Human Gesture Recognition using Computer Vision: Bridging the Gap Between Man and Machine.

A project report submitted
in partial fulfillment of requirement for the award of degree

BACHELOR OF TECHNOLOGY
in
ELECTRONICS & COMMUNICATION ENGINEERING
by

G. Pavan Kumar	**2105A41138**
O. Ranadheer	**2105A41108**
K. Hebsiba	**2105A41141**
R.Rishith Kumar	**2105A41147**

Under the guidance of
Dr. K. Sreedhar
Assoc. Prof, Department of ECE.

Department of Electronics and Communication Engineering

SR UNIVERSITY

Ananthasagar, Warangal.

SR UNIVERSITY

CERTIFICATE

This is to certify that this project entitled **“Human Gesture Recognition using Computer Vision: Bridging the Gap Between Man and Machine”** is the bonafied work carried out by **G. Pavan Kumar, O. Ranadheer, K. Hepsiba, P. Adithya** as a Capstone project for the partial fulfillment to award the degree **BACHELOR OF TECHNOLOGY** in **ELECTRONICS & COMMUNICATION ENGINEERING** during the academic year 2024-2025 under our guidance and Supervision.

Dr. K. Sreedhar
Assoc. Prof. (ECE),
SR University,
Ananthasagar, Warangal.

Dr. Sandip Bhattacharya
Prof. & HOD (ECE),
SR University,
Ananthasagar, Warangal.

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved Vice Chancellor, **Prof. DEEPAK GARG,** for his continuous support and guidance to complete this project in the institute.

We would like to extend our gratitude to **Dr. LMI Leo Joseph**, **Assoc. Professor,** the Dean of the School of Engineering, and **Dr. Sandip Bhattacharya**, **Professor** and Head of the ECE Department, for their encouragement and support.

We owe an enormous debt of gratitude to our project guide **Dr. K. Sreedhar, Assoc. Prof.** as well as project coordinators **Dr. LMI Leo Joseph, Assoc. Prof., Dr. J. Ajayan, Prof.**, for guiding us from the beginning through the end of the Major project with their intellectual advices and insightful suggestions. We truly value their consistent feedback on our progress, which was always constructive, and encouraging and ultimately drove us in the right direction.

Finally, we express our thanks to all the teaching and non-teaching staff of the department for their suggestions and timely support.

ABSTRACT

A unique Hand Gesture Recognition system improves human-computer interaction by tackling common problems including high computing load, lighting sensitivity, and real-time accuracy. This system uses an innovative computer vision approach to classify hand gestures filmed using a conventional camera. The method enhances recognition accuracy over a wide range of lighting situations and backdrops by training on many datasets. Streamlined algorithms allow for faster processing, resulting in real-time performance. The system adjusts well to various hand shapes and settings, offering a dependable solution for gesture-based interaction. Future studies will concentrate on improving system robustness and validating its use in various real-world circumstances.

CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1.1 OVERVIEW

The **Real-Time AI Virtual Mouse System Using Computer Vision** is an innovative project that utilizes hand gestures to control a computer mouse, removing the need for a traditional physical mouse. This system combines computer vision, machine learning, and automation techniques to create a touch-free interface, where users can move the cursor, click, and scroll using specific hand gestures. Developed with Python, the project leverages **OpenCV** for video capture and image processing, **MediaPipe** for accurate hand detection and tracking, and **PyAutoGUI** or **Autopy** to simulate mouse actions on the computer.

Objective and Motivation

The project's main goal is to provide a flexible, accessible, and hygienic alternative to traditional mouse interfaces. In addition to offering new options for everyday computer users, this technology is especially valuable for individuals with limited mobility, enabling them to control computers without the need for a physical device. The virtual mouse is also highly applicable in **public spaces**, where touchless control is preferable for health reasons, as well as in AR/VR environments, where traditional input methods may be less practical.

Technology and Tools

- **Python**: Python serves as the primary programming language for its ease of use and rich ecosystem of libraries.
- **OpenCV**: This open-source library is used for real-time video capture, frame processing, and basic image-processing tasks, allowing the system to analyze video feed from the webcam.
- **MediaPipe**: MediaPipe by Google provides a pre-trained model for accurate hand and finger landmark detection, tracking 21 landmarks on each hand in real time.
- **PyAutoGUI / Autopy**: These libraries allow the program to simulate mouse actions, such as cursor movement, clicks, and scrolling, enabling seamless integration between gesture recognition and actual mouse functions.

System Workflow and Core Functionality

1. **Video Capture and Frame Processing**: The system captures live video from the user's webcam, breaking it down into individual frames for processing.
2. **Hand Detection and Landmark Tracking**: MediaPipe's hand-tracking model identifies hands in each frame, detecting 21 key landmarks on each hand to allow for detailed

gesture recognition. These landmarks, including fingertips, knuckles, and the center of the palm, are used to identify hand shapes and positions.

3. **Gesture Recognition**: Using the positions of these landmarks, the system recognizes gestures. For example:
   - **Cursor Movement**: The position of the index fingertip can be mapped to screen coordinates, allowing users to move the cursor by moving their index finger.
   - **Clicking Gestures**: A pinching motion between the index finger and thumb can simulate a left-click, while other specific gestures, like extending multiple fingers, may signal right-clicking or scrolling actions.
4. **Mapping to Screen Coordinates**: Hand movements are mapped to the screen's dimensions, ensuring smooth and intuitive cursor control. This involves scaling the hand's movements to match the screen resolution, so small hand motions translate correctly to on-screen actions.
5. **Execution of Mouse Actions**: Using PyAutoGUI or Autopy, the recognized gestures are translated into actual mouse events on the computer. The system runs continuously, processing gestures in real time to provide a responsive and immersive user experience.

