ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

Hand Gesture Video Navigation System

B. Sri Ramya[1]; Ch. Bhargavi[2]; P. Dhanumjaya[3]; P. N. J. S. Siri[4];
K. Vyshnavi[5]; G. Kavya[6]; Y. Pujitha [7]

[1] Assistant Professor

[1,2,3,4,5,6,7] Department of Computer Science Engineering, Sri Vasavi Engineering College, Tadepalligudem, Andhra Pradesh, India

Publication Date: 2025/04/30

Abstract: Hand gesture video navigation system control provides a seamless and touch-free way to interact with multimedia systems, enhancing user convenience and accessibility. This project presents an automated system that enables users to control video playback using only hand movements, eliminating the need for physical remotes or touch screens. The system recognizes specific gestures to perform essential functions such as playing and pausing videos, fast-forwarding and rewinding, increasing and decreasing volume, and zooming in and out. A camera captures hand movements in real time, and the system processes these gestures to execute corresponding video control commands. This intuitive approach enhances user experience, making video navigation more efficient and responsive. This project contributes to the advancement of touchless human- computer interaction, making video control more accessible and user-friendly across various applications. Here MediaPipe and OpenCV are playing a key role in the development of Hand gesture video navigation system.

Key words: *OpenCv, MediaPipe, Human ComputerInteraction (HCI), Gesture Recognition, Video playback*

How to Cite: B. Sri Ramya Asst.Professor; Ch. Bhargavi; P. Dhanumjaya; P. N. J. S. Siri; K. Vyshnavi; G. Kavya; Y. Pujitha (2025), Hand Gesture Video Navigation System. *International Journal of Innovative Science and Research Technology,* 10(4), 1781-1788. https://doi.org/10.38124/ijisrt/25apr992

I. INTRODUCTION

With the rapid growth of human-computer interaction, touchless technologies have become an innovative method to improve convenience and accessibility for users. One strategy is the use of a hand gesture-based control system. This approach allows users to navigate some aspects of their digital interfaces naturally and seamlessly. This project outlines the design of a Hand Gesture-Based Video Navigation System developed to provide users with a touch-free experience navigating multimedia content. This system works in place of a standard input system (e.g., remotes, keyboards, touchscreens) to provide a new way of interacting with a digital interface when touch is a nuisance or limitation for user interaction.

In order to interpret particular gestures, the suggested system uses a camera to record hand movements in real time. These movements are then processed using computer vision and machine learning algorithms. Essential video control features like play, pause, fast-forward, rewind, volume control, and zooming in and out are mapped to these identified gestures. The system provides a seamless and effective multimedia control experience by guaranteeing high responsiveness and accuracy in gesture recognition.

Improving accessibility for people with disabilities or mobility impairments is one of the main driving forces behind this project. For those who might find it difficult to use traditional input devices, this system offers an inclusive alternative by doing away with the need for physical remote controls or touch-based interfaces. Additionally, touchless control mechanisms aid in limiting the spread of contaminants and germs in settings like hospitals, public areas, and industrial settings where hygiene is a top concern.

Beyond entertainment, gesture-based video navigation finds use in industrial automation, smart homes, assistive technology, and healthcare. This project contributes significantly to the development of intelligent user interfaces and ubiquitous computing, which in turn promotes the creation of novel hands-free control mechanisms for a variety of applications as industries embrace touchless interaction more and more.

By offering a natural, effective, and intuitive method of controlling multimedia with hand gestures, this research ultimately aims to close the gap between human intuition and digital interaction. This system has the potential to revolutionize how users interact with digital content by doing away with the need for traditional input devices and providing a technologically advanced, easily accessible, and convenient substitute for traditional navigation techniques.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

II. LITERATURE SURVEY

The most important phase of software development is the literature review. This will provide an overview of some of the early research that many authors have done on this pertinent topic. We will also consider some important papers and keep refining our work.

