



AI Gesture Recognition

¹Joel Abram Santhosh, ²Mr. N. S. L Kumar Kurumeti, M.Tech,(Ph.D)

^{1,2}Department of CSE, Aditya Engineering College, Surampalem, A.P., India

ABSTRACT:

Mouse is one of the most used inventions for interaction with a computer system. Currently, a wireless mouse or a Bluetooth mouse is used to control the system, which uses a physical device for the purpose. In the proposed AI Gesture recognition system, these limitations can be overcome by employing a webcam or a built-in camera to capture hand gestures and hand tip detection using computer vision. Based on hand gestures, the computer can be controlled virtually to perform all the functions that are performed by a physical mouse. The algorithm is based on deep learning and machine learning to detect hand movements and the gestures.

1. INTRODUCTION

A physical mouse has been used since the invention of computer systems. There have been different variations in physical mouse throughout the years as technology has advanced. With the development of technologies in the areas of augmented reality, devices are becoming compact in the form of Bluetooth or wireless technologies. In this proposed system, AI gesture recognition is used to control the cursor using hand gestures. The main objective of the system is to perform computer cursor functions and scroll functions using a web camera or a built-in camera in the computer instead of using a traditional mouse device. With the help of an AI gesture recognition system, we can track the fingertip of the hand by using a built-in camera or web camera and perform the mouse cursor operations and scroll functions.

In the proposed system, Python programming language is used for developing the AI gesture recognition system, along with OpenCV which is a library for mapping hand movements. In the AI gesture recognition system, the model makes use of Media Pipe package for tracking of the hands and for the tracking of the tip of the hands, and also, Pypot, Autopy, and Py Auto GUI packages are used for moving around the windows of the screen of the computer for performing functions such as left click, right-click, and scrolling functions. As the proposed system does not require any physical device for controlling the cursor, we can use it in any circumstance. The main objective of the system is to develop an alternative to perform the mouse functions which can be achieved with the help of a built-in webcam to recognize hand gestures.

2. RELATED WORK

There are some related works carried out on virtual mouse using hand gesture detection by wearing data gloves in hand and also by using colour tips in the hands for gesture recognition. These systems have drawbacks as well, as gloves are not suitable for some users and in some case the recognition is not accurate.

Glove based gesture recognition system: The general idea of this particular project is based on the design of a system that allows the recognition of a set of established gestures, such as the alphabet in American Sign Language, performed by a user with a single hand, using a data glove as an input device. In that system, artificial intelligence tools were used, like artificial neural networks, to achieve an interaction between a computer and the user, where the first classifies and recognizes the gestures executed by the latter. Within the alphabet of ASL, there are static and dynamic gestures: static gestures can be accomplished by maintaining a pose of the hand, while dynamic gestures join a specific pose of the hand with a movement. In this system, the user has to wear data gloves for the AI to recognize the hand and the gestures created by the hand. Some people may have difficulty wearing gloves. The system is also proven to be less accurate as the system cannot detect the hand.

RGB based gesture recognition system: This is a project where you will design a cursor that can move through the desktop and perform actions based on hand gestures. The system's object movement will be based on RGB (red, green, and blue) colour – it can detect RGB colour objects that will function as the mouse. It would help if you imported the Java AWT library to coordinate with the cursor. The system setting uses a webcam to track the movement of the red, green, and blue objects and based on the object movement patterns, accordingly trigger an event.

The cursor movement system will acquire a single frame from the video recorded by the webcam and flip the frame for the user to see. It converts the captured image into a binary image wherein the RGB objects will become white. The system further adds a bounding box around the object that the user can move throughout the display. The user needs to stick coloured tapes at the fingertips tips in order for the system to detect the hand gesture. The system is less accurate as the program fails to detect the gestures.

gestures. Monika B. Gandhi, Sneha U. Dudhane, and Ashwini M. Patil in 2013 [4] proposed a study on “Cursor Control System Using Hand Gesture Recognition.” In this work, the limitation is stored frames are needed to be processed for hand segmentation and skin pixel detection.

Vinay Kr. Pasi, Saurabh Singh, and Pooja Kumari in 2016 [5] proposed “Cursor Control using Hand Gestures” in the IJCA Journal. The system proposes the different bands to perform different functions of the mouse. The limitation is it depends on various colors to perform mouse functions.

Chaithanya C, Lisho Somas, Naveen Wilson, and Abhilash SS in 2018 [6] proposed “Virtual Mouse Using Hand Gesture” where the model detection is based on colors. But, only few mouse functions are performed.

3. PROPOSED METHODOLOGY

The proposed system uses AI for controlling the cursor movements and detecting the gestures. The hand gestures are detected by mapping the coordinates in the hand and calculating the requirements using the MediaPipe framework. The hand gesture is detected by processing the image from the camera and creating a rectangle frame around the hand to specify the area of the windows and then the hand landmarks are used to identify the gestures and perform the functions.

