Gesture and Voice Controlled Virtual Mouse



**UNIVERSITY OF ENGINEERING**

**& MANAGEMENT, JAIPUR**

**Gesture and Voice Controlled Virtual Mouse**

Submitted in the partial fulfillment of the degree of

## BACHELOR OF TECHNOLOGY

In

## COMPUTER SCIENCE & ENGINEERING

Under

## UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

BY

## Rohit Pancholi, Subroto Saha, Priyal Jangid

**University Roll no: 12020002001154, 12020002001013, 12020002001097**

UNDER THE GUIDANCE OF

## Prof. Hriday Banerjee

COMPUTER SCIENCE & ENGINEERING



UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

**Approval Certificate**

This is to certify that the project report entitled “**Gesture and Voice Controlled Virtual Mouse**” submitted by **Rohit Pancholi** (Enrollment No.:**12020002001154**), **Subroto Saha** (Enrollment No.:**12020002001013**), **Priyal Jangid** (Enrollment No.:**12020002001097**) in partial fulfillment of the requirements of the degree of **Bachelor of Technology** in **Computer Science & Engineering** from University **of Engineering and Management, Jaipur** was carried out in a systematic and procedural manner to the best of our knowledge. It is a bona fide work of the candidate and was carried out under our supervision and guidance during the academic session of 2020-2024.

### Prof. Hriday Banerjee

Project Guide, Assistant Professor (CSE) UEM, JAIPUR

### Prof. Mrinal Kanti Sarkar Prof. A Mukherjee

HOD (CSE AIML) Dean

UEM, JAIPUR UEM, JAIPUR

**ACKNOWLEDGEMENT**

The endless thanks goes to Lord Almighty for all the blessings he has showered onto us, which has enabled us to write this last note in our research work. During the period of our research, as in the rest of our life, we have been blessed by Almighty with some extraordinary people who have spun a web of support around us. Words can never be enough in expressing how grateful we to those incredible people in our life who made this thesis possible. We would like an attempt to thank them for making time during our research in the Institute a period We will treasure. We are deeply indebted to our research supervisor, Professor Hriday Banerjee such an interesting thesis topic. Each meeting with him added in valuable aspects to the implementation and broadened us perspective. He has guided us with his invaluable suggestions, lightened up the way in my darkest times and encouraged us a lot in the academic life.

Rohit Pancholi Subroto Saha Priyal Jangid

**ABSTRACT**

By controlling cursor movement with a real-time camera and microphone, this project advances the Human Computer Interaction (HCI) paradigm in the field of computer science.

The hand movement and speech is the most effortless and primitive way of communication. It's a replacement for the present ways, which entail manually moving a physical computer mouse or pressing buttons. Instead, the system controls and performs numerous mouse activities using a camera for computer vision technology and a microphone for speech recognition and processing. It can perform all functions that a physical mouse can.

The Virtual Mouse continuously gathers real-time visuals and voice commands, which are then filtered and converted in a number of steps. When the procedure is completed, the programme uses image processing and natural language processing to extract the valid command needed to complete the task.

Specially abled people with hand problems can use this virtual mouse to control the computer's mouse functionalities.

[**Table of Contents**](#_bookmark0)

|  |  |  |
| --- | --- | --- |
| **[Table of Contents](#_bookmark0)** |  | VI |
| **List of figures** |  | **VIII** |
| [**1. CHAPTER**](#_bookmark1) |  | **1** |
| **Introduction** |  | **1** |
| **1.1 Context and Significance** | |  | **1** |
| **1.2 Gesture Recognition Technique** | |  | **2** |
| **1.2.1 Evolution of Gesture Control**  **1.2.2 Applications and Use Cases** | |  | **2**  **2** |
| **1.2.3 Challenges and Solutions** | |  | **2** |
| **1.3 Voice Recognition Techniques** | |  | **3** |
| **1.3.1 Advancement in Voice Recognition** | |  | **3** |
| **1.3.2 Challenges and Mitigations** |  | **3** |
| **1.4 Synergy of Gesture and Voice Interaction** |  | **4** |
| **1.5 Practical Application** |  | **4** |
| **2. CHAPTER** |  | **5** |
| **Background Study** |  | **5** |
| [**3. CHAPTER**](#_bookmark3) |  | [**6**](#_bookmark3) |
| **OBJECTIVE**  **3.1 Seamless Integration of Gesture and Voice Controls**  **3.2 Precision and Reliability**  **3.3 Accessibility and Inclusivity** |  | **6**  **6**  **6**  **7** |
| **4. CHAPTER** |  | **8** |
| **FLOW CHART** |  | **8** |
| **5. CHAPTER**  **METHODOLOGY** |  | **10** |
| **5.1 Computer Vision application for object identifiaction** |  | **11** |
| **5.2 Palm Detector Model** |  | **13** |
| **6. CHAPTER** |  | **15** |
| **APPRATUS** |  | **15** |
| **6.1 Hardware Requirement** |  | **15** |
| **6.2 Software Requirement** |  | **16** |
| **7. CHAPTER** |  | **17** |
| **Experimental Setup** | |  | **17** |
| **RESULTS** | |  | **18** |
| **CONCLUSION** | |  | **29** |
| **FUTURE SCOPE** | |  | **30** |
| **REFRENCES** |  | **31** |

# LIST OF FIGURES

### Figure 4.1: Architecture of Gesture and Virtual Mouse 8

**Figure 4.2: Working of Gesture Virtual Mouse 9**

### Figure 5.1: Gesture Control 10

**Figure 5.1.1: OpenCV work 11**

**Figure 5.1.2: Gray Scale Images 11**

**Figure 5.1.3: BGR Format 12**

**Figure 5.1.4: Palm Detector 12**

**Figure 5.2.1: Hand Landmark 14**

**Figure : Move cursor 18**

**Figure : Right Click 19**

**Figure : Left Click 20**

**Figure : Volume Control 21**

**Figure : Scroll 22**

**Figure : DRAG and Drop 23**

**Figure : Neutral Gesture 24**

**Figure : Double Click 25**

**Figure : Voice command search location 26**

**Figure : Google Search 26**

**Figure : Control Root Directory 27**

**Figure : Time 28**

# 1. CHAPTER

## INTRODUCTION

In the ever-evolving landscape of human-computer interaction, the amalgamation of gesture and voice recognition technologies stands out as a transformative and innovative approach. This project embarks on the ambitious journey of crafting a Gesture and Voice-Controlled Virtual Mouse, redefining how users engage with digital interfaces. The convergence of gesture and voice inputs represents a paradigm shift, offering a natural and intuitive means to interact with computers and devices.

The conventional mouse and keyboard interfaces, while foundational, may fall short in providing a truly immersive and inclusive interaction experience. The Gesture and Voice-Controlled Virtual Mouse project responds to this limitation by harnessing the capabilities of two distinct yet complementary modalities: gestures, the expressive language of body movements, and voice, the rich and nuanced medium of human communication.

