HAND GESTURE AND VOICE ASSISTANTS

B. Latha¹,*, Sri Sowndarya¹, Swethamalya.K¹, Mr. Ashish Raghuwanshi², Rakhmatova Feruza³, Mr. G. Sathish Kumar⁴

¹Department of Information Technology, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai, Tamilnadu
²Department of Electronics & Communication Engineering, IES College Of Technology, Bhopal, MP 462044 India. research@iesbpl.ac.in
³Tashkent State Pedagogical University, Tashkent, Uzbekistan. rahmatovaferuza67@gmail.com
⁴Assistant professor, Department of Mechanical Engineering, K. Ramakrishnan college of technology, Tiruchirappalli, sathishathi@gmail.com

Abstract- Hand gesture mouse control for systems has received a lot of attention in recent years. Because of its many uses and capability to efficiently connect with machines via human-computer interaction, hand gesture mouse control for systems has attracted a lot of attention in recent years. Most visual hand motion detection systems can only be used in specific settings due to the effects of lighting and complicated backgrounds. Together with the quick advancement of computer vision and machine learning, the need for human-machine interaction is also rising. Despite the complicated history, the suggested methodology provides a straightforward but efficient way to perform rapid manual tracking. In addition to hand tracking and gesture detection, this system also eliminates motion blur. The detected gesture is transformed into particular functional inputs, such as clicking and mouse movement, to control other programs. The project's voice assistant feature is one of its additional features. The voice assistants can be instructed verbally or by text. Voice-based intelligent assistants need an activating word, sometimes referred to as a wake word, before the command may be given.

Keywords- Virtual Mouse, Hand Gestures, Media Pipe, Gesture Recognition, Voice Assistant, Machine Learning

INTRODUCTION

This project provides a method for cursor control that is entirely manual and requires no tools. The suggested system is designed to perform a variety of operations, including volume control, right- and left-clicking, double-clicking, drag-and-dropping, multiple item selection, etc. A webcam is the sole input device needed for the suggested setup. Python and OpenCV are the two programmes needed to use the suggested strategy. The output from the camera will be shown on the system's screen so that the user can make additional modifications. The modules NumPy, wx, and Python will be utilized to construct this system. mouse and math. Moreover, voice assistance is a part of this project. These days,
voice searches surpass text searches. The According to analysts, voice searches will make up 50% of all queries by 2024. It seems that voice assistants have improved in intelligence. The voice assistant is helpful for streamlining daily tasks like showing the time and date, performing Google searches, finding any location on a map, opening any software, etc. Both vocal and written commands can be given to the voice assistants. Before accepting a command, voice-based intelligent assistants require an activating phrase, often known as a wake word. Quantum is the wake word for this endeavour.

LITERATURE SURVEY

Munir Oudah et al., proposed a paper titled “Hand Gesture Recognition Based on Computer Vision” that was published in the year 23 July 2020 by “Journal of imaging” This paper describe Hand gestures are a type of nonverbal communication that can be employed in a variety of contexts, including medical applications, human-computer interface, robot control, and communication between deaf-mute individuals. Many different methodologies have been used in research papers based on hand gestures, including computer vision and instrumented sensor technology. It also tabulates the effectiveness of these techniques, concentrating on computer vision methods that address similarity and difference points, utilised dataset, detection range (distance), classification techniques, amount and types of movements, hand segmentation approach, and camera type. With a brief discussion of some potential applications. Hand gesture recognition addresses a fault in interaction systems. Controlling things by hand is more natural, easier, more flexible and cheaper, and there is no need to fix problems caused by hardware devices, since none is required.

Md. Zahirul Islam et al, proposed a paper titled “Static Hand Gesture Recognition using CNN” that was published in the year 24 April 2019 by “Research gate” This paper describe Computer is a part and parcel in our day to day life and used in various fields. The interaction of human and computer is accomplished by traditional input devices like mouse, keyboard etc. Hand gestures can be a useful medium of human-computer interaction and can make the interaction easier. Gestures vary in orientation and shape from person to person. So, non-linearity exists in this problem. Recent research has proved the supremacy of Convolutional Neural Network for image representation and classification. Data augmentation like re-scaling, zooming, shearing, rotation, width and height shifting was applied to the dataset.

Mayur V et al, proposed a paper titled “Human Computer Interaction using Hand Gesture Recognition” that was published in the year 4 April 2014 by “International Journal of Engineering Research & Technology “ This paper describe Operating computer in virtual environment is increasingly getting attention in recent time. In this paper method of operating computer in real time using hand gesture technique is described. Three main application performed to operate computer are mouse operation using hand gesture, controlling Media player & third is creating shortcuts using static hand gesture .In static gesture recognition each gesture is assigned to a specific application such as opening word file or opening control panel. For static gesture recognition principal component analysis.