Anklesh G suggested a research paper on Hand Gesture Recognition for Video Player in 2024 using OpenCV and flutter. He came to the conclusion that this research makes it possible for video players to recognize hand gestures with great precision, guaranteeing effective operation. It improves user involvement by providing a natural and easy way to control media playback with the usage of OpenCV [1]. Sakshi Shinde suggested a research paper on Gesture Based Media Player Controller in 2022 using OpenCV. She came to the conclusion that this program improves accessibility and convenience by providing a touch-free interface for hand gesture-based VLC control. For smooth media management, it enables users—including those with disabilities—to personalize movements [2]. Shruti tibhe suggested a research paper on Media Controlling Using Hand Gestures in 2023 using OpenCV and MediaPipe. She came to the conclusion that One more user-friendly and effective method of managing media players is by using hand gestures. Using Python and OpenCV, the system suggested in this project can recognize hand gestures in real time and utilize them to operate a range of media devices. The system may find use in domains including gaming, public exhibitions, and home entertainment [3]. Manjunath R Kounte suggested a research paper on Video Based Hand Gesture Detection System Using Machine Learning in 2022 using 2D CNN and Temporal Shift Module (TSM). He came to the conclusion that in this research, a hardware-efficient dynamic gesture detection system for smart device control utilizing a Temporal Shift Model and CNN is proposed. By facilitating communication through natural hand gestures, it improves accessibility for individuals with disabilities. The model is intended for gesture-based PowerPoint slide control and can identify sixteen different gestures [4]. Rishabh Agrawal suggested a research paper on Real Time Hand Gesture Recognition for Human Computer Interaction in 2016 using OpenCV API in C++ language. He came to the conclusion that the suggested technique for real-time hand gesture identification yields excellent outcomes with great precision and accuracy, and it may be applied in a practical setting for engagement with a computer. It can precisely identify the fingertips, is far easier to use than a mouse, and has a lot of potential for expansion into other HCI applications [5]. Saransh Sharma suggested a research paper on A Static Hand Gesture and Face Recognition System for Blind People in 2019 using Haar cascade method, Linear Binary Pattern (LBP), Convex hull and convex defects algorithms. He came to the conclusion that with the use of LBPH recognizers and Haar cascade classifiers for real-time face recognition, this system functions as a virtual assistant for the blind and visually impaired. The Convex Hull and Convex Defects approach is used to identify hand gestures, and the YCbCr color scheme guarantees precise skin identification in a range of lighting conditions. The system identifies faces with 92% accuracy and recognizes gestures with 95.2% accuracy [6]. Siddharth Swarup Rautaray suggested a research paper on A Vision based Hand Gesture Interface for Controlling VLC Media Player in 2010 using Pyramid Lucas-Kanade Optical Flow, PCA, K-Mean, KNN Algorithms. He came to the conclusion that using hand gestures, this application presents a clever, touch-free interface for VLC control. By allowing users to control the media player from a distance without a keyboard or mouse, it improves accessibility and convenience [7]. Serkan Genç suggested a research paper on HandVR: a hand-gesture-based interface to a video retrieval system in 2014 using C++ using the OpenCV library. He came to the conclusion that We investigated how well a hand-based interface could formulate queries for video retrieval systems. Some queries are difficult for traditional mouse-based interfaces to handle because of their restricted flexibility in interaction. In order to improve engagement, we suggested a hand-gesture-based interface that capitalizes on the innate dexterity of human hands [8]. Ahmad Puad Ismail suggested a research paper on Hand gesture recognition on python and OpenCV in 2020 using OpenCV and Haar cascade method. He came to the conclusion that In summary, the project effectively used Python and OpenCV to recognize hand gestures by utilizing hand segmentation and detection methods with the Haar-cascade classifier. In line with the project's aims, it accomplished two major goals: (1) building a comprehensive system for computer vision-based hand gesture detection, recognition, and interpretation, and (2) producing a system that can identify numbers and sign language movements[9]. Yuting Meng suggested a research paper on Real-Time Hand Gesture Monitoring Model Based on MediaPipe's Registerable System in OpenCV,MediaPipe and FingerNet. He came to the conclusion that This study introduces the Registerable Gesture Recognition Dataset (RGDS), which consists of 1,600 photos and 32 gesture types. To improve feature extraction, a canonical transformation-based normalizing technique was presented. The suggested FingerComb block sped up model convergence and enhanced feature robustness. Furthermore, RGDS was used to test and construct an improved ResNet-based FingerNet model, which showed excellent gesture recognition accuracy. For realistic gesture-based interactions, these developments provide a workable approach[10].

III. EXISTING SYSTEM

To enable user-computer contact, the majority of the existing computer interaction systems rely on new technology and physical input devices. Even while these gadgets and user interfaces have changed a lot over time, they still have built-in drawbacks that affect accessibility, usability, and efficiency.

