The Design of Hand Gesture Controlled Virtual Mouse Using Convolutional Neural Network Technique

Nagaraj M. Lutimath [1], Niharika K [2], K Bhuvanesh[3], Mohan M[4] and Muhammad Zidan K M[5]
[1] *Associate professor Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore*
[2345] *Students of Computer Science and Engineering, Dayananda Sagar Academy of Technology and Management Bangalore*

***Abstract*— This exploratory paper suggests a method for using computer vision and deep learning models to implement the functionalities of the cursor based on hand gestures. These models aim to replicate the swaying motion of human hand gestures and hold the key to further improving the performance of such computer vision solutions. In today's technological environment, several technologies are constantly evolving. One such promising idea is the human-machine interface. The concept is to use hand gestures to emulate mouse functionality on the screen without the use of any hardware, just by utilizing finger motions, a process known as gesture recognition. In this paper, we introduce a novel Human-Computer Interaction (HCI) strategy. For the implementation of this system, Python will be utilized, and its dependencies include OpenCV, MediaPipe, and the latest packages such as PyAutoGUI.**

***Index Terms*— Computer Vision, Deep Learning, Human- Computer Interaction (HCI), MediaPipe, PyAutoGUI**

I. INTRODUCTION

These hand motions are efficient means of expressing words, ideas, and emotions. Nonverbal communication use gestures to convey specific messages. A person can convey this information by moving their hands, face, or body. Because of developments in computer vision and artificial intelligence, computers can now record and identify human hand movements, bridging the gap between human-machine contact. The conventional human-machine interface has evolved as a result of virtualization and the long-term shift towards immersive technologies like the metaverse. As technology develops, devices get smaller and smaller. Although some gadgets are invisible, others have gone wireless. The goal is to create a virtual mouse that can recognize gestures. The aim is to use a simple camera rather than a standard or regular mouse to control various mouse cursor functions. This study proposes an effective hand gesture segmentation algorithm based on image preprocessing methods.

II. SCOPE OF PROJECT

The two most common methods for hand gesture recognition are vision-based, which employs image processing techniques with input from a camera, and hardware-based, which requires the user to wear a device. The suggested system is obviously a vision-based system that makes use of camera inputs and image processing algorithms.

Our objective is to offer more gestures so that users may complete more tasks quickly in the future. This proposal suggests a system that only makes use of the proper hand when making gestures. As a result, it will be feasible to use both hands for certain gestures in the future thanks to improvements made to the method that is presently in use.

The end user will save time and effort using this suggested approach, and people who are blind or disabled will be able to utilize computers more easily.

III. IMPLEMENTATION

The following are the steps of implementation of oursystem.

We take real-time pictures with the webcam during this process real-time photos from the computer webcamare mostly acquired.

We process the video frames after breaking them down into individual frames using techniques like image segmentation, object recognition and picture compression.

We extract several fingertips from the picture.

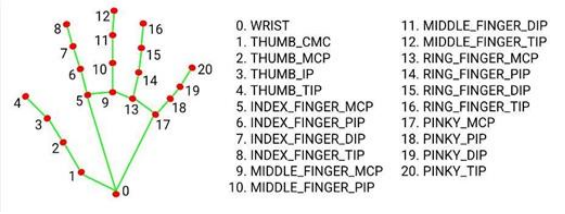


Fig.1. Commonly used hand coordinates

4.Using the tip id as shown in the figure of the relevant finger, detecting the fingertips.

5.Follow the pointer's movement.

6.Executing various mouse operations by assigning a finger id to each action.

IV. METHODOLOGY

Image Acquisition

OpenCV is a popular open-source computer vision library that can be used for a wide range of applications, including image acquisition. The library provides a set of functions that allow you to interface with different types of cameras, video streams, and image files, and to acquire images and video frames from them.

Hand recognition

Hand recognition is a technology that involves detecting and interpreting human hand gestures and movements through the use of computer vision algorithms and machine learning techniques.

Mediapipe is a computer vision library developed by Google that can recognize hand gestures using machine learning algorithms. The library uses a combination of deep neural networks and computer vision techniques to accurately detect and track hand landmarks in real-time.

Mediapipe uses a single-shot detector (SSD) object detection algorithm to detect the presence of a hand in an image or video frame. Once a hand is detected, the library uses a series of convolutional neural networks (CNNs) to estimate the 3D coordinates of 21 hand landmarks or key points, such as the tip of the thumb or the base of the wrist.

These hand landmarks are then used to determine the hand pose and recognize hand gestures. To recognize hand gestures, Mediapipe uses a combination of heuristics and machine learning models trained on large datasets of hand poses and gestures. The library provides pre-trained models for recognizing a variety of hand gestures, including open hand, closed fist, thumbs up, and others.