Implementation Steps

The project's implementation involves setting up the required libraries and defining each step:

1. **Initialize Libraries**: Import OpenCV, MediaPipe, PyAutoGUI/Autopy, and configure the camera and hand-tracking model.
2. **Capture Video and Detect Hands**: In each video frame, detect the hand and identify landmark positions using MediaPipe.
3. **Define Gesture Functions**: Define specific gestures based on landmark positions. For instance, calculate the distance between the thumb and index finger to detect a click gesture.
4. **Map and Execute Mouse Actions**: Map gestures to screen coordinates, execute mouse movements, and trigger click actions using PyAutoGUI.
5. **Run the System in Real Time**: Continuously capture frames, process them for gesture recognition, and simulate mouse actions, creating a seamless experience for the user.

Applications and Benefits

This AI Virtual Mouse has wide-ranging applications:

- **Accessibility**: Enables hands-free computer interaction for users with mobility challenges, providing greater independence.
- **Touchless Public Interfaces**: Useful in public terminals, kiosks, and ATMs, where reducing physical contact is crucial for hygiene, especially in high-traffic environments.
- **AR/VR Interaction**: In virtual and augmented reality, gesture-based control can replace traditional input devices, enhancing immersion and usability.

- **Creative Tools**: Artists, designers, and engineers can use the virtual mouse for hands-free control in digital workspaces, potentially transforming creative processes.

Challenges and Limitations

Despite its advantages, the AI Virtual Mouse system faces some challenges:

- **Real-Time Performance**: Processing frames in real time requires efficient coding and hardware. If the system is too slow, the delay will reduce the user experience, making gestures feel unresponsive.
- **Lighting and Background Dependency**: For optimal accuracy, the system needs good lighting and a simple background. Poor lighting or cluttered backgrounds can interfere with hand detection, affecting reliability.
- **Gesture Sensitivity and Customization**: Defining intuitive gestures is key to preventing accidental commands. Customizing gestures for user preferences may also require additional development and user testing to enhance usability.

Potential Future Enhancements

To further improve the AI Virtual Mouse, future updates could include:

- **Expanded Gesture Library**: Adding more gestures to handle complex commands and creating a more versatile system.
- **Voice Command Integration**: Integrating voice commands to allow for multimodal input, making the system even more flexible.
- **Better Accuracy with Advanced Models**: Using more advanced deep learning models could enhance hand-tracking accuracy and reduce misinterpretation.
- **Cross-Platform Support**: Ensuring compatibility across different operating systems for wider accessibility and usability.

2. LITERATURE SURVEY

2.1 EXISTING METHODS

In recent years, computer vision and deep learning methods are used for hand gesture recognition system. There are different techniques that involve in recognition of hand gesture. Here are some existing methods:

In [1] the study of development of a Smart Mouse system that leverages computer vision and gesture recognition to control computer functions without a traditional physical mouse. This approach addresses the limitations of standard input devices, particularly in enhancing accessibility and usability in various environments. The Smart Mouse utilizes image processing techniques to detect and interpret hand gestures, translating these movements into corresponding cursor actions on the screen. The authors highlight the system's application in reducing physical contact with devices, making it suitable for public use and touchless control scenarios, and describe the technology stack, including machine learning algorithms and computer vision tools, that enables accurate, real-time hand-tracking. Their findings emphasize the potential of such AI-driven systems to improve human-computer interaction, particularly in contexts where hygiene or accessibility is a priority. This research adds valuable insights into the evolving field of gesture-based interfaces, demonstrating the practicality and potential impact of gesture-controlled smart devices.

In [2] their study, present a Virtual Mouse Using Hand Gestures Recognition system that leverages computer vision and gesture recognition to enable touchless control of a computer interface. Their research outlines the development of a virtual mouse that recognizes hand gestures to perform standard mouse functions, such as cursor movement, clicks, and scrolling, offering a practical alternative to traditional input devices. Utilizing a combination of image processing techniques and real-time hand-tracking algorithms, the authors emphasize the system's ability to detect and interpret finger positions and hand gestures with precision. This virtual mouse is proposed as a more accessible and hygienic input method, with potential applications in fields like assistive technology and public interfaces. The study contributes to the body of work on gesture-based interfaces by demonstrating how AI and computer vision can transform human-computer interaction, enhancing user experience and addressing physical device limitations.

In [3] their study, which is presented at the International Conference on Communication and Signal Processing explores a real-time system for controlling brightness, contrast, and volume using hand gestures, developed with OpenCV and Python. This research focuses on enhancing human-computer interaction by eliminating the need for physical controls, providing a touch-free approach that is especially beneficial in settings where direct contact is impractical or undesirable. The system employs computer vision techniques to capture and analyze hand movements, allowing users to adjust screen brightness, contrast, and volume through predefined gestures. This application showcases the effectiveness of gesture recognition in real-time processing environments, demonstrating the viability of such systems for intuitive, hands-free control of multimedia settings. The study contributes to ongoing advancements in gesture-based control systems, highlighting how AI and computer vision can enhance user experience in interactive, responsive ways, particularly in situations where hands-free interaction is desirable.

