project aims to seamlessly translate hand movements into precise cursor control, offering users a natural and immersive alternative to traditional input methods. In this endeavor, we explore the fusion of hardware and software, pushing the boundaries of technology to create a virtual mouse that not only enhances user experience but also opens doors to a new era of intuitive computing.

II. LITERATURE REVIEW :

Ahmed, Muhammad, et al., (2021) [6] provide a very good explanation of various techniques of object detection such as Object Detection (OD), Salient Object Detection (SOD), and Category-specific Object Detection (COD). It reviews different existing methods that employ deep learning and compares them based on their performance in challenging environments. It depicts recent research and the advancement in this field comparing their performance using output, and time required and comparing them with traditional methods. It also discusses publicly available datasets for this task as well as explains evaluation metrics used for

comparison. It also gives a brief idea about the shortcomings and direction in which further advancements can be made.V. Tiwari et al., (2020) [7] use VGC16 as a pre-trained model to achieve image classification. It compares its results with different models such as Baseline CNN, and the block VGG model. The paper also incorporated the data augmentation in the VGC3 model to study its effect on accuracy. The implemented VGC16 model has an accuracy of 98.97% which is quite an improvement over the accuracy of 55.075%, 74.561%, and 61.404% of Baseline CNN, VGC3, VGC3+Data augmentation respectively. T. Palleja et al. (2008) the system controls the mouse through the movements of the head and facial features. It uses an image processing algorithm and computes four motion matrices. A frame cumulative image is analyzed and the region of interest is used for the detection of the movement. The process takes some time thus reducing the speed at which the mouse responds. Li Wensheng, Deng Chunjian, and Lv Yi (2010) [8] proposed a model consisting of two parts-server side and client side. Server-side: adaptive online training is used for the detection of a finger using a color histogram, tracking of a finger using a mean shift algorithm, constructing appropriate messages, and sending such messages to the client, and the mouse movement is done by the BP neural network. Client-side responds to the messages and obtains the coordinates of multiple fingertips by calling the API function provided by the server, then does the appropriate message processing to achieve control of the application. The results are not consistently using adaptive online training because of the varied skin color.Kumar, P. & Verma, J. & Prasad, Shitala (2012) proposed an additional device called a data glove. The data captures the current position of the hand and the angles between the joints and then these features are used to classify the gestures using the K-NN classifier. K. S. Varun et al., (2019) [11] developed models which are based on color detection and mouse movement based on highlighted color which is given by the user for the movement of the mouse. A two-figure input can be seen where it is forming two rectangles and an average point from both figures. That point will be acting like a mouse pointer. As the point shifts, the mouse pointer concurrently adjusts its position during runtime.

. So using this the movement of the mouse can be implemented. The updating of the mouse pointers depends on the position of the prespecified colored caps in the mask that is created for understanding the system. The created mask is converted from an RGB background to a black and white that will be used for the detection of the prespecified colored objects that will help with the movement of the mouse. If the prespecified colored caps camouflage with the background they will not be detected and thus no movement of the mouse is possible. S. R. Chowdhury et al., (2020) [12] proposed a model consisting of both mouse and keyboard. The Mouse uses a convex hull process for its working, defects are captured or read, and using these defects the functions of the mouse are mapped. The process of this image recognition process solely focuses on defects and conditional statements, the convex hull takes the gap of the fingers as defects, so it can be used for multiple gestures and mapping commands. Sai Mahitha G. and colleagues (2021) introduced an innovative model for mouse cursor control, utilizing hand gestures captured by a computer's webcam. This model relies on cutlet gestures or colored caps to direct the mouse cursor, enabling operations such as line dragging and left clicking. Additionally, it facilitates the transfer of lines between two computers within a single network, and it operates effectively with a low-resolution webcam for hand tracking.