- *Wireless and Wired Mice*

Because of their accuracy and ergonomic design, mice continue to be one of the most popular input devices. Conventional wired mice use a USB connection, which guarantees continuous use without battery issues. However, obstacles include limited mobility and cable clutter. Although wireless mouse with 2.4 GHz or Bluetooth communication provide greater flexibility, they also come with drawbacks

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

such possible latency, battery dependence, and network interference. Furthermore, extended mouse use can lead to repetitive strain injuries (RSI), which is why ergonomic designs have been developed.

➢ *Trackpads and Touchpads*

Commonly featured in laptops, integrated touchpads enable scrolling, multi-touch gestures, and cursor movement without the need for an external device. To improve the user experience, some sophisticated touchpads offer haptic feedback and force touch. Touchpads, on the other hand, may not be as effective for extended work, lack the accuracy of a dedicated mouse, and cause discomfort for those used to more conventional pointing devices. Although they are frequently restricted to particular environments, external trackpads provide an experience comparable to built-in ones.

➢ *Interfaces with touchscreens*

Touchscreens allow for direct device engagement, they have completely changed computing. A common feature of smartphones, tablets, and certain computers, they offer an easy-to-use interface that supports tapping, pinching, and swiping. Although touchscreens are incredibly portable and user-friendly, they are less appropriate for applications that demand a high level of precision, including complicated data entry or graphic creation. In addition to displays being prone to smudges and scratches, extended touchscreen use can wear out fingers.

➢ *Voice Command*

AI-powered assistants like Siri, Google Assistant, Alexa, and Cortana, voice-based communication has become more and more common. These technologies enable users to manage smart devices, dictate text, and carry out tasks. Voice control is very helpful for accessibility since it allows people with disabilities to use devices hands-free. However, because voice recognition systems need internet access and data processing, which raises security threats, they have limitations in terms of speed, precision, and privacy. Furthermore, voice instructions are not the best option for intricate jobs like programming or design, and background noise might affect accuracy.

IV. PROPOSED SYSTEM

Real-time hand gesture detection is used by a gesture-based video navigation system to carry out different orders. Among the main features are the following:

➢ *Video Control Using Gesture Recognition*

- *Swipe Gesture:*

Without pushing buttons, you may easily move through material by swiping left or right to fast-forward or rewind the movie.

- *Hand Open Gesture:*

Without having to look for a pause button, you can easily interact with the video by pausing it with an open hand.

- *Thumbs-Up Gesture:*

By using a thumbs-up, you can tell the system to start playing the movie without using a remote control or tapping the screen.

- *Fist Gesture:*

To end a video or shut down a media player, clench your fist.

- *Volume Control:*

You can map hand actions, like raising or lowering a hand or rotating a fist, to change the volume levels.

➢ *Hand Tracking in Real Time with Computer Vision*

- Live video input from a webcam or built-in camera is processed by the system.
- It uses computer vision techniques based on deep learning to detect and track hand movements.
- For real-time tracking and gesture classification, libraries like OpenCV, MediaPipe, and TensorFlow are used.

➢ *Libraries for Computer Vision*

- *OpenCV:*

An open-source package for tracking in real time, object detection, and image processing.

- *MediaPipe:*

A framework created by Google that is ideal for gesture recognition and real-time hand tracking.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

