VIRTUAL MOUSE: A Literature Review

ASWIN SIVADAS
B.tech Student
Dept.of CSE
FISAT
Angamaly

ASWIN C
B.tech Student
Dept.of CSE
FISAT
Angamaly

ASHUAL JESON
B.tech Student
Dept.of CSE
FISAT
Angamaly

ADHIL MUHAMMAD K
B.tech Student
Dept.of CSE
FISAT
Angamaly

ANURANJ P
PROFESSOR
Dept.of CSE
FISAT
Angamaly

***Abstract*—The development of technology in the area of human-computer interaction (HCI) is best illustrated by the PC mouse. A wireless or contactless mouse depends on the device or an external power source, such as a battery, even if it may not be entirely free from gadgets. In today's technological age, it is also advised to maintain social distance and refrain from handling items that have been handled by several persons, particularly during the COVID pandemic.**

In this study, a novel method for employing a real-time camera to operate a mouse is presented. We propose rebuilding the hardware, not only adding more buttons or changing the tracking ball of the mouse. Our suggested approach uses a camera and computer vision technologies to accomplish tasks.

Scrolling and other actions, including clicking, can imitate the capabilities of conventional mouse devices. The project describes how to construct a mouse control system with this technique.

Keywords- OpenCV, Hand recognition, Gesture recognition

I. Introduction

Using a virtual mouse, a user can operate a computer, laptop, or smart pad without actually touching it. As an alternative, the user only makes movements in front of a camera that is either separate or built into the system or any other Human Interface Device (HID). Due to its dual functionality as a mouse and a marker, the gadget can be used in meetings and classrooms in place of conventional wired mouse, chalkboards, and markers.

For people who desire to operate a computer or other system in a more intuitive and natural way, the virtual mouse gadget is a helpful tool. It can be especially useful in educational contexts because it enables teachers to connect with pupils and demonstrate topics without the use of physical boards and markers.

The ability to collaborate and exchange ideas without the use of conventional mouse and keyboard input devices makes it effective in conference settings as well. The virtual mouse can also be a practical choice for folks who favour using gestures to operate their devices rather than actual buttons.The virtual mouse, as a whole, is a flexible and creative instrument that provides various advantages for users in a number of contexts.

Identify applicable funding agency here. If none, delete this.

II. Literature Survey

A. Video Hand Gestures Recognition Using Depth Camera and Lightweight CNN

[1]In their study, David Gonzalez Leon, Jade Grol, and Sreenivasa Reddy Yeduri suggest utilising a depth camera and a compact convolutional neural network (CNN) model to recognise hand motions in video. In recent years, significant research has been conducted on the recognition of human actions or gestures, and the employment of sensor and artificial intelligence technologies has grown in popularity as a means of enhancing human autonomy. Potential uses for hand gesture recognition include the identification of sign languages, smart home technology, autonomous vehicles, healthcare, augmented and virtual reality, driver monitoring in autonomous vehicles, and robotic surgery.

B. Real-Time Gesture Detection Based on Machine Learning Classification of Continuous Wave Radar Signals

[2] This study compares traditional machine learning (CML) techniques with cutting-edge machine learning (NML) strategies like neural networks (NNs) in order to investigate several approaches for accurately categorising simple human motions captured by a low-cost radar system. It examines the interplay between several approaches, such as threshold detection (THD) algorithms, and seeks to ascertain the extra features needed for each technique (THD, CML, NML) to function well, as well as the accuracy that can be anticipated and attained with each strategy. The study also looks into the viability of using inexpensive standalone computers and micro controller units (MCUs) to run the envisioned networks.

C. Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation

[3] Making systems that can recognise and translate the hand gestures and body motions used in sign languages for deaf communities is the goal of sign language recognition. These tools can be used to translate sign languages into spoken or written languages or to improve communication between deaf and hearing people. However, due to the complexity and diversity of sign languages, as well as the variation in how individual users make signs, sign language detection is a challenging undertaking. The approach for recognising sign language described in this research makes use of deep learning, a type of machine learning that involves training

artificial neural networks on sizable data sets. Deep learning models are useful for a variety of tasks, including speech recognition, image and video recognition, and natural language processing because they can understand complex patterns and correlations in the data. Deep learning models may be taught to recognise and classify diverse hand gestures in the context of sign language identification utilising a variety of hand gesture representation methodologies. These methods could include skeletons, which are reduced representations of the hand and finger bones, depth maps, which display the distance of objects from the camera, or 2D photographs of the hand. The performance of the model can be influenced by the type of representation method chosen, as well as by the model's complexity and the quantity of training data required.