**1.1 Context and Significance:**

In recent years, gesture and voice recognition technologies have matured, thanks to advancements in computer vision, machine learning, and natural language processing. These technologies have found applications in diverse domains, from gaming consoles to virtual assistants. However, the seamless integration of these modalities into a unified, responsive, and versatile virtual mouse system remains a relatively unexplored frontier.

The significance of this project lies in its potential to revolutionize how users interact with digital interfaces. By eliminating the reliance on physical peripherals and embracing a multimodal approach, the Gesture and Voice-Controlled Virtual Mouse aims to enhance user experience, accessibility, and productivity. This project is not merely about creating a novel input method but about paving the way for a more natural, adaptive, and personalized interaction between humans and computers.

**1.2: Gesture Recognition Technologies**

**1.2.1 Evolution of Gesture Recognition**

The evolution of gesture recognition has traced a captivating journey from the early stages of basic hand tracking to the sophisticated and nuanced systems we encounter today. Initially confined to simple, predefined hand movements, gesture recognition has undergone transformative advancements spurred by developments in computer vision and machine learning. Early systems primarily focused on recognizing distinct gestures for specific commands, laying the foundation for applications in gaming and interactive interfaces. Over time, the evolution has seen a shift toward more complex and natural interactions, enabling systems to discern subtle nuances in hand and body movements. This progression has been fueled by advancements in depth-sensing technologies, allowing for a more accurate interpretation of gestures in three-dimensional space. As a result, gesture recognition has transcended its initial limitations, finding applications in diverse fields such as healthcare, automotive interfaces, and virtual reality, offering users a more intuitive and immersive interaction experience.

**1.2.2: Applications and Use Cases**

Gesture recognition technology has found an array of compelling applications and use cases, reshaping how humans interact with digital interfaces. In the realm of consumer electronics, gesture recognition has become integral to gaming consoles, enabling users to control games and devices with intuitive hand and body movements. Moreover, in healthcare, it plays a pivotal role in touchless interfaces for medical equipment, reducing the risk of contamination. Automotive interfaces utilize gesture recognition for hands-free controls, enhancing driver safety and convenience. Retail and public spaces leverage this technology for interactive displays and customer engagement. As the technology continues to advance, its applications extend to augmented and virtual reality, allowing users to manipulate digital environments seamlessly. The versatility of gesture recognition underscores its transformative impact across various industries, promising a future where human-computer interaction is not bound by physical interfaces but is instead guided by natural, gesture-based communication.

**1.2.3: Challenges and Solutions**

Gesture recognition, while transformative, grapples with distinct challenges that researchers and developers tirelessly address to enhance its effectiveness. One prominent challenge lies in the diversity and complexity of gestures, making it difficult to create universally applicable recognition models. Ambiguity and variability in individual movements pose obstacles, especially in real-world, dynamic environments. Moreover, occlusion, where parts of the body or hand are hidden from the sensor's view, adds another layer of complexity. To counter these challenges, ongoing research focuses on developing robust machine learning algorithms capable of adapting to diverse gestures and environmental conditions. Advances in depth-sensing technologies, such as 3D cameras, provide solutions for improved spatial understanding, mitigating issues related to occlusion and enhancing the accuracy of gesture recognition systems. Additionally, collaborative efforts in creating standardized gesture libraries and frameworks contribute to a more cohesive and interoperable landscape, overcoming the challenges associated with the diverse nature of human gestures. Addressing these hurdles ensures that gesture recognition continues to evolve as a reliable and versatile technology in various applications.

**1.3: Voice Recognition Technologies**

**1.3.1: Advancements in Voice Recognition**

Advancements in voice recognition technology have propelled it from early stages of basic speech-to-text conversion to sophisticated systems that comprehend natural language with remarkable accuracy. Initial iterations struggled with accents, ambient noise, and contextual nuances, limiting their practicality. However, contemporary developments have been marked by the integration of advanced machine learning algorithms, neural networks, and natural language processing techniques. Modern voice recognition systems, such as those employed by virtual assistants, exhibit a deep understanding of spoken language, distinguishing between accents and dialects while adapting to individual speaking styles. Cloud-based architectures and expansive datasets have significantly improved the overall performance, enabling real-time processing and fostering seamless integration across devices. The evolution of voice recognition stands as a testament to its transformative impact, permeating various aspects of our daily lives, from voice-controlled smart devices to hands-free operation in automotive interfaces. As these advancements continue, voice recognition technology is poised to play an even more central role in reshaping human-computer interaction.

**1.3.2: Challenges and Mitigations**

Voice recognition technology, while making substantial strides, grapples with notable challenges that impact its widespread adoption. One primary concern is the inherent variability in human speech patterns, including accents, dialects, and individual nuances, which can lead to misinterpretation by recognition systems. Background noise in diverse environments poses another obstacle, affecting the accuracy of voice commands. Furthermore, privacy and security concerns regarding the storage and processing of voice data raise apprehensions among users. To address these challenges, ongoing research focuses on developing robust algorithms that adapt to diverse linguistic variations and environmental conditions. Advanced machine learning techniques, such as deep learning, contribute to improving the accuracy of voice recognition systems. Additionally, advancements in noise-canceling technologies and signal processing aim to enhance the performance of these systems in noisy surroundings. Stricter data privacy measures, including on-device processing and anonymization of stored data, serve as mitigations for the privacy concerns associated with voice recognition technology. As developers continue to refine these solutions, the journey toward seamless and secure voice interactions remains at the forefront of technological innovation

**1.4: Synergy of Gesture and Voice Interaction**

The synergy between gesture and voice interaction represents a compelling frontier in human-computer interaction, offering a multimodal approach that capitalizes on the strengths of both modalities. Combining the expressiveness of gestures with the richness of voice commands creates a more immersive and intuitive user experience. This synergy allows users to seamlessly switch between gestural inputs and vocal instructions, providing a versatile means of interaction tailored to different contexts and preferences. The complementarity of gesture and voice controls enhances the overall efficiency and naturalness of the interaction, enabling users to convey nuanced commands and navigate digital environments with unprecedented fluidity. As these modalities converge, the synergy between gesture and voice interaction not only broadens the scope of possibilities in virtual interfaces but also contributes to the development of more inclusive and accessible technologies that cater to diverse user needs and preferences.