In [6] their study, which is presented at the International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies explores a novel approach to controlling cursor movement through object detection powered by Vision Transformers. This research leverages the advanced capabilities of Vision Transformers to identify and track specific objects in real time, enabling precise cursor movement based on object positioning. By using Vision Transformers, known for their effectiveness in capturing spatial relationships and complex visual patterns, the system enhances the accuracy and responsiveness of cursor control without requiring physical input devices. This approach highlights the potential of deep learning models, specifically transformers, in improving human-computer interaction by offering touchless, gesture-based control. The study demonstrates how Vision Transformers can effectively process visual data to enable intuitive control interfaces, contributing to advancements in computer vision applications and furthering the development of hands-free digital interaction systems.

In [7] their study, it offers a comprehensive methodological and structural review of hand gesture recognition (HGR) across various data modalities. The authors explore a wide range of techniques used to capture hand gestures, including vision-based approaches (such as RGB cameras, depth sensors, and infrared cameras), as well as sensor-based modalities like accelerometers, gyroscopes, and electromyography (EMG). They analyze the advantages and limitations of each modality, emphasizing their application in different contexts like human-computer interaction, sign language interpretation, and assistive technologies. The review further discusses the challenges involved in HGR, such as dealing with noisy data, real-time processing requirements, and the need for robust algorithms to accommodate diverse gesture variations and environmental conditions. The authors also highlight recent advancements in deep learning methods, which have significantly improved the accuracy and efficiency of HGR systems. Overall, the study provides valuable insights into the state-of-the-art techniques, while proposing directions for future research in this rapidly evolving field.

In [8] their study, they present a novel approach to hand gesture recognition (HGR) for multi-cultural sign languages, integrating graph-based models with general deep learning networks. The authors address the challenge of recognizing diverse hand gestures across different sign languages, which vary not only in gesture structure but also in cultural context. They propose a hybrid model that combines the strengths of graph theory, which can effectively capture spatial and temporal relationships between hand gestures, with the flexibility and adaptability of deep learning networks for learning complex patterns in data. The paper reviews the limitations of traditional gesture recognition systems, such as difficulties in handling variations in sign language due to different hand shapes, orientations, and speeds, and introduces their method's ability to generalize across multiple languages. The authors also discuss the importance of cross-cultural data integration and training strategies to improve system performance and accuracy. Their work contributes to advancing the application of HGR in multi-cultural settings, paving the way for more inclusive sign language recognition systems that can bridge communication gaps in diverse global contexts.

In [9] this paper presents an innovative approach to vision-based hand gesture recognition (HGR) by combining Single Shot Detector Convolutional Neural Networks (SSD CNN) with deep dilated masks for processing video sequences and also addresses the challenges inherent in gesture recognition from dynamic video data, where variations in lighting, background, and hand movement can complicate accurate detection and classification. The authors provide a detailed review of existing vision-based HGR methods, highlighting the limitations of conventional approaches that struggle with real-time performance and robustness to diverse gesture variations. Their proposed method leverages SSD CNN to efficiently detect hand gestures while employing deep dilated masks to enhance the model's ability to capture fine-grained spatial features in video sequences. By combining these techniques, the approach achieves improved accuracy and speed in recognizing gestures, even in complex and cluttered environments. The paper contributes significantly to the field by offering a robust solution for real-time hand gesture recognition, with potential applications in areas such as human-computer interaction, sign language recognition, and virtual reality systems.

In [10] this paper it explore a novel approach to hand gesture recognition for smart healthcare applications, utilizing a Convolutional Neural Network (CNN)-based detector in combination with a Deep Belief Network (DBN). The paper provides a detailed review of existing techniques in healthcare HGR, highlighting the need for accurate, real-time gesture recognition systems to support hands-free operation in medical environments. The authors discuss the challenges faced by traditional methods, including sensitivity to varying lighting conditions, hand occlusions, and the need for high computational efficiency. Their proposed approach integrates the CNN-based detector for effective feature extraction and gesture localization, with the DBN for advanced classification capabilities, enabling the system to recognize a wide range of hand gestures with high accuracy. The study emphasizes the potential of this hybrid model in applications such as patient monitoring, assistive devices, and remote healthcare systems, where gesture-based interfaces can enhance the usability and accessibility of healthcare technologies. Through their work, the authors contribute to advancing gesture recognition in smart healthcare, offering a robust solution for hands-free, intuitive interaction in critical medical settings.

In [11] this paper they highlighting the evolution of methods and their applications across various domains. The authors categorize existing approaches into three primary groups: feature-based methods, appearance-based methods, and deep learning techniques, analyzing their respective strengths and weaknesses. They discuss feature-based methods' reliance on handcrafted descriptors, which offer interpretability but struggle with variability in gestures and environmental conditions. Appearance-based methods, utilizing image-based features like shape and texture, are noted for their simplicity but lack robustness to occlusions and dynamic gestures. The review emphasizes the transformative impact of deep learning, which has enabled automatic feature extraction and significantly improved recognition accuracy, particularly in complex and dynamic scenarios. Additionally, the authors identify challenges such as real-time processing, computational efficiency, and adaptability to diverse datasets, suggesting future directions for enhancing HGR systems. Their work serves as a valuable resource for understanding the state-of-the-art in vision-based HGR, providing a foundation for further advancements in the field.