V. IMPLIMENTATION PROCESS DIAGRAM

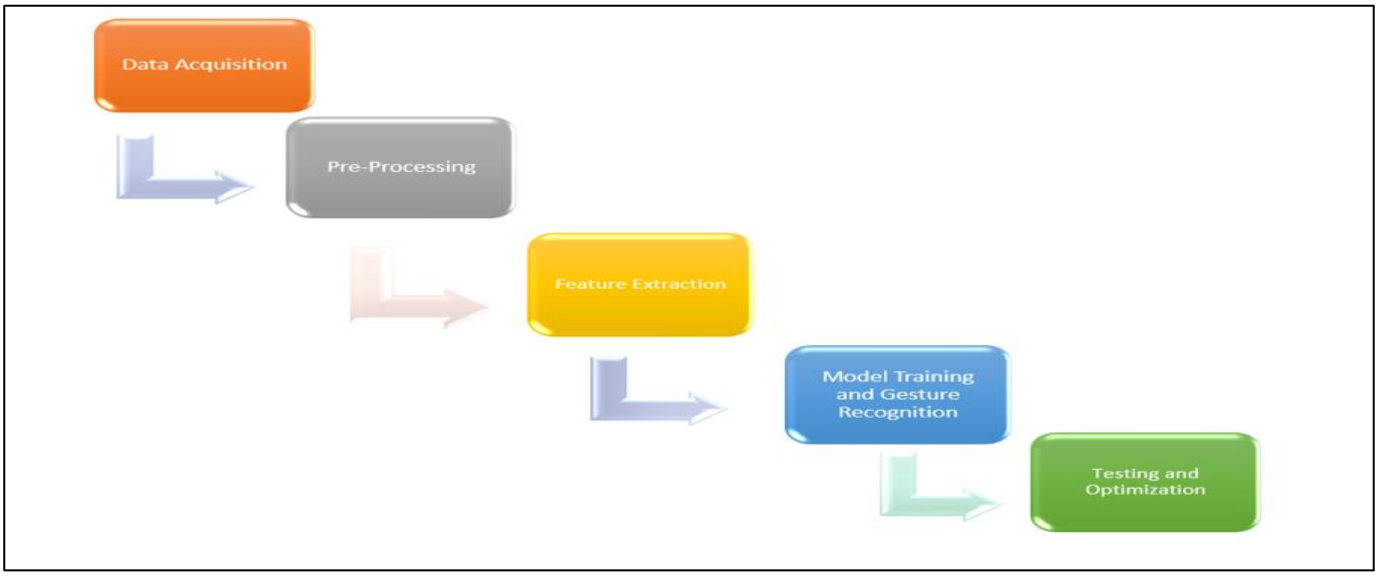

Fig 1 Implementation

Figure 1 Implementation illustrates the step-by-step process involved in hand gesture recognition for a media player.

➤ *Data Acquisition*

Real-time hand detection and tracking using a webcam is the only way the gesture-based video navigation system works. Without requiring labeled training data, the system recognizes hand motions by dynamically processing live video frames. In order to maximize performance, OpenCV extracts frames at a predetermined rate from continuously recorded video input from a camera. Advanced computer vision methods like MediaPipe Hand Tracking are used for hand detection, which enables the system to identify hands quickly without consulting existing datasets. While Haar Cascades provide pattern-based detection for real-time recognition, MediaPipe offers exact tracking by recognizing 21 important landmarks on the hand. After the hand has been identified, the features that have been retrieved are examined, and gesture-based commands are mapped to operate the video playback.

➤ *Preprocessing*

The collected frames are preprocessed before being entered into the machine learning model in order to improve accuracy. The first stage is RGB conversion, in which the Luminance Method is used to convert frames from OpenCV's default BGR format to RGB: $Y=0.2989R+0.5870G+0.1140B$. Better feature extraction is ensured by this conversion, which also helps to modify contrast and visibility. After that, resizing is done to keep all input data consistent by standardizing photos to a fixed resolution, like 128 by 128 pixels. Lastly, undesired distortions are removed using noise reduction techniques like Gaussian blur or median filtering, which improves hand detection accuracy and system performance in general.

➤ *Feature Extraction*

Finding important details in the hand image to differentiate between various gestures depends heavily on feature extraction. By identifying 21 important landmarks on each hand, MediaPipe Hand Tracking provides real-time tracking with little computing overhead and allows for accurate finger position and orientation recognition. Furthermore, form and texture data are extracted using the Histogram of Oriented Gradients (HOG) approach, which strengthens gesture detection in a variety of backgrounds. Edge detection techniques like Sobel and Canny are used to draw attention to hand boundaries in order to increase accuracy. This helps the system better identify different hand forms and movements.

➤ *Model Training and Gesture Recognition*

To guarantee precise hand gesture identification, the system makes use of a Convolutional Neural Network (CNN). Preprocessed images are received by the input layer, which then transforms them into an appropriate processing format. Convolutional layers extract crucial spatial data from the photos by applying many filters to identify edges, textures, and distinctive hand patterns. By reducing dimensionality while maintaining important features, a pooling layer (Max Pooling) increases computational efficiency. After being retrieved, the features are converted into a one-dimensional vector by passing them through fully linked layers. Lastly, the output layer determines the detected hand movement with high accuracy by assigning probabilities to various motions.

➤ *Video Navigation Control*

Once gestures are recognized, they are mapped to specific video playback controls.