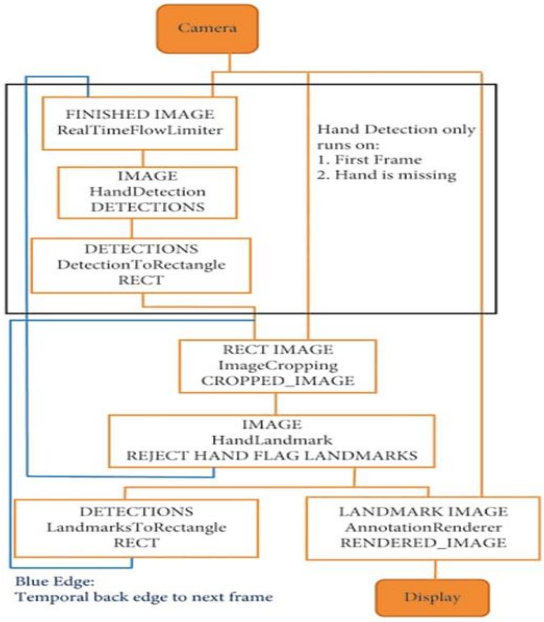
Hand tracking

CSRT(Channel and Spatial Reliability Tracker) is a real-time object tracking algorithm used in computer vision. The CSRT algorithm uses a multi-channel feature representation to improve tracking performance. The algorithm is trained on a sample of the target object using a discriminative correlation filter. The filter is then used to track the object by searching for the most similar response in the subsequent frames. The CSRT algorithm has been shown to achieve state-of-the-art tracking performance on a wide range of object tracking benchmarks, including the VOT2016, VOT2017, and OTB-100 datasets. It is widely used in applications such as surveillance, robotics, and augmented reality.

Mouse control

To stimulate the actions of the mouse we make use of PyAutoGUI library. PyAutoGUI is a Python library for automating GUI interactions on the desktop. It can simulate keyboard and mouse inputs, and control the movement and clicking of the mouse. Some common use cases for PyAutoGUI include automating repetitive tasks such as data entry or testing, creating macros or scripts for applications that don't have built-in automation capabilities, and controlling the desktop for accessibility purposes.

Flow chart



V. RESULT ANALYSIS

The next seven figures demonstrate the different hand movements that are employed in our system to carry out the different functions of the virtual mouse.
If all five figures in Fig. 3 are up as they should be, the system interprets this as Neutral and activates the cursor. The index finger must be spread wide, as shown in Fig. 4, to enable the left button to be clicked. The index finger must be spread widely, as shown in Fig. 5, to enable the right button to be clicked. Closing all fingers as indicated in Fig. 6 will enable the drag and drop action, which can be helpful for moving files from one spot to another. If both index and middle fingers as shown in Fig.7 are up and held wide-open it performs the cursor control operation. By pinching the index and Thumb fingers as shown in Fig.8 we can perform scrolling action. Double- click is equal to pressing a mouse button twice which performed by joining both index and middle finger as Gesture Controlled virtual Mouse is simulated using CSRT tracking algorithm to obtain better accuracy.
shown in Fig.9.The Hand

VI. UNIT TESTING

Serial Number	Input	Expected Output	Actual Output	Pass/Fail
1	No hand	No hand detected	No hand detected	Pass
2	5 fingers up	No change in cursor movement.	No change in cursor movement.	Pass
3	Index and middle finger is up	Cursor movement	Cursor movement	Pass
4	Index finger is up	Right click	Right click	Pass
5	Middle finger is up	Left click	Left click	Pass
6	Both index and middle finger is close	Double click	Double click	Pass