Abhay Dekate et al, proposed a paper titled “Study of Voice Controlled Personal Assistant earDevice” that was published in the year 10 December 2016 by “International Journal of Computer Trends and Technology” this paper describe the fast moving technology we can do things which we never thought we could do before but, to achieve and accomplish these thoughts there is a need for a platform which can automate all our tasks with ease and comfort. Thus we need to develop a Personal Assistant having brilliant powers of deduction and the ability to interact with the surroundings just by one of the materialistic form of human interaction The Hardware device captures the audio
request through microphone and processes the request so that the device can respond to
the individual using in-built speaker module.

Lokesh Gagnani et al, proposed a paper titled “Gesture Controlled Mouse and Voice
Assistant” that was published in the year 10 Oct 2022 by “International Journal for
Research in Applied Science & Engineering Technology” this paper describe Human
Computer Interaction method

where cursor movement can be controlled via real- time camera by making use of human
hand postures recognition. This method is an alternative to current techniques, which
include manually pressing buttons or using physical computer mouse. The process then
compares current color schemes within the frames to a list of color combinations. In
addition to that authors have also developed a voice assistant to improve user productivity
by managing routine tasks of the user and by providing information.

Steven Raj N et al, proposed a paper titled “Implementing Hand Gesture Mouse Using
OpenCV” that was published in the year 6 June 2020 by “International Journal of
Engineering Research & Technology” this paper describe Computer mouse is an input
device that helps to point and to interact with whatever that is being pointed. There are so
many types of mouse in the current trend, there’s the mechanical mouse that consists of a
single rubber ball which can rotate in any direction and the movement of the pointer is
determined by the motion of that rubber ball. Optical Mouse consists of a led sensor to
detect the movement of the pointer. An activating word, also known as a wake word, is
required for voice-based intelligent assistants before the order may be given. For this
project the wake word is quantum

Viraj Shinde et al, proposed a paper titled “Hand Gesture Recognition System Using
Camera” that was published in the year January – 2014 by International Journal of
Engineering Research & Technology In this paper, we focus on using pointing behavior
for a natural interface, Hand gesture recognition based human-machine interface is being
developed vigorously in recent years. Due to the effect of lighting and complex
background, most visual hand gesture recognition systems work only under restricted
environment. To classify the dynamic hand gestures, we developed a simple and fast
motion history image based method. In recent years, the gesture control technique has
become a new developmental trend for many human-based electronics products. This
technique let people can control these products more naturally, intuitively and
conveniently. In this paper, a fast gesture recognition scheme is proposed to be an
interface for the human-machine interaction (HMI) of systems. recognition complexity
and be more suitable for controlling real-time computer system.

**SYSTEM ANALYSIS**

**EXISTING SYSTEM**

Muscle-Gesture Computer Interface requires hand gesture identification using surface
electromyography (sEMG). It typically focuses on how well it performs in terms of
accuracy and robustness. Unfortunately, there hasn't been much discussion of how reliable
these classifiers are. This might be because there isn't agreement on how reliable models
are defined in this field. They have shown that ECNN has a lot of potential for identifying
finger movements by establishing the model reliability measures and offering an offline
framework to evaluate it.

**3.1.2 EXISTING SYSTEM DISADVANTAGE**

Accuracy is low.

Electromyography Signals are used to prepare the dataset.

It is a complex process.
PROPOSED SYSTEM

For the purpose of automatically identifying hand gesture movements, we put forth a fresh and reliable deep learning model based on a convolutional neural network (CNN). If more feature extraction techniques are added to assist the CNN method, along with the effective detection of hand gesture movements, it is projected that the success of the results would grow. To recognise hand gestures and vocal instructions, we are employing computer vision and machine learning techniques, which function without the need for any additional hardware. Using pybind11 as its foundation, Media Pipe's implementation of models like CNN is utilised. It comes in two varieties: one that uses gloves of any uniform colour and Media Pipe Hand detection to work directly on hands, and the other that does so. We additionally use more feature in voice assistant.

PROPOSED SYSTEM ADVANTAGE

Accuracy may be improvised.
No signals are required to prepare the dataset.

HARDWARE REQUIREMENTS:
Processor : Intel i3 or later
Hard disk : minimum 10 GB
RAM : minimum 4 GB

SOFTWARE REQUIREMENTS

Operating System: Windows 10 or later
Simulation Tool : Anaconda with Jupyter Notebook
Language: Python
Pre-requisites: pyttsx3=2.7, Speech Recognition=3.8.1, pynput==1.8.1

LIST OF MODULES
1) Initialization of Camera
2) Mapping and Extraction of all Hand Gestures
3) Co-relating the Resolution of Camera Window and Screen.
4) Recognition and Execution of Hand Gestures.
5) Voice assistant