The camera used in the system : The proposed system is based on the frames that are captured by the webcam or the built-in camera of the laptop and with the help of python programming and OpenCV, the video capture object is created.

Capturing and video processing: from the image captured the mapping of the hand is done using the transformational algorithm to detect the movements of the fingertips.

Detecting the hand gestures: the gestures are detected by using the fingertip id that are provided during the mapping of the fingertips found using the MediaPipe. The position of each finger and points are evaluated to identify the particular gesture.

4. ALGORITHM USED

For detection of hand gestures and hand tracking, MediaPipe framework is used, OpenCV library is used for computer vision. The algorithm makes use of machine learning to detect and perform the gestures.

MediaPipe: MediaPipe is a framework which is used for applying in a machine learning pipeline, and it is an open-source framework of Google. The MediaPipe framework is useful for cross platform development since the framework is built using the time series data. The MediaPipe framework is multimodal, where this framework can be applied to various audios and videos. The MediaPipe framework is used by the developer for building and analysing the systems through graphs, and it has also been used for developing the systems for the application purpose. The steps involved in the system that uses MediaPipe are carried out in the pipeline configuration. The pipeline created can run in various platforms allowing scalability in mobile and desktops. The MediaPipe framework is based on three fundamental parts; they are performance evaluation, framework for retrieving sensor data, and a collection of components which are called calculators, and they are reusable. A pipeline is a graph which consists of components called calculators, where each calculator is connected by streams in which the packets of data flow through. Developers are able to replace or define custom calculators anywhere in the graph creating their own application. Single-shot detector model is used for detecting and recognizing a hand or palm in real time. The single-shot detector model is used by the MediaPipe. First, in the hand detection module, it is first trained for a palm detection model because it is easier to train palms. Furthermore, the non maximum suppression works significantly better on small objects such as palms or fists. A model of hand landmark consists of locating 20 joint or knuckle co-ordinates in the hand region.

OpenCV: OpenCV is a computer vision library which contains image-processing algorithms for object detection. OpenCV is a library of python programming language, and real-time computer vision applications can be developed by using the computer vision library. The OpenCV library is used in image and video processing and also analysis such as face detection and object detection.