Fig 3 .Neutral Gesture

Fig.4. Left Click

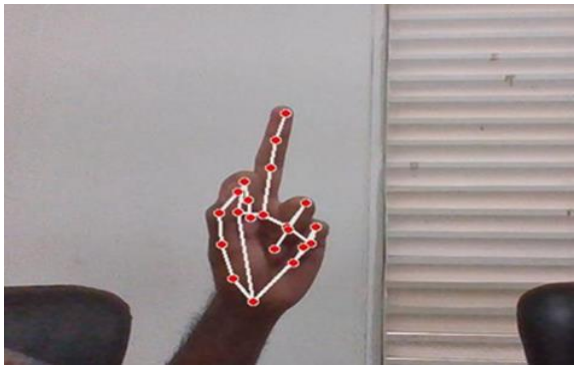


Fig.5. Right Click

Fig.6. Drag and Drop

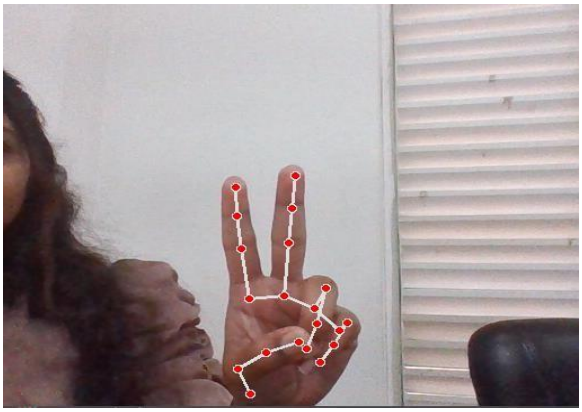


Fig.7. Cursor Control

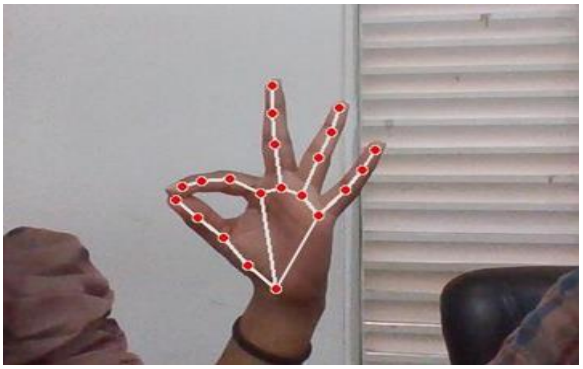


Fig.8. Scrolling

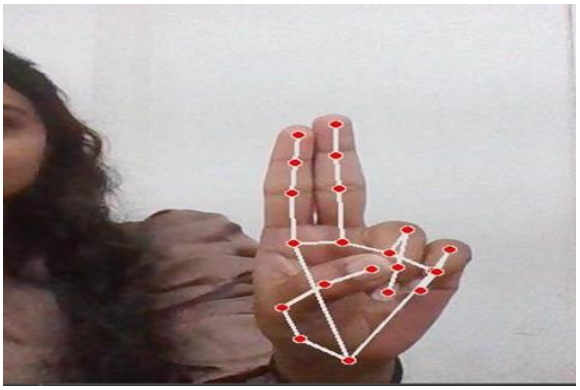


Fig.9. Double Click

The current version of the Mediapipe Hands model has an average precision of 95.7% and an average recall of 91.6% on the public benchmark dataset.

The performance of the CSRT algorithm which is used in our hand tracking model with a human-robot interaction scenario is found to be achieved an accuracy of 98.7%.

VII.CONCLUSION

In this project, we are developing a system that uses a live camera to mimic the functions of a mouse. Our technology, which is based on a computer vision algorithm, is capable of handling the majority of mouse activities. Convolutional Neural Network would be used by our system to recognize the hand landmarks. This landmarks are the tracked by the CSRT algorithm to identify the various gestures performed by the user. This technique would streamline presentations while saving space. This project was created using Python 3.7 (64-bit) and open-source modules, making it ideal for improvements in the future.

REFERENCE

[1] Abhisek Banerjee, Abhirup Ghosh, Koustuvmoni Bharadwaj," Mouse Control using a Web Camera based on Color Detection", IJCTT, vol.9, Mar 2014.

[2] Amrutha D Bhumika M Shivani Hosangadi, Shravya, Manoj. H. M, "Time Static and Dynamic Hand Gesture Recognition using CNN", ISSN No : 2278-0181, 2022

[3] CN Sujatha , S.P.V Subbarao , P.Preetham , P.Surajvarma Y.Upendra, "Virtual Mouse Using Hand Gestures", ISSN No: 2320-2882, 2022

[4] Abhik Banerjee, AbhirupGhosh, Koustuvmoni Bharadwaj," Mouse Control using a Web Camera based on Color Detection", IJCTT, vol.9, Mar 2014

[5] Angel, Neethu.P.S, "Real Time Static & Dynamic Hand Gesture Recognition", International Journal of Scientific & Engineering Research Volume 4, Issue3, March-2013.

[6] Chen-Chiung Hsieh and Dung-HuaLiou," A Real Time Hand Gesture Recognition System Using Motion History Image"icsps, 2010

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

[8] Baohuaqiang, yijiezhai, mingliangzhou." SqueezeNet and Fusion Network-Based Accurate Fast Fully Convolutional Network for Hand Detection and Gesture Recognition". IEEE.2021.

[9] Mohammad Naveed Khan,Sheikh Auqid Irfan,Shah Faisal Ahmad,Mr. Khalid Makhdoomi, "Mouse Cursor Control Using Hand Gesture Recognition", 2021.

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