**1.5: Practical Applications**

The Gesture and Voice-Controlled Virtual Mouse holds immense promise for revolutionizing user interfaces across various domains, with one particularly impactful practical application being in the field of accessibility. For individuals with motor impairments or physical disabilities that may limit their ability to use conventional input devices, this innovative technology serves as a gateway to a more inclusive computing experience. By interpreting a diverse range of gestures and voice commands, users can effortlessly navigate, click, and interact with digital interfaces without the need for physical manipulation. This application goes beyond convenience; it empowers users with disabilities, providing them with a more independent and seamless means to access and control digital devices, fostering increased participation in educational, professional, and recreational activities. The Gesture and Voice-Controlled Virtual Mouse, through its accessibility-focused application, not only redefines the parameters of user interaction but also contributes significantly to building a more equitable and inclusive digital landscape.

# CHAPTER

## BACKGROUND STUDY

We have come a long way in the field of human computer interaction. Gesture based mouse control was carried out by wearing gloves initially. Later, colour tips were also used for gesture recognition. Although such systems were not very accurate. The recognition accuracy is less due to use of gloves. Some users may not feel comfortable wearing gloves, and in some situations, recognition is not as accurate as it may be due to colour tip detection failure. Computer-based gesture detection systems have recently received some attention.

In 1990, Quam introduced a hardware-based approach that required the user to wear a DataGlove. Despite the fact that Quam's proposed method generates more precise results, certain of the gesture controls are difficult to execute with the system.

Zhengyou et al. proposed the Visual Panel interface system (2001). A quadrangle-shaped plane is used in this system, allowing the user to perform mouse operations with any tip-pointed interface instrument. Though the system can be operated contact free yet it does not solve the problem of surface area requirement and material handling.

Color tracking mouse stimulation was proposed by Kamran Niyazi et al. (2012). Using computer vision technology, the system tracks two colour tapes on the user's fingertips. One of the tapes will be used to control the cursor's movement, while the other will act as a trigger for the mouse's click events. Despite the fact that the proposed system handled the bulk of the issues, it only has a limited range of capabilities, as it can only perform fundamental actions such as cursor movements, left/right clicks, and double clicks.

To replicate click events, the system requires three fingers with three colour pointers, according to Kazim Sekeroglu (2010). The suggested system can detect pointers using colour information, track their motion, change the cursor according to the position of the pointer, and simulate single and double left or right mouse click events.

Chu-Feng Lien (2015) proposed a way for controlling the mouse cursor and click events using only one's fingertip. To interact with the system, the suggested system does not require hand motions or colour tracking; instead, it uses a feature called Motion History Images (MHI). Because the frame rates can't keep up with quickly moving objects, the proposed system can't detect them. Furthermore, because mouse click events occur when the finger is held in particular positions, this may cause the user to move their fingers constantly to avoid false alarms, which can be inconvenient.

S. Shriram(2021); the model employs the MediaPipe package for tracking the hands and the tips of the hands, as well as the Pynput, Autopy, and PyAutoGUI packages for moving around the computer's window screen and performing actions like left click, right click, and scrolling.

# CHAPTER

## OBJECTIVE

The primary objective of the Gesture and Voice-Controlled Virtual Mouse project is to design, develop, and implement an intuitive and efficient human-computer interaction system that leverages the synergies of gesture and voice recognition technologies. The project aims to achieve the following specific goals:

**3.1 Seamless Integration of Gesture and Voice Controls**:

The seamless integration of gesture and voice controls marks a significant advancement in human-computer interaction, fostering a more intuitive and natural engagement with digital systems. By harmonizing the expressive power of gestures and the contextual richness of voice commands, this integration eliminates traditional constraints on user inputs. Users can effortlessly transition between gestural movements and spoken instructions, creating a fluid and adaptable interaction experience. The synergy between gestures and voice enables a dynamic and responsive virtual interface, where users can navigate, command, and interact with digital environments with remarkable ease. This seamless integration not only enhances the efficiency of user inputs but also contributes to a more inclusive and user-friendly computing paradigm, catering to a diverse range of preferences and accessibility needs. As technology advances, the seamless fusion of gesture and voice controls promises to redefine the way we interact with and control The seamless integration of gesture and voice controls marks a significant advancement in human-computer interaction, fostering a more intuitive and natural engagement with digital systems. By harmonizing the expressive power of gestures and the contextual richness of voice commands, this integration eliminates traditional constraints on user inputs. Users can effortlessly transition between gestural movements and spoken instructions, creating a fluid and adaptable interaction experience. The synergy between gestures and voice enables a dynamic and responsive virtual interface, where users can navigate, command, and interact with digital environments with remarkable ease. This seamless integration not only enhances the efficiency of user inputs but also contributes to a more inclusive and user-friendly computing paradigm, catering to a diverse range of preferences and accessibility needs. As technology advances, the seamless fusion of gesture and voice controls promises to redefine the way we interact with and control digital devices digital devices

**3.2 Precision and Reliability**:

The paramount objectives of the Gesture and Voice-Controlled Virtual Mouse project are precision and reliability in user interaction. Precision is pivotal to accurately interpret and execute user gestures and voice commands, ensuring the virtual mouse res The paramount objectives of the Gesture and Voice-Controlled Virtual Mouse project are precision and reliability in user interaction. Precision is pivotal to accurately interpret and execute user gestures and voice commands, ensuring the virtual mouse responds seamlessly to the user's intent. By employing advanced algorithms and machine learning techniques, the project aims to minimize errors and false positives, enhancing the overall precision of gesture and voice recognition. Reliability is equally crucial, emphasizing consistent and dependable performance across various environmental conditions and user contexts. The system's reliability will be rigorously tested to ensure that it consistently recognizes and executes commands, fostering a robust and trustworthy virtual mouse experience. In prioritizing precision and reliability, the project aims to establish a new standard for user interaction, where the Gesture and Voice-Controlled Virtual Mouse becomes a dependable and precise tool for navigating and controlling digital interfaces. ponds seamlessly to the user's intent. By employing advanced algorithms and machine learning techniques, the project aims to minimize errors and false positives, enhancing the overall precision of gesture and voice recognition. Reliability is equally crucial, emphasizing consistent and dependable performance across various environmental conditions and user contexts. The system's reliability will be rigorously tested to ensure that it consistently recognizes and executes commands, fostering a robust and trustworthy virtual mouse experience. In prioritizing precision and reliability, the project aims to establish a new standard for user interaction, where the Gesture and Voice-Controlled Virtual Mouse becomes a dependable and precise tool for navigating and controlling digital interfaces.