The proposed model builds upon the foundation laid by V. Reddy et al. (2020). Key features from their background deduction algorithm have been integrated into the design. Notably, the convex housing algorithm in the original model has been replaced with a convolutional neural network (CNN) to enhance sensitivity. For implementation, an erected webcam captures live feed, and the system activates processes upon detecting a hand within the webcam content range. The model's workflow involves hand detection using a convex housing algorithm, followed by image pre-processing, including background deduction. Subsequently, the model links the hand boundary and removes extraneous elements. To account for the challenge of moving the cursor across the desktop screen from a lower webcam position, a relative match system is employed. The pre-processed image is then forwarded to the CNN model for gesture recognition. This model is trained on a dataset featuring pre-processed images of colorful hand gestures, encompassing actions like navigation, clicking, zooming, and scrolling. Mouse control is executed through the auto-py Python library based on the CNN model's predictions. This system allows us to direct the system pointer by using our finger bearing color caps or tapes and the operations like dragging of files and the left click would be performed by using distinct finger gestures. Additionally, the system facilitates file transfer between two PCs within the same network. This innovative system relies on a low-resolution webcam serving as a sensor to track the user's hands in two dimensions. If the prespecified colored caps camouflage with the background they will not be detected and thus no movement of the mouse is possible.

Tran, DS., Ho, NH., Yang, HJ. et al.,(2021) [14] proposed virtual mouse method using RGB-D images and fingertip detection. The initial step involves extracting the hand region of interest and identifying the center of the palm by utilizing detailed skeleton-joint information from a Microsoft Kinect Sensor version 2. Subsequently, this information is converted into a binary image. Following this, the contours of the hands are extracted and characterized using a border-tracing algorithm. The Kcosine algorithm is then applied to identify the fingertip location, relying on the coordinates of the hand contour. Ultimately, the determined fingertip location is mapped onto RGB images to manipulate the mouse cursor on a virtual screen. However, it is important to note that this study still grapples with several limitations, primarily stemming from the inherent constraints of the Microsoft Kinect. V. V. Reddy et al., (2020) [15] paper implements algorithm which uses two types of methods for implementation of mousecontrol. One includes the usage of color caps and other uses the recognition of bare hand gesture recognition. It is divided into two methods; ‘fingertip detection’ using colored caps and‘gesture recognition’. Fingertip detection involves the coloridentification, circle identification and guessing the fingers. Gesture recognition involves skin color identification, contourdetection and convex hull formation and then guessing the gesture. Respective mouse operation can be done accordingly. We have used this model as our base for the research. We havestudied the background subtraction from the model to be incorporated in our project. In this model convex hull is used for the recognition of the gesture, to increase the accuracy of this model we will be replacing convex hull with convolutional neural network.

III. PROPOSED SYSTEM:

Through the integration of hand gesture recognition, the proposed system seamlessly amalgamates cutting-edge technology to deliver a responsive and user-friendly virtual mouse. The hand detection methodology employs Convolutional Neural Networks (CNNs) or Haar cascades, utilizing a webcam for instantaneous video input, to accurately identify and isolate the user's hand within

the frames. Subsequently, an advanced hand landmark detection model, such as Media-Pipe, pinpoints key hand landmarks in three dimensions, including fingertips and knuckles.

The system then utilizes a machine learning model, whether a support vector machine or neural network, to categorize specific hand movements based on the spatial arrangement of landmarks. This amalgamation ensures a coherent system that operates in real-time, providing users with a seamless and engaging digital interface experience. To enhance user engagement, feedback—be it visual or auditory—signals the successful recognition of gestures and corresponding cursor activities. All in all, this approach represents a creative and efficient means of digital interaction without necessitating the use of traditional input devices.