- Five→ Play/Pause
- One/Two→ Forward/Rewind
- Three/Four→ Volume Increase/Decrease

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

- Wrist → Zoom in
- Five fingers move backward → Zoom out

➢ *Testing and Optimization*

Real-world testing is done on the system to improve its performance and accuracy. To guarantee flexibility, lighting conditions are assessed in a range of indoor and outdoor environments. For accurate hand detection, backdrop variations are evaluated with both simple and complicated backgrounds. Variations in distance and angle aid in confirming precise gesture detection from various viewpoints. Testing involves several users with different hand sizes, skin tones, and orientations. Generalization is enhanced by data augmentation methods including flipping, rotating, and introducing noise. The hand tracking parameters are adjusted for increased efficiency because deep learning is not utilized. Algorithm improvement improves real-time execution speed and accuracy. Smooth operation on common hardware is ensured by computational load reduction. To increase accuracy and resilience, the system is iteratively improved depending on test results.

VI. EXPERIMENTAL RESULTS

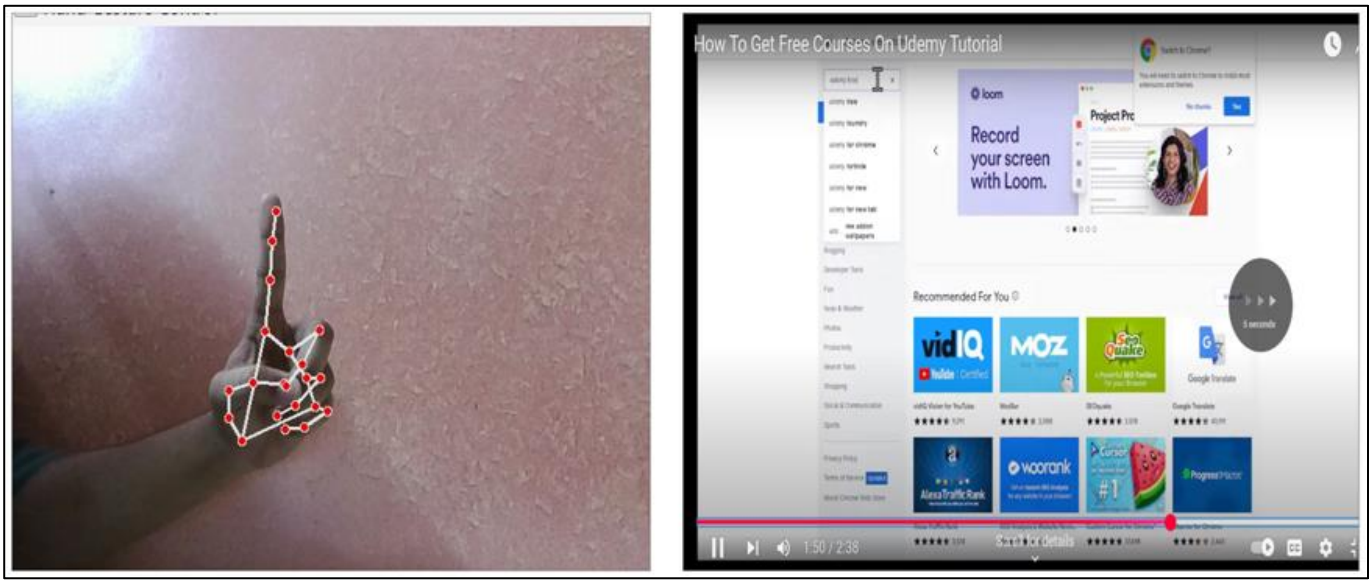

Fig 2 Forward

In the above image (Figure2), the recognized gesture successfully triggers the forward action, skipping the video by 5 seconds.