The [12] paper reviews the limitations of traditional gesture recognition methods, particularly those relying on vision-based systems, which are often constrained by environmental factors like lighting and occlusions. The authors discuss the emergence of wearable systems that leverage biosensors to capture physiological signals such as electromyography (EMG) and inertial data, providing more robust and context-independent gesture recognition. Their proposed system incorporates adaptive machine learning directly within the sensor, enabling real-time and energy-efficient processing without relying on external computational resources. The review highlights the advantages of this approach, including improved accuracy, low latency, and enhanced adaptability to user-specific gestures, making it ideal for applications in healthcare, prosthetics, and human-computer interaction. This work represents a significant advancement in wearable technology and its potential to revolutionize gesture recognition by addressing the limitations of conventional methods while expanding its applicability in various real-world scenarios.

The [13] paper reviews prior SLR methods, emphasizing the limitations of traditional approaches that rely on handcrafted features and rule-based algorithms, which often fail to generalize across varying sign languages and dynamic environments. The authors highlight the effectiveness of CNNs in automatically learning complex patterns and visual features from sign language data, enabling higher accuracy in recognition tasks. The study discusses the use of computer vision to preprocess and segment hand gestures, ensuring robust feature extraction even in challenging conditions such as occlusions and background noise. By combining CNNs with vision-based preprocessing, the proposed system addresses key challenges in SLR, such as inter-user variability and real-time performance. The work is positioned as a significant contribution to the field, offering insights into the potential of deep learning and computer vision to create scalable and inclusive sign language recognition systems.

The [14] paper explore the innovative use of a three-dimensional (3D) virtual mouse model to generate synthetic training data for behavioral analysis, offering a novel perspective on addressing data scarcity challenges in machine learning applications and also reviews existing methods for behavioral data generation, emphasizing their reliance on labor-intensive manual annotations or limited real-world datasets that often lack variability. The authors propose the use of a 3D virtual mouse to simulate complex and diverse behavioral patterns, providing a rich, annotated dataset for training machine learning algorithms. This approach is particularly beneficial for scenarios where ethical considerations or logistical constraints limit access to live data. The review highlights the potential of synthetic data to enhance model robustness, improve generalization, and reduce biases. By bridging the gap between simulated and real-world scenarios, the study significantly contributes to advancing behavioral analysis techniques and underscores the broader applicability of synthetic training data in fields such as neuroscience, robotics, and human-computer interaction.

The [15] paper reviews prior advancements in human-computer interaction (HCI), focusing on the growing demand for intuitive and contactless control systems. Traditional input devices, such as physical mice and keyboards, are increasingly being complemented or replaced by virtual and voice-based systems to cater to accessibility and usability requirements. The authors examine existing gesture recognition techniques, highlighting limitations in accuracy and responsiveness due to environmental factors and device constraints. By employing deep learning for gesture recognition, the proposed system achieves enhanced accuracy in detecting and interpreting hand movements, while incorporating voice commands for additional functionality. This hybrid approach leverages convolutional neural networks (CNNs) for gesture detection and natural language processing (NLP) for voice assistant integration. The study demonstrates the system's potential in applications such as assistive technology, remote work setups, and smart environments, showcasing its ability to improve HCI experiences through seamless and efficient control mechanisms.

2.2 MOTIVATION AND SCOPE OF THE WORK

The motivation for developing a Real-Time AI Virtual Mouse System Using Computer Vision arises from the increasing demand for contactless, efficient, and intuitive human-computer interaction (HCI). Traditional input devices like physical mice and touchpads have limitations, particularly in terms of hygiene, ergonomics, and accessibility for individuals with physical disabilities. A virtual mouse system powered by computer vision eliminates the need for physical contact, addressing hygiene concerns and enabling hands-free operation.

This technology is particularly relevant in environments such as healthcare, shared workspaces, and public kiosks, where minimizing surface contact is critical. Additionally, advancements in computer vision and deep learning have made real-time gesture recognition feasible, allowing for the development of responsive and user-friendly systems. By leveraging hand tracking and gesture-based controls, the project aims to enhance productivity and accessibility while showcasing the potential of AI to transform everyday computing tasks into seamless, futuristic interactions.

SCOPE

The scope of the project is to create a contactless, efficient alternative to traditional input devices by leveraging real-time hand gesture recognition and tracking. The system will utilize computer vision and deep learning to detect hand movements and execute mouse operations like clicking, scrolling, and dragging. It aims to provide a hygienic solution for touchless interaction, particularly valuable in healthcare, shared workspaces, and public kiosks. Additionally, the project focuses on enhancing accessibility for users with physical disabilities and scalability for applications in gaming, virtual reality, and IoT ecosystems. By integrating advanced technologies, the project seeks to redefine human-computer interaction with a seamless and intuitive user experience.

2.3 PROBLEM STATEMENT

The problem addressed by the project is limitation of traditional input devices, such as physical mice and touchpads, which require direct contact and can be cumbersome, particularly in environments where hygiene, accessibility, or hands-free control is critical. Physical input devices are also not ideal for individuals with physical disabilities or those in need of more ergonomic solutions. This project aims to develop a touchless, gesture-controlled virtual mouse system that uses computer vision to detect and interpret hand movements in real-time. The system seeks to overcome challenges related to accurate and responsive gesture recognition, environmental variations, and the need for seamless, intuitive human-computer interaction, thereby providing a versatile solution for diverse applications in healthcare, smart environments, and assistive technologies.