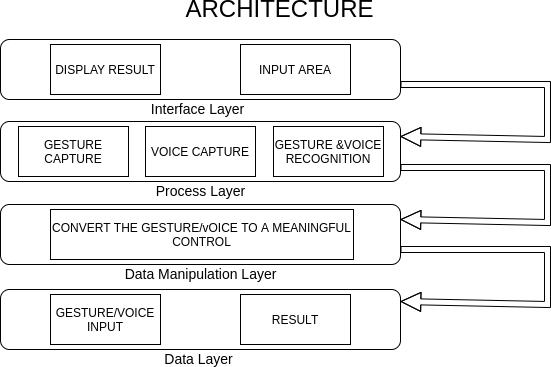
**3.3 Accessibility and Inclusivity**:

The core focus of the Gesture and Voice-Controlled Virtual Mouse project is to enhance accessibility and inclusivity in human-computer interaction. By integrating gesture and voice controls, the project aims to create an alternative input method that caters to users with diverse abilities and preferences. Accessibility features will be embedded to accommodate individuals with motor impairments, providing an inclusive solution that promotes independence and engagement with digital interfaces. The customization options for gestures and voice commands further contribute to inclusivity, allowing users to tailor the virtual mouse experience according to their specific needs. Through a user-centered design approach, the project aspires to break down barriers, making technology more accessible to a broad spectrum of users and fostering a more inclusive digital environment.

# 4. CHAPTER

## FLOWCHART

## Working of Gesture Virtual Mouse:

****

**4.1 Figure: Working of Gesture and Virtual Mouse**

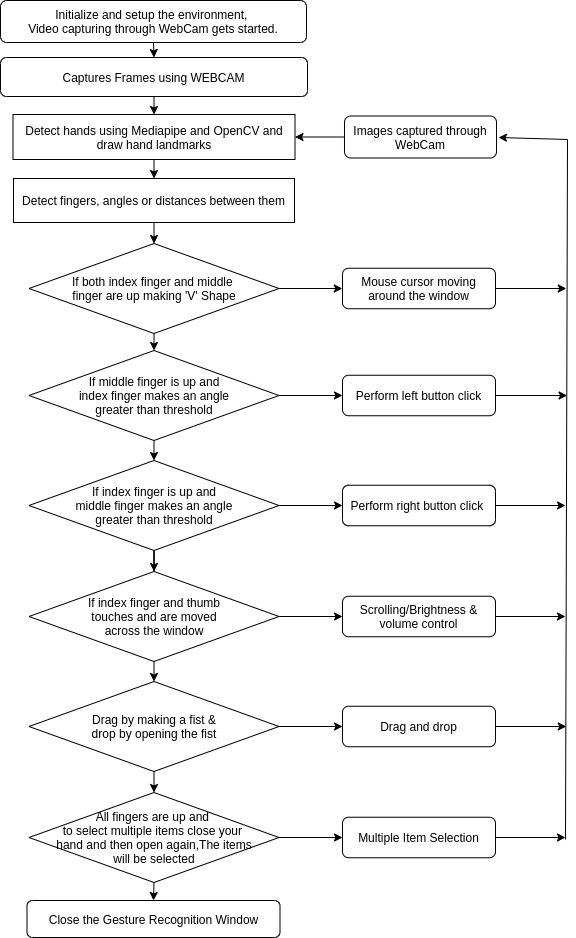
## Flowchat of Voice virtual Mouse:

## 

* 1. **Figure: Working of Gesture Virtual Mouse**

## CHAPTER

# Methodology

** Gesture Control**

**5.1 Figure: Gesture Control**

5.1 **Computer Vision application for object identification**

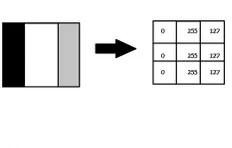
The frames captured by the webcam on a laptop or PC are used to create our virtual mouse environment. To capture the video object that is being formed, we used OpenCv, a Python computer vision package, and the web camera was used to begin capturing video. The web camera captures the frames and sends them to the virtual environment.

**Working of OpenCV**

Computer works only with numbers. Everything we save in a computer like video, images, documents, etc is saved in a computer in the form of numbers.

In image processing pixels are converted into numbers. A pixel is the smallest unit of a digital image.

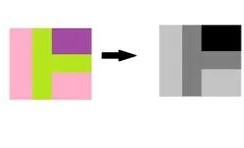
The numbers can be used to calculate a number's intensity at any given pixel. OpenCV can work in a grey scale or BGR(Blue, Green, Red) format.



**5.1.1 Figure: OpenCV work**

**Gray Scale Images:**

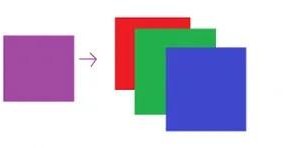
Image processing using this method involves converting the image into black and white format, where black is 0 and white is 255.



**5.1.2 Figure: Gray Scale Images**

**BGR format**

The photos feature three colours: blue, green, and red. The computer extracts that value from each pixel and puts the results in an array to be interpreted. Images are represented as three channels blue, green and red.



**5.1.3 Figure: BGR Format**

**ML Pipeline(Mediapipe) for Hand Tracking and Gesture Recognition**

Mediapipe is a Machine Learning system built on the collaboration of pipeline models.

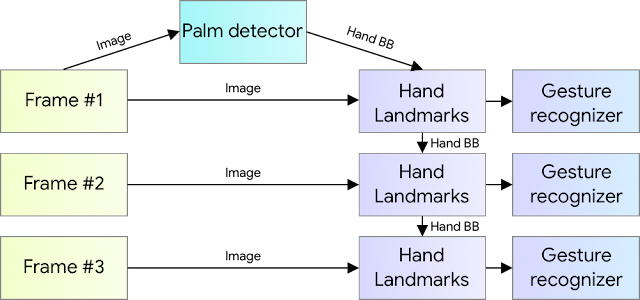
**What is ML Pipeline?**

A pipeline joins several stages together so that the output of one is used as the input for the next. Pipeline makes it simple to train and test using the same preprocessing.

**The hand tracking method makes use of a machine learning pipeline that consists of two models that work together:**

A palm detector that uses an aligned hand bounding box to locate palms on a whole input image.

A hand landmark model that uses the palm detector's clipped hand bounding box to produce high-fidelity results. landmarks in 2.5D



**5.1.4 Figure: Palm Detector**

**5.2 Palm Detector Model**

Hand detection is a tedious process as it requires identifying hands of various sizes, shapes, with deformities, etc. It is more complex than face detection as the contrast in features is far less than that in face.

We use palm detection model first as detecting palm or a fist is much easier than detecting a full hand with articulated fingers. Also palms are smaller therefore non suppression algorithm works better for it.