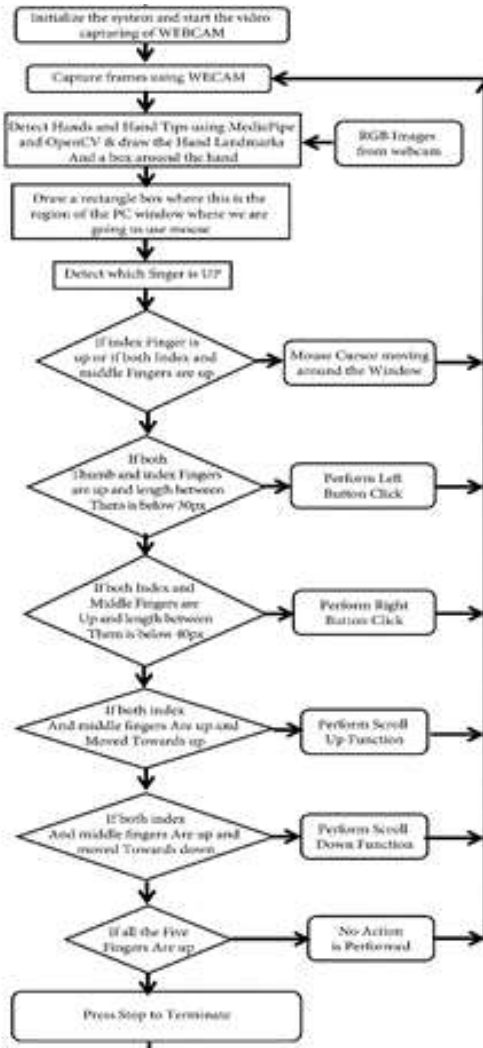


Fig.1 flow chart of AI gesture system

5. EXPERIMENTAL RESULTS

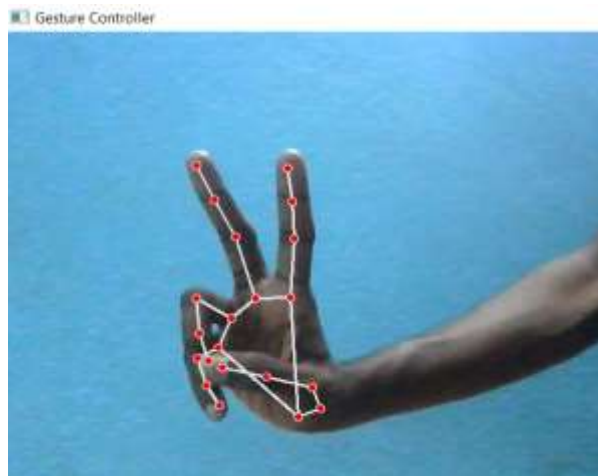


Fig. 2 gesture for moving the cursor



Fig.3 gesture for drag and drop



Fig .4 gesture for scrolling

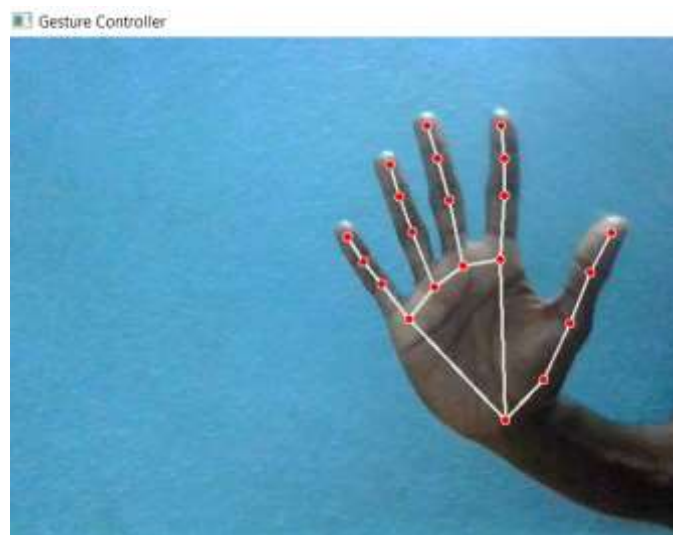


Fig.5 gesture for no movement

5. ADVANTAGES

The AI recognition system can detect the hand gestures without the use of any additional gloves or colour tips.

The accuracy of this system is much more precise than that of all the other proposed systems.

It can be used in areas where there is less space for a physical mouse .

6. CONCLUSION

The main objective of the AI virtual mouse system is to control the mouse cursor functions by using hand gestures instead of using a physical mouse. The system can be achieved by using a webcam or a built-in camera which detects the hand gestures and hand tips and processes these frames to perform the particular mouse functions.

From the results of the model, we can come to the conclusion that the AI virtual mouse system has performed very well and has greater accuracy compared to the existing models and also the model overcomes most of the limitations of the existing systems. Since the model has greater accuracy, the AI virtual mouse can be used for real-world applications, and also, it can be used to reduce the spread of COVID-19, since the mouse system can be used virtually using hand gestures without using the traditional physical mouse.

The model has some limitations such as a small decrease in accuracy in right-click mouse function and some difficulties in clicking and dragging to select the text. Hence, we will work next to overcome these limitations by improving the fingertip detection algorithm to produce more accurate results.

7. FUTURE SCOPE

The method can be developed to handle the keyboard functionalities along with the mouse functionalities virtually which is another future scope of Human-Computer Interaction (HCI).

Amidst the COVID-19 situation, it is not safe to use the devices by touching them because it may result in a possible situation of the spread of the virus by touching the devices, so the AI virtual mouse can be used to control the PC mouse functions without using the physical mouse.

The system can be used to control robots and automation systems without the usage of devices.

2D and 3D images can be drawn using the AI virtual system using hand gestures.

AI virtual mouse can be used to play virtual reality- and augmented reality-based games without wireless or wired mouse devices.

8. REFERENCES

- 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. View at: [Publisher Site](#) | [Google Scholar](#)
- D. L. Quam, "Gesture recognition with a DataGlove," *IEEE Conference on Aerospace and Electronics*, vol. 2, pp. 755–760, 1990. View at: [Publisher Site](#) | [Google Scholar](#)
- 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. View at: [Publisher Site](#) | [Google Scholar](#)
- S. U. Dudhane, "Cursor control system using hand gesture recognition," *IJARCCCE*, vol. 2, no. 5, 2013. View at: [Google Scholar](#)
- K. P. Vinay, "Cursor control using hand gestures," *International Journal of Critical Accounting*, vol. 0975–8887, 2016. View at: [Google Scholar](#)
- L. Thomas, "Virtual mouse using hand gesture," *International Research Journal of Engineering and Technology (IRJET)*, vol. 5, no. 4, 2018. View at: [Google Scholar](#)
- 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. View at: [Publisher Site](#) | [Google Scholar](#)
- 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. View at: [Google Scholar](#)
- 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. View at: [Google Scholar](#)
- 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. View at: [Google Scholar](#)
- J. T. Camillo Lugaresi, "MediaPipe: A Framework for Building Perception Pipelines," 2019, <https://arxiv.org/abs/1906.08172>. View at: [Google Scholar](#)