3. PROPOSED METHODOLOGY

3.1 DESCRIPTION

The proposed method for developing a laptop controller using hand gesture recognition involves several stages, each aimed at accurately detecting, recognizing, and translating hand gestures into actionable laptop commands. Here's a detailed outline of the method:

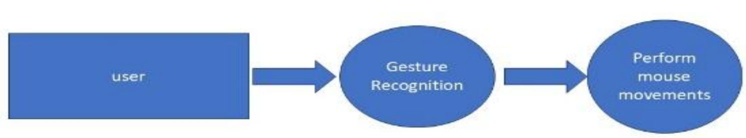


Fig.1. The general block diagram for the proposed technique

3.1.1. Dataset

To build a real-time AI virtual mouse system with computer vision, datasets are essential for training and testing the model to ensure it accurately recognizes hand gestures and finger movements. Specifically, you'll need datasets that include hand gesture photos, videos, hand landmarks, and finger detection data. These datasets will help the system understand various hand positions, finger orientations, and gestures that can be mapped to virtual mouse commands. Here are some examples of relevant datasets.

Hand Gesture Dataset:

The Hand Gesture Recognition Dataset provides a comprehensive collection of tagged photos and videos that capture a wide array of hand movements, essential for developing a virtual mouse control system. This dataset includes specific gestures such as pointing, pinching, swiping, and other dynamic movements that can be mapped directly to mouse activities, enhancing the system's usability for intuitive, touch-free control. Each gesture type in the dataset is tagged with metadata describing the hand position, orientation, and movement pattern, allowing for precise recognition and differentiation between similar gestures.

Hand Pose Estimation Dataset:

Hand pose estimation is essential for enabling accurate recognition of hand gestures in virtual mouse systems, as it provides precise information on hand and finger positions. These datasets offer rich annotations for keypoints, landmarks, and segmentation masks that support various hand pose estimation tasks, making them valuable for developing an AI-powered virtual mouse. Here's an expanded overview of key datasets and their contributions to hand pose estimation.

Finger Detection Dataset:

The Fingerspelling Recognition Dataset includes images and videos of hand signs used in fingerspelling, primarily from sign languages, capturing detailed finger positions and configurations. Although originally developed for fingerspelling recognition, this dataset can be adapted for finger recognition and tracking in virtual mouse applications. It contains diverse hand poses with specific finger arrangements, making it highly valuable for training models to distinguish between various finger positions and motions. By recognizing distinct finger configurations, the virtual mouse system can be trained to interpret individual finger movements as commands, such as single or double clicks, swipes, or scrolling.