D. A Hand Gesture Based Interactive Presentation System Utilizing Heterogeneous Cameras

[4]It is possible to use a hand gesture-based interactive presentation system in a variety of locations, such as conference rooms, classrooms, or other venues where presentations are delivered. This system employs a network of cameras with varied capabilities or types to detect hand gestures for controlling presentations. Because it enables presenters who must be mobile to manage the presentation without the usage of a tangible object like a remote control, this hand gesture-based interactive presentation system can be helpful. The presenter may be able to walk around the room while presenting the presentation thanks to this. The precision and dependability of detecting hand motions can be enhanced by the use of numerous cameras, each with a separate set of capabilities, in this hand gesture-based interactive presentation system. This is due to the possibility that various cameras can record hand movements in a range of settings, such as in low light or with high resolution. The system may be more effective at identifying hand movements in a range of situations by utilising a network of cameras with these various capabilities. The accuracy and reliability of recognising the motions could be increased by a system that enables users to control presentations using hand gestures that are recognised by a network of cameras with different capabilities. This kind of technology might be advantageous for presenters who need to be mobile and unrestricted by physical gadgets, and it might be employed in a variety of locations including conference rooms or classrooms.

E. Robust Hand Gesture Recognition Based on RGB-D Data for Natural Human–Computer Interaction.

[5]The approach for identifying both static and moving hand motions is presented in this paper using RGB-D data. The technique involves applying the Distance Transform algorithm to locate the palm centre and extract the hand gesture shape. The K-Curvature-Convex Defects Detection technique is used to locate the fingertips, and the

The angles between the fingertips and the distances between the pixels on the palm centre and the hand gesture contour are employed as supplementary features to build a multi modal feature vector. The static hand movements are then reliably identified using a recognition algorithm. The approach described in this research uses RGB-D data to identify both static and moving hand gestures. Extraction of the hand gesture contour and identification of the palm centre and fingertips are steps in the static hand gesture recognition process. After that, a multi modal feature vector is created using auxiliary information such the angles between the fingertips and the lengths between pixels on the palm contour to the palm centre. The static hand movements are then reliably recognised by a recognition algorithm. The Euclidean distance between the hand joints and the shoulder centre joint, along with the modulus ratios of skeletal features, are used to create a unifying feature descriptor for the detection of dynamic hand gestures. The dynamic hand motions are then recognised using a better dynamic time warping method. Finally, a low-cost, real-time application enabling natural interaction with a virtual environment using hand gestures is developed using the static and dynamic hand gesture detection algorithm, which has been validated and proven through extensive trials.

F. Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks

[6]A deep learning model for identifying hand movements in videos is presented in the article. In order to identify both short-term and long-term elements in the visual input while minimising computing demands, the model integrates a number of strategies.

In order to generate an RGB image and an optical flow snapshot, the model first separates the video into groups of frames. From each group, one frame is then chosen. To extract features, these photos are merged and fed through a convolutional neural network (ConvNet). The Long Short-Term Memory (LSTM) network then generates the final classification prediction for the hand gesture using the output from all of the ConvNets.

G. Continuous Finger Gesture Spotting and Recognition Based on Similarities Between Start and End Frames

[7]Based on the similarity between the start and finish frames of the gesture, a method for continually recognising and recognising finger movements is provided in this study. In order to recognise and categorise movements in real-time, a convolutional neural network (CNN) trained on a large dataset of finger motions is used. The results of the experiments demonstrate that the suggested method works effectively and achieves high accuracy in a variety of lighting and background situations. The suggested approach has the potential to be applied in a variety of fields, such as virtual reality, sign language recognition, and human-computer interface. Humans communicate and interact with computers and other systems naturally and intuitively by using finger gestures.

H. Gesture-Based Human Machine Interaction Using RCNNs in Limited Computation Power Devices

[8] Manual hand gestures can be used in real-time to allow humans to communicate with and control computers

or other electronic devices using hand movements. This is made possible through the use of motion tracking, computer vision, and machine learning technologies. Hand gestures can be used to carry out a range of tasks, such as navigating menus, controlling media playback, and interacting with virtual or augmented reality environments. While hand gesture-based interaction has the potential to enhance the usability and accessibility of computing devices, it also presents challenges in terms of gesture design and technology accuracy and reliability.

I. A hand Gesture Based Interactive Presentation System Utilizing Heterogeneous cameras.

[9]Hand gestures are detected by a network of cameras with different types or capabilities. This system could be used in various settings, such as conference rooms, classrooms, or other locations where presentations are given. One potential use of this type of system is to enable the presenter to control the presentation without the need for a physical remote control or other device. This could be especially beneficial for individuals who are giving a presentation and need to be mobile, as it would allow them to move around the room without being tied to a device. The use of a network of cameras with different capabilities, or heterogeneous cameras, in the system could lead to improved accuracy and reliability in detecting hand gestures. Different cameras may have unique features, like higher resolution or the ability to capture hand gestures in low light conditions. By using a variety of cameras, the system may be able to more effectively detect hand gestures in various situations. In summary, a hand gesture-based interactive presentation system that utilizes heterogeneous cameras could be a useful tool for individuals giving presentations, allowing them to control their presentations in a more intuitive and natural way.