**Hand Landmark Model**

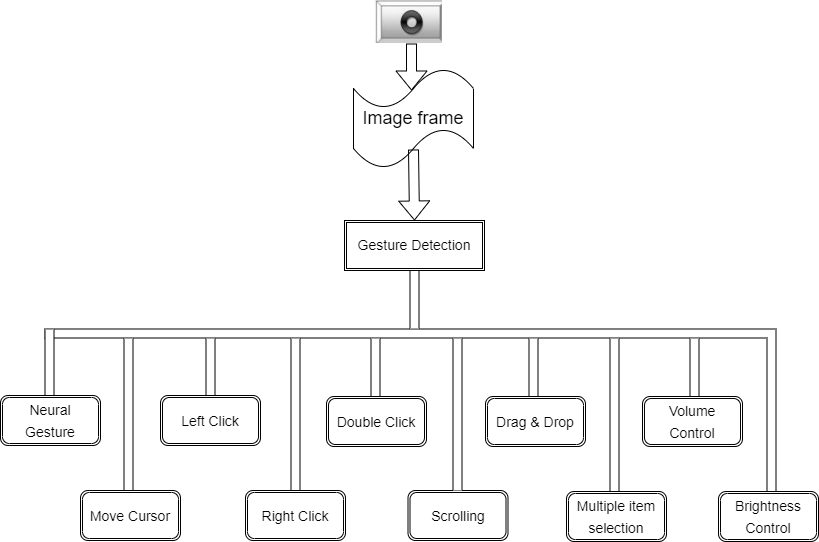
After detecting the palm using a palm detection model next a hand landmark model is used to detect 21 landmark points in 2.5 dimension. The Z depth is analysed using a Image Depth Map. The model recognises both partially and fully acclusioned hands perfectly.

The model has three outputs:

1. hand landmarks consisting of x, y, and relative depth.
2. A hand flag indicates the existence of a hand in the input image.
3. A binary classification of handedness, e.g. left or right hand.

The topology is the same as for the 21 landmarks. To avoid performing hand detection over and over for the entire frame, the probability of hand presence in a bounded crop is determined. The detector is triggered to reset tracking if the score falls below a threshold. We constructed a binary classification head to predict whether the input hand is left or right. Only the first frame or when the hand prediction shows that the hand is lost is the detector used.

For our project, the fingers are given Ids from 0 to 4.



**5.2.1 Figure: Hand Landmark**

# CHAPTER

## APPRATUS

**Hardware Requirement**

The Virtual Mouse application requires the following hardware for development and execution:

**Laptop or Computer Desktop**

To display what the webcam has taken, the virtual software will be started on the laptop or computer desktop.

The system will make use of (minimum requirements) Core2Duo processor (2nd generation)

2 GB RAM (Main Memory) 320 GB hard drive

14-inch LCD monitor

**Webcam**

The image is acquired with a camera that will continue to take photos

endlessly so that the application may process the image and calculate pixel position.

Resolution: 1.3 megapixels is the minimum required.

**Microphone**

Voice commands are recorded via the microphone. Until it is switched off or placed to sleep, it will continue to listen to all commands.

Microphone should be capable of recognizing frequency range of 40Hz - 16KH

**Software Requirement**

The following software is required for the development and execution of the Virtual Mouse application:

**Python Language**

The Virtual Mouse application is coded in Python with the help of Microsoft Visual Studio Code, an integrated development environment (IDE) for programming computer applications.

Basic arithmetic, bit manipulation, indirection, comparisons, logical operations, and more are all available in the Python library.

**Open CV Library**

This software is also created with the help of OpenCV.

OpenCV (Open Source Computer Vision) is a real-time computer vision library. OpenCV is capable of reading picture pixel values as well as real-time eye tracking and blink detection.

Software will be using:

OS : Window 10 Ultimate 64-bit Language : Python

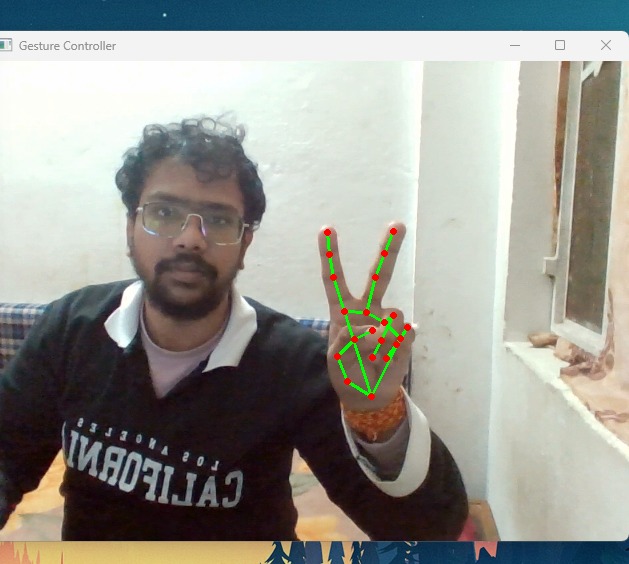
Tool Used : OpenCV and MediaPipe

# CHAPTER

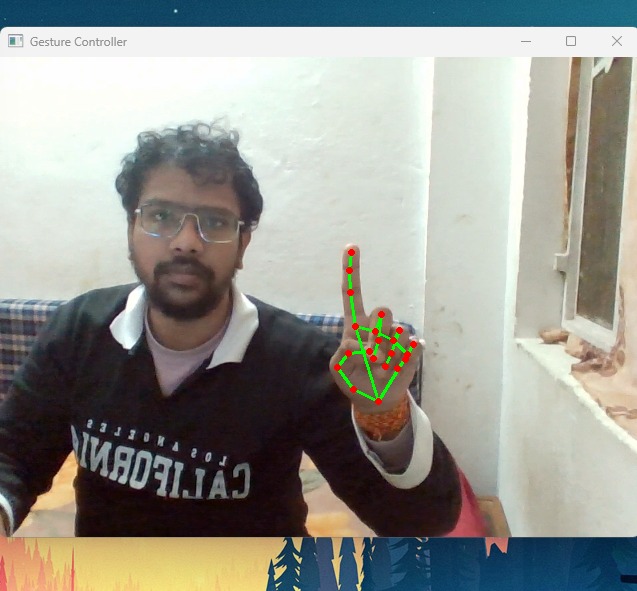
## EXPERIMENTAL SETUP

* **Camera Placement:** Position the camera at an optimal angle and distance to capture the user's hand gestures effectively.
* **Calibration Process:** Implement a calibration step to align the camera view with the virtual space where gestures will be recognized.
* **Python Script Execution:** Develop a main Python script that captures video frames from the camera, processes them using OpenCV for hand tracking, and feeds the data to the gesture recognition algorithm.
* **Gesture Mapping:** Define a set of gestures and map them to specific mouse actions (e.g., move left, move right, click, scroll) within the Python script.
* **User Interface:** Display a simple user interface using a graphical library like Tkinter to showcase the virtual mouse cursor and any relevant information.
* **Feedback Mechanism:** Integrate feedback mechanisms, such as visual indicators or sounds, to notify users when a gesture is recognized and acted upon.
* **Testing and Refinement:** Conduct extensive testing with users, gather feedback, and refine the Python scripts to enhance accuracy and responsiveness.
* **Power Supply and Connectivity:** Ensure the hardware components are adequately powered and establish reliable connections between the Raspberry Pi, camera, and other sensors.