3.1.2. METHODDOLOGY

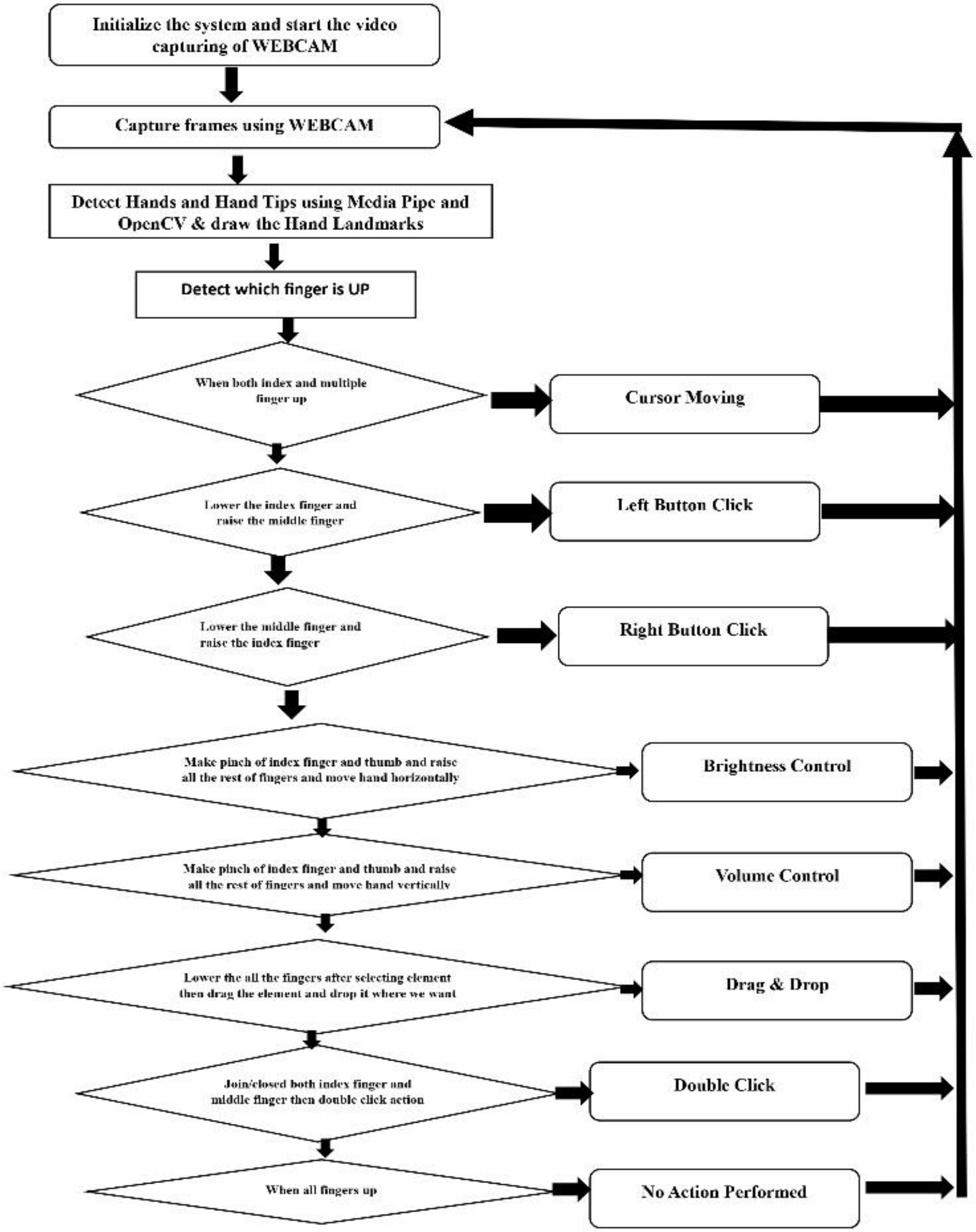


Fig 2: Flowchart

Mediapipe

The creation of a Real-Time AI Virtual Mouse System with MediaPipe to enable touchless human-computer interaction. Using MediaPipe's hand-tracking capabilities, the system identifies 20 hand landmarks in real time, allowing users to direct the mouse pointer with natural hand gestures. The index fingertip is used to move the pointer on the screen, while a pinch gesture between the thumb and index finger simulates mouse clicks. The Python-based system takes livevideo input, detects hand landmarks, and utilizes PyAutoGUI to transform gestures into mouse motions and actions. With applications in virtual presentations, gaming, public kiosks, and accessibility, this method offers a touchless interface that is responsive and easy to use. The system displays good accuracy and responsiveness in a variety of situations, indicating that it has the potential to improve user engagement without the usage of physical input devices. Future development could include multi-gesture recognition and more advanced control options to enhance usability and accuracy**.**

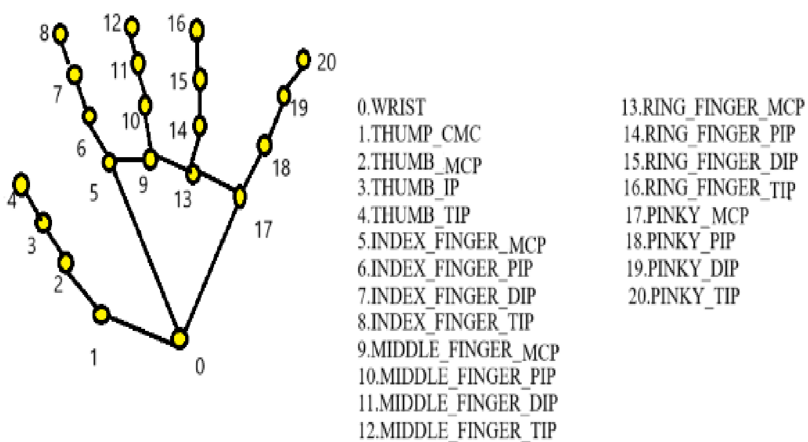


Fig 3: Landmarks of the hand

OpenCV

The computer vision library OpenCV includes object detection picture processing methods. The computer vision library OpenCV is a Python programming language library that may be used to create real-time computer vision applications. The OpenCV library is used for analysis tasks including object and face detection as well as image and video processing.

The Camara Used in AI Virtual Mouse System

The camera in an AI virtual mouse system is critical in capturing real-time hand motions for processing. A depth-sensing camera is commonly used to properly track hand movements. High-resolution cameras boost the system's capacity to perceive fine gestures, resulting in more precision and responsiveness from the virtual mouse.

Capturing the Video Processing

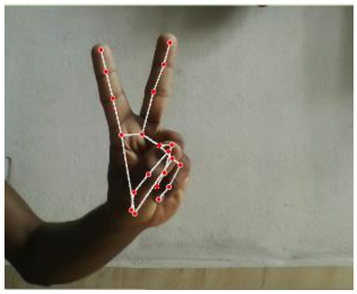
In an AI virtual mouse system, video processing entails recording continuous frames from the camera and detecting hand gestures in real time. Advanced image processing techniques such as background reduction, contour detection, and colour segmentation are used to detect and track hand movements. The processed data is then utilized to imitate mouse operations like clicking and pointer movement.

Detecting Which Finger Is Up and Performing the Particular Mouse Function

The AI virtual mouse system detects which finger is up using gesture recognition algorithms that examine hand features and finger placements. Specific fingers correspond to distinct mouse operations, such as using both index and multiple fingers to move the cursor and a combination of fingers to click or scroll. This mapping allows for seamless interaction by allocating mouse actions to different finger motions.

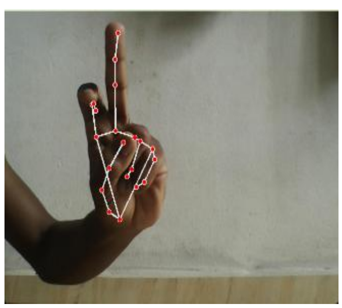
Mouse Functions

For Mouse Cursor Moving



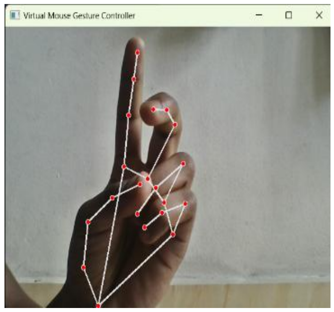
In the AI virtual mouse system, the pointer travels based on real-time hand motion tracking, specifically the position of both index and multiple fingers up. As the user moves their hand, the system calculates the coordinates and converts them into cursor motions on the screen. This allows for touchless, smooth navigation controlled purely by gestures.