J. On-Air Hand-Drawn Doodles for IoT Devices Authentication

[10]This research paper proposes a new method for authenticating human interaction with Internet of Things (IoT) devices using hand gestures in the air to create virtual hand-drawn passwords. This approach is particularly relevant during the COVID-19 pandemic as it allows for secure authentication without physical contact. The proposed method utilizes a computer vision technique with a single camera, two lightweight deep CNN models, and a Kalman filter for signal processing to correct the path of the drawn line in the air. The proposed method for authenticating human interaction with IoT devices using hand gestures in the air is intended to be a simple and secure alternative to traditional authentication methods. By using hand gestures to create a virtual hand-drawn password, users are able to interact with the device in a natural and intuitive way. The method utilizes a computer vision technique with a single camera and two lightweight deep CNN models to accurately detect and interpret the hand gestures. One potential benefit of the proposed method for authenticating human interaction with IoT devices using hand gestures in the air is its potential to improve accessibility for users with disabilities. Traditional authentication methods, such as passwords or fingerprint scanners, may be difficult or impossible for some users to access. By using hand gestures as an alternative method, more users may be able to successfully authenticate and interact with the device. In addition, the proposed method may be more convenient for users in certain situations. For example, if a user is wearing gloves or has dirty hands, they may be unable to use a traditional touch-based authentication method. In these cases, the ability to authenticate using hand gestures in the air could be a useful alternative. Overall, the proposed method for authenticating human interaction with IoT devices using hand gestures in the air offers a simple, secure, and potentially more accessible option for protecting against unauthorized access. Its combination of computer vision techniques, deep CNN models, and signal processing with a Kalman filter allows it to be more accurate and efficient than existing approaches, and it has been shown to be acceptable to users in terms of usability and satisfaction.

III. COMPARISON

Title	Dataset	Methodology	Disadvantages
Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks,2021	Jester dataset and Nvidia.	STSNN model	The current optical flow algorithm is still costly..
Continuous Finger Gesture Spotting and Recognition Based on Similarities Between Start and End Frames ,2022	VIVA and Nvidia & Deep network approach,Hand crafted based approach	This proposed model have robustness to scale and orientation change	Gesture spotting is not perfect
Gesture-Based Human Machine Interaction Using RCNNs in Limited Computation Power Devices,2021	EMG Dataset	Region based CNN approach	High computation capaciies are needed for the majority of the problems that use CNN
Human-Computer Interaction Using Manual Hand Gestures in Real Time,20	Nvidia	PPI together to DMA	When implemeting a recursive funtion,a space limitation problem was faced since the ADSP BF533 processor must save the context in each iteration
A hand Gesture Based Interactive Presentation System Utilizing Heterogeneous Cameras	Nvidia	ML and Computer Vision	Currently one gesture interaction is limited to one hand ,web camera is only used for caliberation——
A Hand Gesture Recognition Sensor Using Reflected Impulse	Jester dataset	(1-d)CNN,IR sensors	Low accuracy for some letters.
On-Air Hand-Drawn Doodles for IoT Devices Authentication During COVID-19	Nvidia	CNN and Kalman filter.	Cant use in the dark room Causes zigzag line due to unstable .
Video Hand Gestures Recognition Using Depth Camera and Lightweight CNN	VIVA	Depth camera and light weight CNN	Camera Hardware's are expensive Need separate hardware's.
Real-Time Gesture Detection Based on Machine Learning Classification of Continuous Wave Radar Signals	Nvidia and jester	Threshold detection (THD) classical machine learning (CML) Support vector machine (SVM) Stochastic gradient descent (SGD).	Need separate radar system. .
Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation	Nvidia	3D CNN	3DCNN requires more parameters than 2DCNN which is one of its disadvantages.

TABLE I
COMPARISON TABLE

IV. Methodology

This paper presents a proposed AI virtual mouse system that uses hand gestures and hand tip detection to perform mouse functions on a computer using computer vision. The goal of the proposed system is to enable computer mouse cursor functions and scroll functions to be performed using a web camera or built-in camera on the computer, instead of using a traditional mouse device. Hand gestures and hand tip detection are used as a human-computer interface (HCI) with the computer. The AI virtual mouse system allows for the tracking of the fingertip of the hand gesture using a built-in camera or web camera, and enables the performance of mouse cursor operations and scrolling functions, as well as the movement of the cursor.