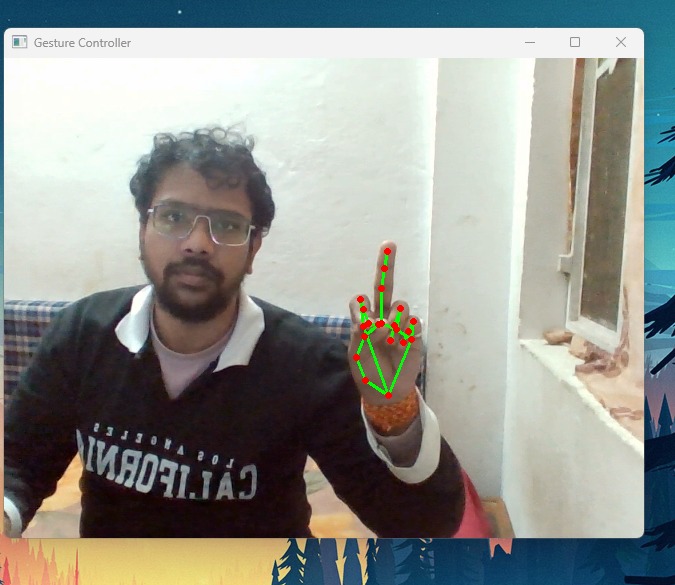
# RESULT



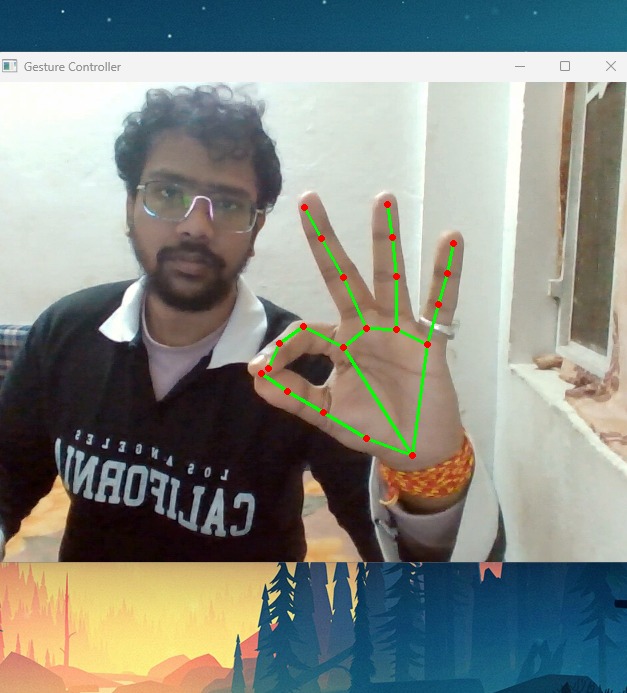
**Figure: Move cursor**



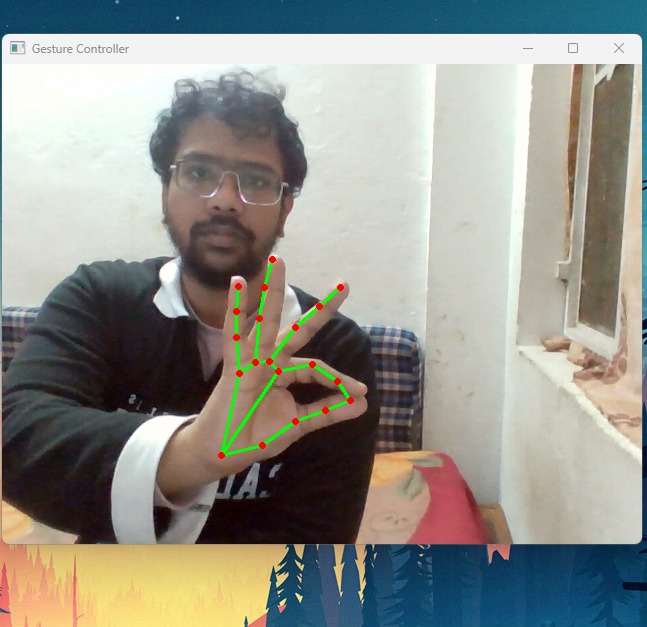
**Figure:** **Right Click**



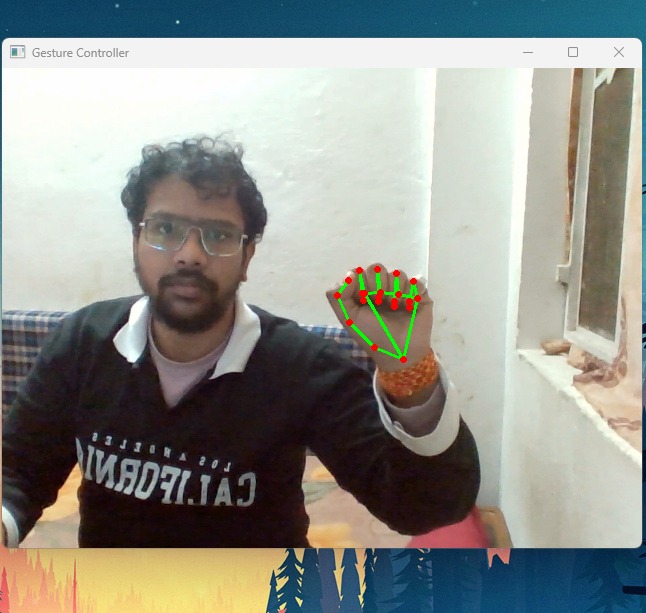
**Figure: Left Click**



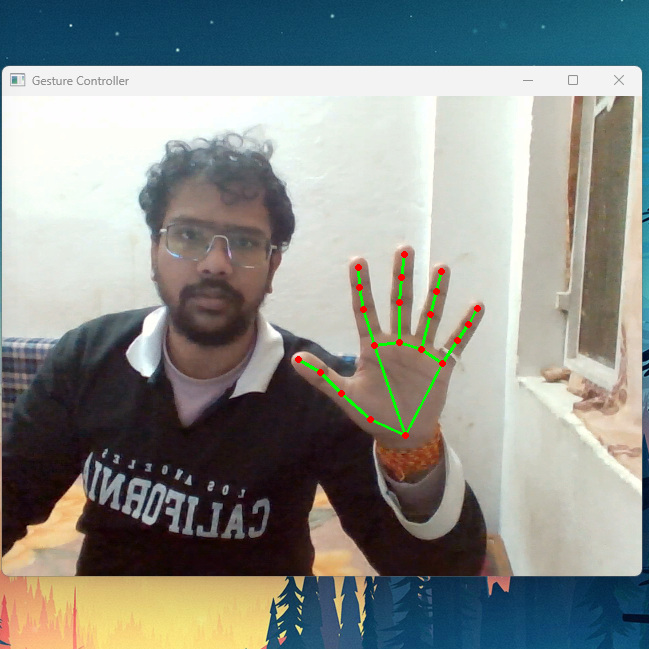
**Figure: Volume Control**



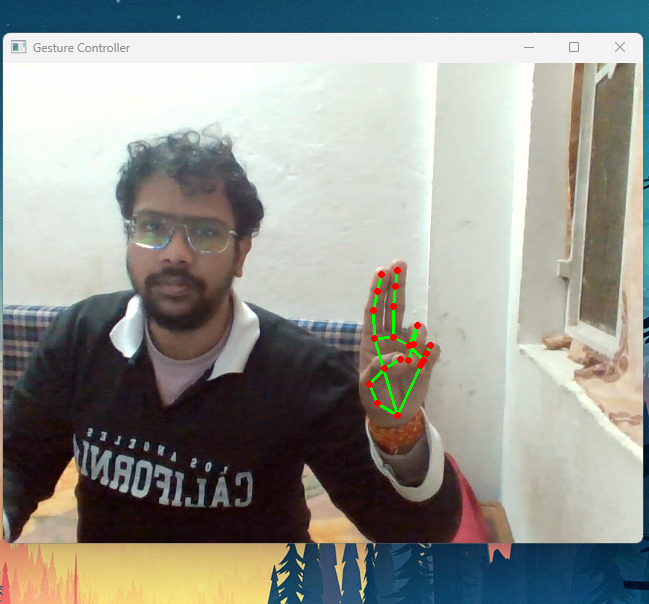
**Figure: Scroll**



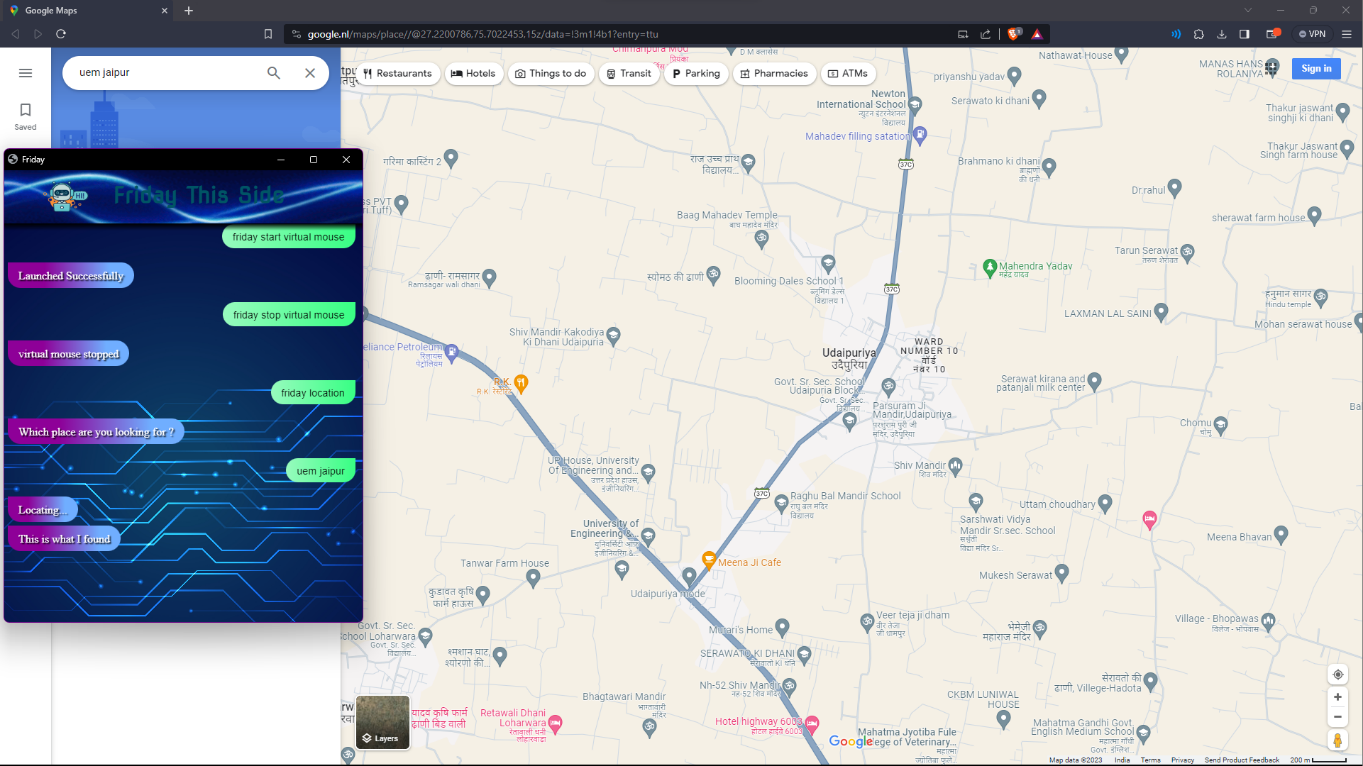
**Figure: Drag and Drop**

****

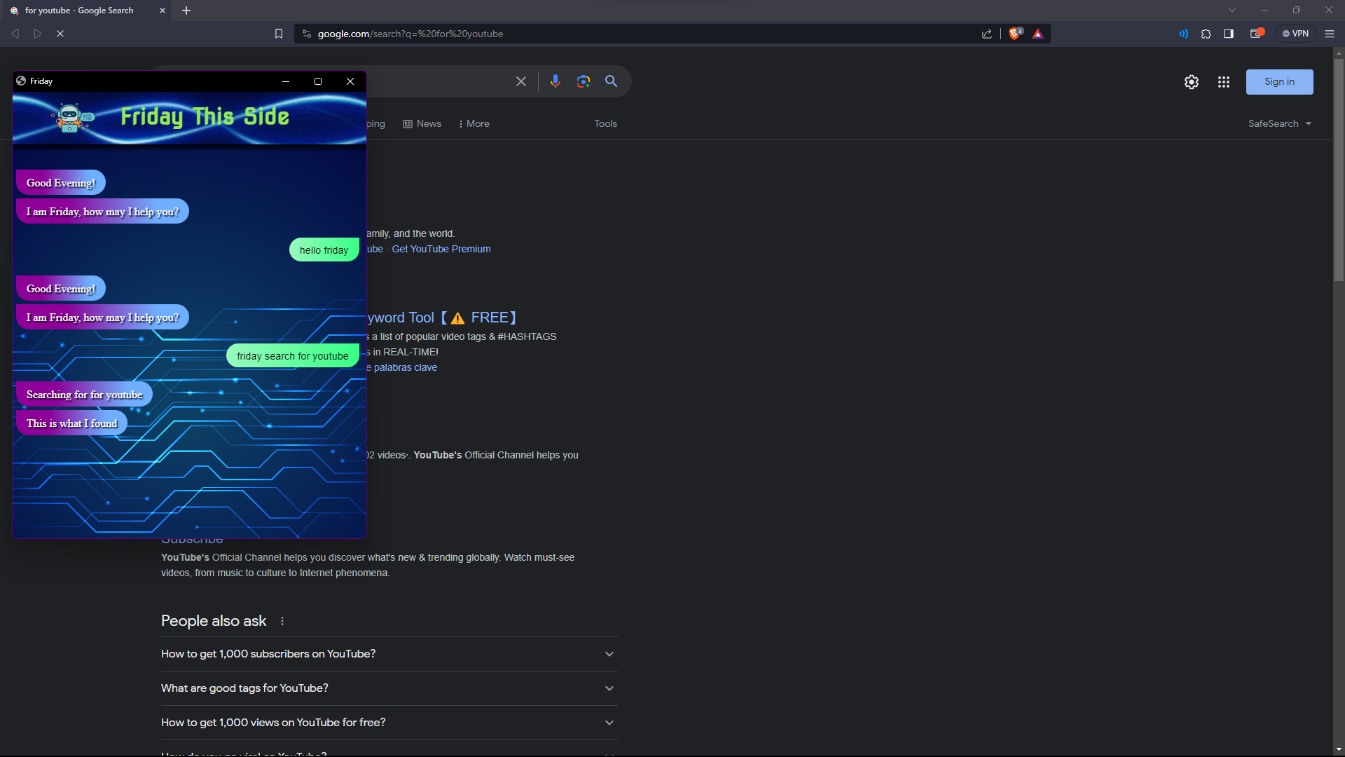
**Figure: Neutral Gesture**

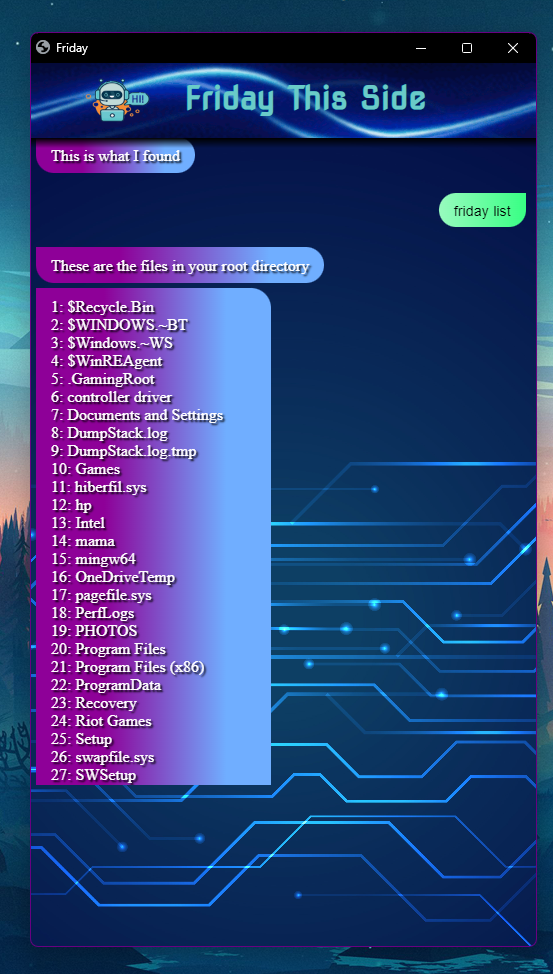
****

**Figure: Double Click**

****

**Figure: Voice command search location**

****

**Figure: Google Search**

**Figure: Control Root Directory**

****

**Figure: Time**

# CONCLUSION

The basic goal of the virtual mouse system is to control the mouse cursor and complete activities without needing a physical mouse by using hand gestures and voice commands. This proposed system is created by using a webcam (or any built-in camera) that recognises hand gestures and hand tip movement and processes these frames to perform the relevant mouse actions using the notion of speech recognition to quickly follow voice commands and perform mouse activities.

The model upon rigorous testing has come out to be highly accurate and sophisticated showing enormous improvements with respect to prior existing models. Since the proposed model has been tested for high sophistication, the virtual mouse can be used for real-time applications. Because the proposed mouse system may be operated digitally utilising hand gestures and voice commands rather than the traditional physical mouse.