1) Initialization of Camera
It is crucial to match the resolution of the system screen with the resolution of the camera window being used because the camera is used through an image as the user moves their finger in the camera window. So, it becomes important to determine how many pixels are moved on the actual screen if the user moves by 1 pixel in the camera window. This is necessary so that the user's finger can completely cover the screen while being photographed by the camera. With the commands, the resolution of the screen can be estimated.
For the purpose of automatically identifying hand gesture movements, we put forth a fresh and reliable deep learning model based on a convolutional neural network (CNN). If more feature extraction techniques are added to assist the CNN method, along with the effective detection of hand gesture movements, it is projected that the success of the results would grow. To recognise hand gestures and vocal instructions, we are employing computer vision and machine learning techniques, which function without the need for any additional hardware. Using pybind11 as its foundation, Media Pipe’s implementation of models like CNN is utilised. It comes in two varieties: one that uses gloves of any uniform colour and Media Pipe Hand detection to work directly on hands, and the other that does so. We additionally use more feature in voice assistant.

### Proposed System Advantage

- Accuracy may be improvised.
- No signals are required to prepare the dataset.

### Hardware Requirements:

- Processor: Intel i3 or later
- Hard disk: minimum 10 GB
- RAM: minimum 4 GB

### Software Requirements:

- Operating System: Windows 10 or later
- Simulation Tool: Anaconda with Jupyter Notebook
- Language: Python
- Pre-requisites: pyttsx3=2.7, Speech Recognition=3.8.1, pynput==1.8.1

### List of Modules

1. Initialization of Camera
2. Mapping and Extraction of all Hand Gestures
3. Co-relating the Resolution of Camera Window and Screen.
4. Recognition and Execution of Hand Gestures.
5. Voice assistant

---

It is crucial to match the resolution of the system screen with the resolution of the camera window because the camera is used through an image as the user moves their finger in the camera window. So, it becomes important to determine how many pixels are moved on the actual screen if the user moves by 1 pixel in the camera window. This is necessary so that the user's finger can completely cover the screen while being photographed by the camera. With the commands, the resolution of the screen can be estimated.

---

Mapping & Extraction of Hand Gesture

The mapping is assigned binary integers because the cursor will move across the entire screen. A number of hand features that are utilised to recognise gestures are extracted using the Mediapipe programme. The subtraction method, which is simple and efficient for hand detection from the original image, is used. The fingers and palms can be divided using the following binary integers. The palm point, which is the centre of the palm, may be found using both the distance transform method and the landmark detection method of the Mediapipe module. The distance to the border pixel nearest to each pixel in the distance transform image is recorded.

---

2) Co-relating the Resolution of Camera Window and Screen.

The many motions that the acquisition toolkit will employ to record the user's finger as it moves. The camera must be able to record at least 15 frames per second.

---

3) Recognition and Execution of Hand Gestures.

The identification and execution of purposeful hand motions are made possible by the
successful mapping and extraction of distinct hand features. To do this, the Mediapipe and Open CV modules are both utilised. The quantity and variety of fingers present can be used to predict hand actions. What types of fingers are detected based on the content of the fingers? A gesture made with five fingers is classified as a neutral gesture, one made with two fingers as a move cursor gesture, one made with the right finger pulled down as a right click, etc.

**Figure: Recognition and execution**

4) **Voice assistant**

The initial phase in the process involved examining the audio commands that the user had delivered through the microphone. This might entail gathering data or controlling the computer's internal data and other thing.

**Figure: Vo Figure 7.**
2. WORKFLOW

SYSTEM ARCHITECTURE
RESULT AND OUTPUT

GESTURE RIGHT CLICK:
In this paper, we have discussed an alternative to the conventional physical mouse that provides mouse functions with the help of computer vision that houses a web camera that recognizes fingers and hand gestures and processes the captured frames and uses a machine learning algorithm to execute the defined mouse functions like moving the cursor, right-click, left click and scrolling function. After testing authors have concluded that the proposed virtual mouse system has worked exceedingly well and with great accuracy and the current system has overcome the drawbacks of the other systems. Through the Voice Assistant, authors have automated various services using a single-line command. It eases most of the tasks of the user like searching the web, file navigation, finding a location on google maps, opening any application, etc. The future plans include integrating quantum with mobile using React Native to provide a synchronized experience between the two connected devices. It is basically designed to minimize human efforts and control the device with just a human Voice.

Our project 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, future improvements to the technique currently in use will allow for the use of both hands for various gestures. Many applications have been substantially improved through the rapid development of hand gesture recognition systems.

REFERENCES

[1] Munir Oudah, Electrical Engineering Technical College, Middle Technical University, Baghdad 10022, Iraq; “Hand Gesture Recognition Based on Computer Vision” Article. 23 July 2020
[6] Lokesh Gagnani, Hetvi Patel, Sakshi Chaturvedi, Rajeshwari Jaiswal Computer Engineering Department, LDRP Institute of Technology and Research Gandhinagar Gujarat India; “Gesture Controlled Mouse and Voice Assistant” Article. 10 Oct 2022