To Perform Left Button Click



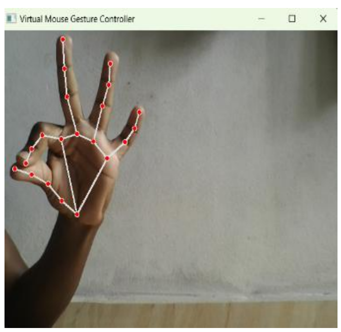
To accomplish a left button, click in an AI virtual mouse system, the user usually holds up a certain finger gesture, such as squeezing the lower index finger and raise the middle finger together. The system detects this gesture via image processing and activates the left-click function. This gesture-based interaction eliminates the requirement for physical mouse clicks.

To Perform Right Button Click



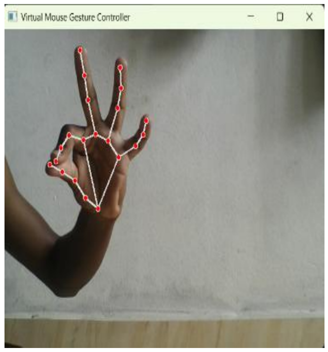
To click the right button in an AI virtual mouse system, a specific gesture, such as both lower the middle finger and raise the index finger, is used. The system recognizes this gesture using hand tracking and interprets it as a right-click operation. This allows for hands-free control of right-click functions without the usage of a hardware mouse.

Brightness Control



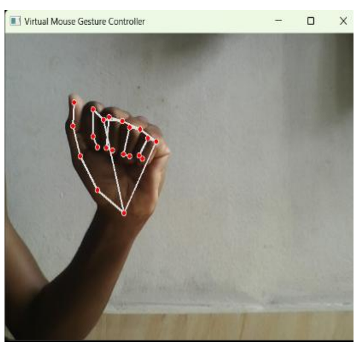
In an AI virtual mouse system, brightness control is frequently accomplished using hand movements that modify screen brightness. The system detects pinch of index finger and thumb finger and raises all the rest of the fingers moves horizontally, and adjusts the brightness accordingly. This enables users to adjust brightness settings without using any physical controls.

Volume Control



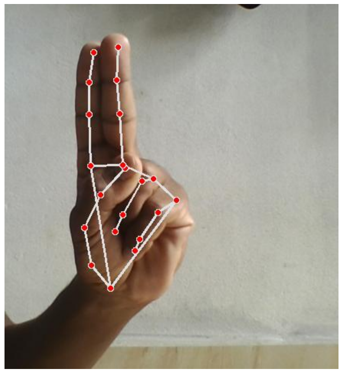
Volume control in an AI virtual mouse system is managed by gesture recognition, which adjusts audio levels based on precise hand movement. Make a pinch of the index finger and thumb and raise all the rest of the fingers and move the hand vertically. This hands-free control provides an easy way to manage sound without using any physical buttons.

Drag and Drop



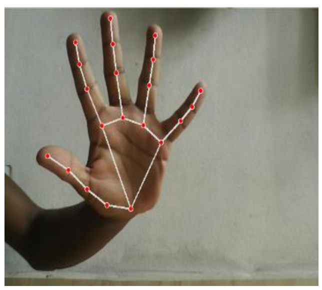
An AI virtual mouse system supports drag-and-drop by continually tracking hand gestures. The user performs a certain motion, lowering all fingers after picking the element, then dragging and dropping it wherever we desire. Releasing the motion completes the drop, allowing for flawless item control without the need for a real mouse.

Double Click



In the AI virtual mouse technology, a double-click is achieved by squeezing the index and middle fingers together. The system recognizes this gesture and activates the double-click function on the screen.

No action performed



When all five fingers are lifted, the AI virtual mouse system interprets the motion as a "stop" command, halting cursor movement. This feature eliminates undesired activity and allows users to temporarily pause their involvement without manually resetting the system.

3.1.3. Transfer Learning

Transfer learning can greatly improve the creation of a Real-Time AI Virtual Mouse System employing computer vision. This method eliminates the need for intensive data gathering and training by fine-tuning a model that has already been trained on huge datasets, like Media Pipe's hand-tracking model, for particular tasks like mouse control. The system employs transfer learning, which leverages previously acquired knowledge to precisely identify hand motions and landmarks in real time. The movement of the index finger, while clicks are simulated by squeezing the thumb and index finger together. This strategy increases the system's efficiency and accuracy, making it more durable in a variety of contexts and users. Transfer learning speeds up the deployment of such AI-driven systems and allows them to adapt to a variety of real-world applications such as accessibility, virtual interfaces, and touchless interaction systems without requiring large processing resources.