IV. METHODOLOGY :

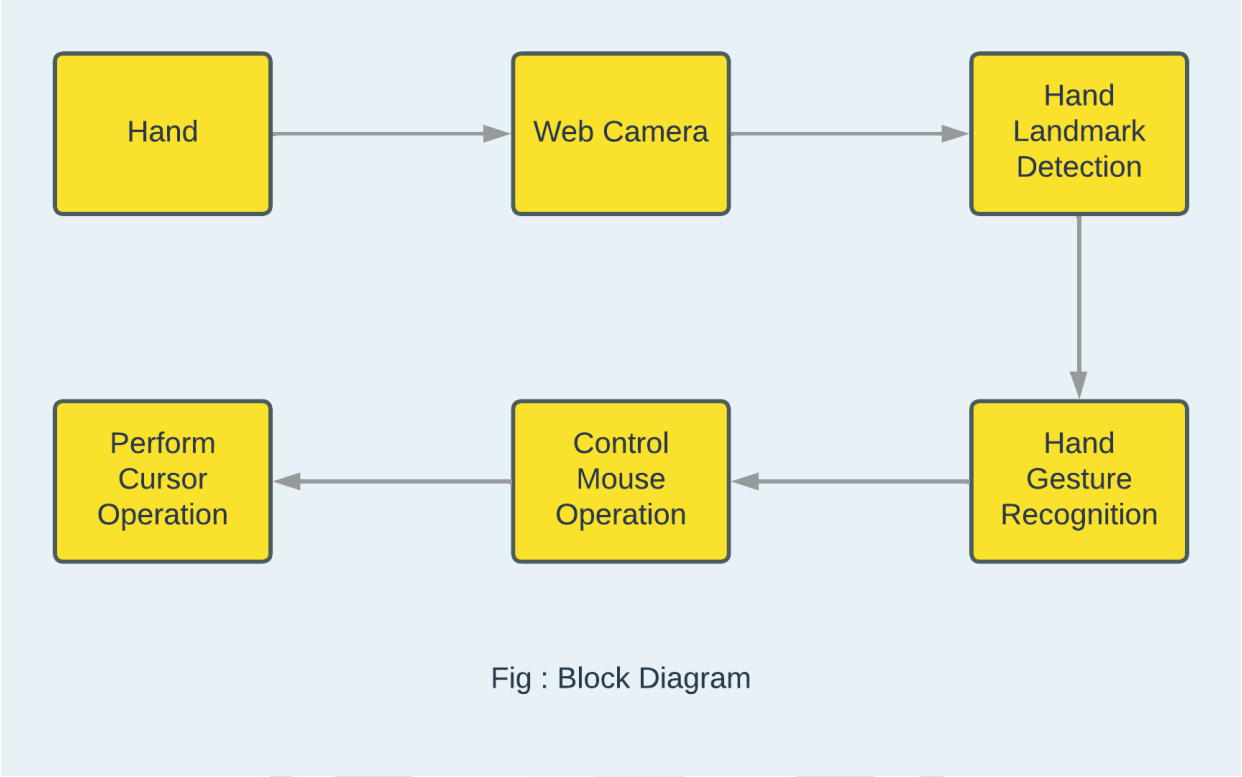

Fig : Block Diagram

Methodology Description:
Here's a revised version of your methodology description:

1. Hand Detection:
   - Input: Utilize a webcam to capture real-time video feed.
   - Algorithm: Apply a hand detection algorithm, such as Convolutional Neural Networks (CNNs) or Haar cascades, to identify and locate the hand within the video frames.
   - Procedure: Extract the region of interest containing the detected hand.

2. Hand Landmark Detection:
   - Input: Use the region of interest obtained from the previous step.
   - Algorithm: Employ a hand landmark detection model, such as MediaPipe, to identify key landmarks on the hand, including fingertips, knuckles, and the palm.
   - Procedure: Obtain the 3D coordinates of the detected hand landmarks.

3. Hand Gesture Recognition:
   - Input: Utilize the 3D coordinates of hand landmarks.
   - Algorithm: Implement a machine learning model, such as a neural network or a support vector machine, to classify and recognize specific hand gestures based on the spatial arrangement of landmarks.
   - Procedure: Identify the recognized hand gesture from a predefined set (e.g., swipe, pinch, etc.).

4. Control Mouse Operation:
   - Input: Use the recognized hand gestures.
   - Algorithm: Map each recognized gesture to a corresponding mouse operation (e.g., cursor movement for swipe, click for pinch).
   - Procedure: Generate commands for mouse control based on the recognized gestures.

5. Perform Cursor Operation:
   - Input: Utilize mouse control commands from the previous step.
   - Operation: Apply cursor movement and action functionalities based on the entered commands. Utilize 2D or 3D coordinates of the hand to control the cursor position on the screen.
   - Procedure: Execute cursor movements, clicks, and other operations as required by the recognized hand gestures.

6. Integration:
   - Combine all components into a cohesive system where hand detection, landmark detection, gesture recognition, and mouse control seamlessly work together.
   - Optimize the system for real-time performance to ensure minimal delay between hand gestures and cursor operations.