It functions as a useful user interface and contains all mouse features. Research on advanced mathematical materials for image processing and investigating different hardware solutions has made possible more accurate hand detections. Not only this project shows the different gesture operations and voice commands that can be done by the users but it can also demonstrate the potential in simplifying user interactions with personal computers and hardware systems. Yet a major extension to this work could be to be able to work at a more complex background and compatible with different light conditions.

**Future Scope Of Gesture Voice Control Virtual Mouse**

The future scope of Gesture and Voice-Controlled Virtual Mouse holds tremendous potential as advancements in technology continue to redefine human-computer interaction. As artificial intelligence and machine learning algorithms evolve, the precision and versatility of gesture and voice recognition are expected to further enhance, offering users even more natural and seamless interactions with digital interfaces. The integration of these modalities may extend beyond traditional computing devices, finding applications in augmented and virtual reality environments. Moreover, the potential for incorporating contextual understanding and emotional cues into the interaction dynamics opens avenues for more personalized and intuitive user experiences. The future may witness the widespread adoption of Gesture and Voice-Controlled Virtual Mouse in diverse sectors, from healthcare and education to gaming and beyond, ultimately shaping a more immersive and accessible digital landscape. As technology progresses, the project's innovative approach to multimodal interaction is poised to play a pivotal role in influencing the trajectory of user interface design and human-computer interaction paradigms.

Virtual Mouse will be introduced soon to replace the conventional computer mouse, making it easier for users to connect with and administer their computers. In order to correctly track the user's gesture, the software must be fast enough to capture and process every image and speech command.

Other features and improvements could be added to make the application more user-friendly, accurate, and adaptable in different contexts. The following are the enhancements and functionalities that are required:

**Smart Recognition Algorithm**

Using the palm and numerous fingers, additional functions such as enlarging and reducing the window, and so on, can be implemented.

**Better Performance**

The response time is largely influenced by the machine's hardware, which includes the processor's processing speed, the amount of RAM available, and the webcam's characteristics. As a result, when the software is performed on a respectable machine with a webcam that operates well in various lighting conditions and a better quality microphone that can detect voice instructions correctly and rapidly, the programme may perform better.

.

# REFERENCE

Google: [www.google.com](http://www.google.com/) Youtube: [www.youtube.com](http://www.youtube.com/) Wikipedia: [www.wikipedia.com](http://www.wikipedia.com/)