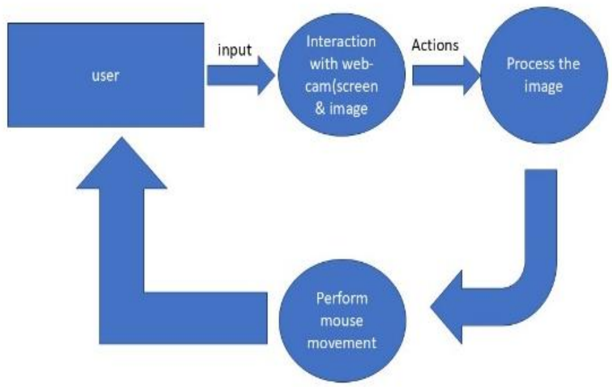


Fig. 4. Flow chart of the real-time AI Virtual Mouse System.

4. CONCLUSION

4.1. Conclusion

The development of a hand gesture-based laptop controller using computer vision demonstrates the potential for creating intuitive, non-contact interfaces that enhance user interaction. By leveraging image processing and gesture recognition algorithms, the system successfully interprets gestures in real-time to control various laptop functions, offering an accessible and hygienic alternative to traditional input devices. This innovation holds promise for applications in areas such as smart home systems, gaming, and robotics. While challenges like varying lighting conditions and environmental noise exist, future advancements in deep learning and

AR/VR integration can improve accuracy and scalability. Overall, this project highlights a significant step toward more intuitive and inclusive human-computer interaction technologies.

4.2. Future Scope

The hand gesture-based laptop controller using computer vision holds significant potential for future advancements, including enhanced gesture recognition accuracy under diverse lighting and environmental conditions through advanced AI models. Personalization features can be introduced to adapt to individual user preferences, such as hand sizes and gesture styles. Expanding the gesture library to include complex multi-gesture commands and integrating the system with AR/VR platforms can enable more immersive and versatile interactions. Additionally, optimizing the system for low-latency performance on lightweight hardware and ensuring cross-platform compatibility with devices like smartphones and IoT systems can broaden its applications and accessibility.

BIBLIOGRAPHY

[1]. Vatari, H., Narsude, S., Tonde, V., & Ramteke, J. (2024). Smart Mouse. Grenze International Journal of Engineering & Technology (GIJET), 10.

[2]. PRAMODH, M., SATYA, M., RAJESH, M., & SUBRAMANYAM, M. (2024). VIRTUAL MOUSE USING HAND GESTURE RECOGNITION.

[3]. Parimala, N., Teja, K. K. S., Kumar, J. A., Keerthana, G., Meghana, K., & Pitchai, R. (2024, April). Real-time Brightness, Contrast, and Volume Control with Hand Gesture Using Open CV Python. In 2024 10th International Conference on Communication and Signal Processing (ICCSP) (pp. 726-730). IEEE.

[4]. Shrivastava, A., Pundir, S., Sharma, A., Srivastava, A., Kumar, R., & Khan, A. K. (2023, June). Control of A Virtual System with Hand Gestures. In 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN) (pp. 1716-1721). IEEE.

[5]. Jagnade, G., Ikar, M., Chaudhari, N., & Chaware, M. (2023, January). Hand Gesture-based Virtual Mouse using Open CV. In 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT) (pp. 820-825). IEEE.

[6]. Surya, S. S., Kalaiselvi, S., Bernice, T., & Gunasekran, V. (2023, May). Cursor Movement Based on Object Detection Using Vision Transformers. In 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN) (pp. 1-5). IEEE.

[7]. Shin, J., Miah, A. S. M., Kabir, M. H., Rahim, M. A., & Al Shiam, A. (2024). A Methodological and Structural Review of Hand Gesture Recognition Across Diverse Data Modalities. IEEE Access.

[8]. Miah, A. S. M., Hasan, M. A. M., Tomioka, Y., & Shin, J. (2024). Hand gesture recognition for multi-culture sign language using graph and general deep learning network. IEEE Open Journal of the Computer Society.

[9]. Al Farid, F., Hashim, N., Abdullah, J., Bhuiyan, M. R., Kairanbay, M., Yusoff, Z., ... & Ramasamy, G. (2024). Single Shot Detector CNN and Deep Dilated Masks for Vision-Based Hand Gesture Recognition from Video Sequences. IEEE Access.

[10]. Alonazi, M., Ansar, H., Al Mudawi, N., Alotaibi, S. S., Almujally, N. A., Alazeb, A., ... & Min, M. (2023). Smart healthcare hand gesture recognition using CNN-based detector and deep belief network. IEEE Access, 11, 84922-84933.

[11]. Oudah, M., Al-Naji, A., & Chahl, J. (2020). Hand gesture recognition based on computer vision: a review of techniques. journal of Imaging, 6(8), 73.

[12]. Moin, A., Zhou, A., Rahimi, A., Menon, A., Benatti, S., Alexandrov, G., ... & Rabaey, J. M. (2021). A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition. Nature Electronics, 4(1), 54-63.

[13]. Murali, R. S. L., Ramayya, L. D., & Santosh, V. A. (2020). Sign language recognition system using convolutional neural network and computer vision. International Journal of Engineering Innovations in Advanced Technology ISSN, 2582-1431.

[14]. Bolaños, L. A., Xiao, D., Ford, N. L., LeDue, J. M., Gupta, P. K., Doebeli, C., ... & Murphy, T. H. (2021). A three-dimensional virtual mouse generates synthetic training data for behavioral analysis. Nature Methods 18(4), 378-381.

[15]. Raja, M., Nagaraj, P., Sathwik, P., Khan, K. M. A., Kumar, N. M., & Prasad, U. S. (2023, March). Voice Assistant and Gesture Controlled Virtual Mouse using Deep Learning Technique. In 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 156-161).