7. User Feedback:
   - Provide visual or auditory feedback to the user to indicate successful recognition of gestures and corresponding cursor operations.

By following this methodology, a virtual mouse using hand gesture recognition can be developed, offering an intuitive and interactive way for users to control digital interfaces.

V. Conclusion:

In this design we proposed an effective way of controlling mouse nearly using real-time camera. We've proposed operations like mouse navigation, clicking, scrolling (over and down), drone heft and out. Our approach is landing the live feed, abating the background and passing it to the CNN model. The CNN Model gives High accurate results. With using background detection, we can also give good results in complex background. The fashion can be enforced in colorful systems like gaming, discovery of sign language and communication. This type of system can be enforced with emphasis over movement of eyes which can be salutary for cases suffering from palsy.

VI. References:

1. Kumar,P. & Verma,J. & Prasad, Shitala, " Hand data glove A wearable real- time device for mortal- computer commerce " International Journal of Advanced Science and Technology, 2012.

2.X. Zhang andS. Xu," Research on Image Processing Technology of Computer Vision Algorithm" 2020 International Conference on Computer Vision, Image and Deep Learning(CVIDL), 2020,pp. 122- 124.

3. Dinh- Son Tran, Ngoc- Huynh Ho, Hyung- Jeong Yang, Eu- Tteum Baek, Soo- Hyung Kim and Gueesang Lee " Real- Time Hand Gesture Spotting and Recognition Using RGB- D Camera and 3D Convolutional Neural Network "Appl.Sci. 2020, 10, 722.

4.L.D. Singh,P. Das andN. Kar," Apre- processing algorithm for hastily convex casing computation" confluence 2013 The Next Generation Information Technology Summit(4th international Conference), 2013,pp. 413- 418

5.T. Palleja etal." Simple and Robust performance of a Relative Virtual Mouse Controlled by Head Movements" Conference on mortal System relations, 2008,pp. 221- 224.

6. Ahmed, Muhammad etal. " Survey and Performance Analysis of Deep Learning predicated Object Discovery in Challenging surroundings. " Sensors(Basel, Switzerland)vol. 21,15 5116. 28Jul. 2021, doi10.3390/ s21155116

7.V. Tiwari,C. Pandey,A. Dwivedi andV. Yadav," Image Bracket Using Deep Neural Network" 2020 2nd international Conference on Advances in Computing, Communication Control and Networking(ICACCCN), 2020,pp. 730- 733.

8. Li Wensheng, Deng Chunjian and Lv Yi," performance of virtual mouse predicated on machine vision" The International Conference on Apperceiving Computing and Intelligence Analysis Proceeding, 2010,pp. 367- 371.

9.A. Mhetar,B.K. Sriroop,A.G.S. Kavya,R. Nayak,R. JavaliandK.Suma," Virtual mouse" International Conference on Circuits, Communication, Control and Computing, 2014,pp. 69- 72.

10.S.M.S. ShajideenandV.H. Preetha," Hand Gestures- Virtual Mouse for Human Computer Interaction" International Conference on Smart Systems and Inventive Technology(ICSSIT), 2018,pp. 543- 546.

11.K.S. Varun,I. PuneethandT.P. Jacob," Virtual Mouse performance using Open CV" 3rd International Conference on Trends in Electronics and Informatics(ICOEI), 2019,pp. 435- 438.

12.S.R. Chowdhury,S. PathakandM.D.A. Praveena," Gesture Recognition predicated Virtual Mouse and Keyboard" 4th international Conference on Trends in Electronics and Informatics(ICOEI), 2020,pp. 585- 589.

13. SaiMahithaG.,RevanthB.,GeethaG., SirishaR. " Hand Gesture Recognition to apply Virtual Mouse Using Open Source Computer Vision Library Python " Proceedings of International Conference on Advances in Computer Engineering and Communication Systems. Learning and Analytics in Intelligent Systems, vol 20. Springer, Singapore, 2021,pp. 435- 446.

14. Tran,DS., Ho,NH., Yang, HJ. etal. " Real- time virtual mouse system using RGB- D images and fingertip discovery " Multimed Tools Appl 80, 10473 – 10490, 2021.