In this system, the user can control the mouse using hand gestures and hand tip gestures captured by a built-in camera or webcam. The captured frames are processed, and the recognized gestures are used to perform specific mouse functions. This method of mouse control is an alternative to using a wireless or Bluetooth mouse, which requires the use of additional devices such as a mouse, dongle, and battery.

The AI virtual mouse system was created using Python and OpenCV for computer vision. It used MediaPipe to track the hands and the tips of the hands. Pynput, Autopy, and PyAutoGUI were used to move around the window screen and perform actions like left click, right click, and scrolling. The model was able to work accurately and effectively in a real-world environment without requiring a GPU, only a CPU.

V. Conclusion

The AI virtual mouse system aims to allow users to control their computer's mouse cursor and perform various functions using hand gestures instead of a physical mouse. This is done by using a webcam or built-in camera to capture and detect hand gestures and hand tip movements made by the user. The system processes these captured frames to perform specific mouse functions, such as left clicking, right clicking, and scrolling. The AI virtual mouse can be used in a variety of realworld situations, including those where there is limited space to use a physical mouse or where individuals may have difficulty using one due to hand mobility issues. It can also be useful during the COVID-19 pandemic as it allows users to control their computer's mouse functions without the need for physical touch, reducing the risk of virus transmission through shared devices.

References

[1] Gibran Benitez-Garcia , Muhammad Haris, Yoshiyuki Tsuda , and Norimichi Ukita, , "Continuous Finger Gesture Spotting and Recognition Based on Similarities Between Start and End Frame" ,VOL. 23, NO. 1, JANUARY 2022 [2] Wenjin Zhang, Jiacun Wang, Senior Member,"Dynamic Hand Gesture Recognition Based on Short-Term Sampling Neural Networks",VOL. 8, NO. 1, JANUARY 2021 [3] David Gonz´alez Le´on , Jade Gr¨oli , Sreenivasa Reddy Yeduri , Daniel Rossier , Romuald Mosqueron , Om Jee Pandey , and Linga Reddy Cenkeramaddi , "Video Hand Gestures Recognition Using Depth Camera and Lightweight CNN", VOL. 22, NO. 14, 15 JULY 202 [4] J. P. Giuffrida, A. Lerner, R. Steiner, and J. Daly, "Upper-extremity stroke therapy task discrimination using motion sensors and electromyography," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 16, no. 1, pp. 82–90, Feb. 2008. [5] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," J. Neuroeng. Rehabil., vol. 9, no. 1, p. 21, 2012 [6]] A. Hollinger and M. M. Wanderley, "Evaluation of commercial forcesensing resistors," in Proc. Int. Conf. New Interfaces Musical Expression, Jun. 2006, pp. 1–4. [7] H.-B. Xie, Y.-P. Zheng, and J.-Y. Guo, "Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multifunction prosthesis control," Physiolog. Meas., vol. 30, no. 5, pp. 441–457, 2009 [8] N. Li, D. Yang, L. Jiang, H. Liu, and H. Cai, "Combined use of FSR sensor array and SVM classifier for finger motion recognition based on pressure distribution map," J. Bionic Eng., vol. 9, no. 1, pp. 39–47, 2012 [9]] R. Xu, S. Zhou, and W. Li, "MEMS accelerometer based nonspecificuser hand gesture recognition," IEEE Sensors J., vol. 12, no. 5, pp. 1166–1173, May 2012 [10] Y. Tenzer, L. P. Jentoft, and R. D. Howe, "The feel of MEMS barometers: Inexpensive and easily customized tactile array sensors," IEEE Robot. Autom. Mag., vol. 21, no. 3, pp. 89–95, Sep. 2014 [11] P. P lawiak, T. So´snicki, M. Nied´zwiecki, Z. Tabor, and K. Rzecki, "Hand body language gesture recognition based on signals from specialized glove and machine learning algorithms," IEEE Trans. Ind. Informat., vol. 12, no. 3, pp. 1104–1113, Jun. 2016. [12] A. R. Fugl-Meyer, L. J¨a¨ask¨o, I. Leyman, S. Olsson, and S. Steglind, "The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance," Scand. J. Rehabil. Med., vol. 7, no. 1, pp. 13–31, 1975 [13]] A. D. Roche, H. Rehbaum, D. Farina, and O. C. Aszmann, "Prosthetic myoelectric control strategies: A clinical perspective," Current Surgery Rep., vol. 2, no. 3, p. 44, 2014. [14] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, Jan. 1967 [15] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: A survey," Artif. Intell. Rev., vol. 43, no. 1, pp. 1–54, Jan. 2015.