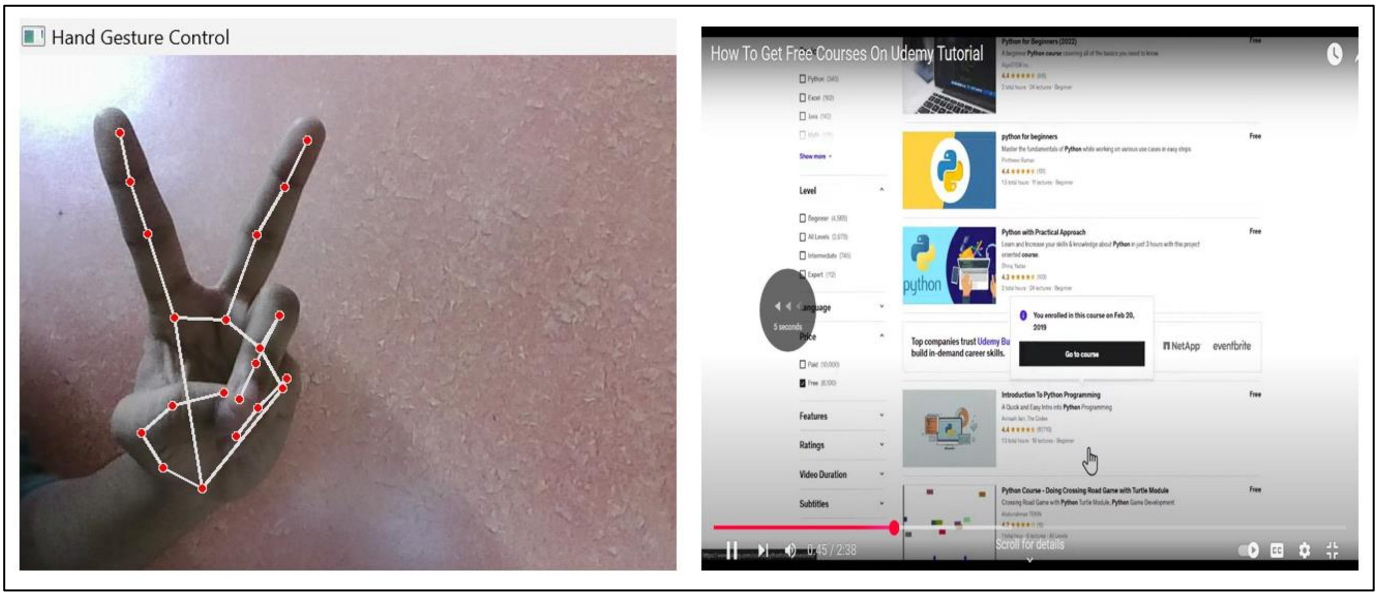

Fig 3 Backward

In the above image (Figure 3), the recognized gesture successfully triggers the backward action, rewinds the video by 5 seconds.

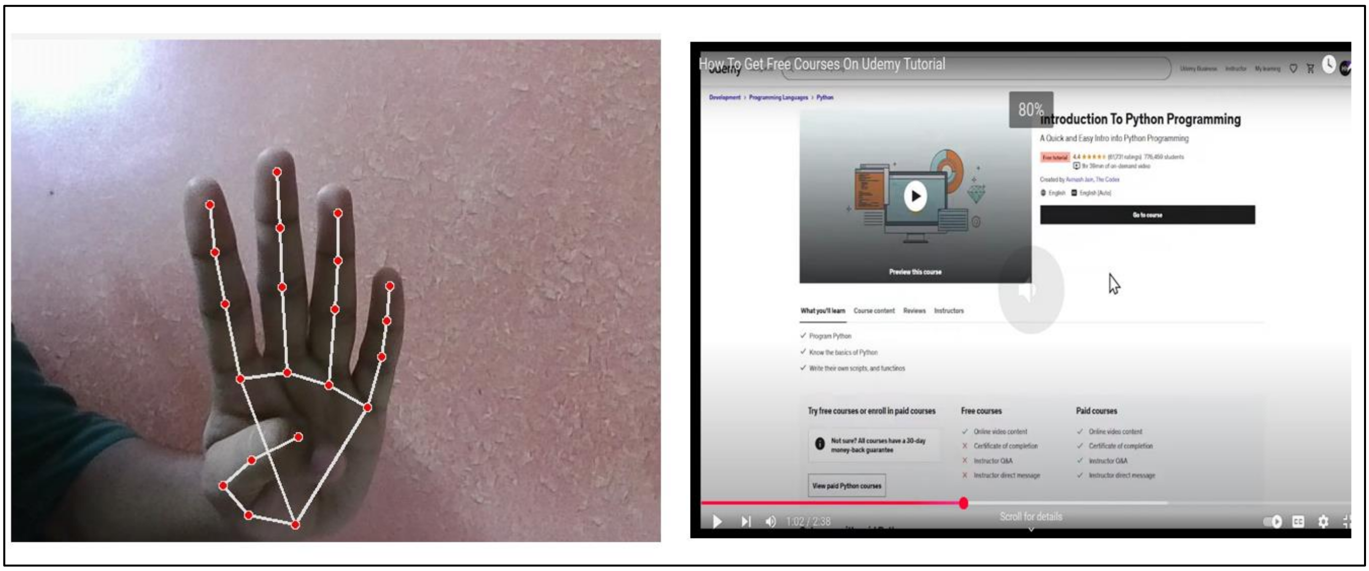

Fig 4 Volume Increase

The system accurately detects the hand gestures and increases the video volume accordingly. This enhances user interaction by providing seamless volume control through gestures (Figure 4).

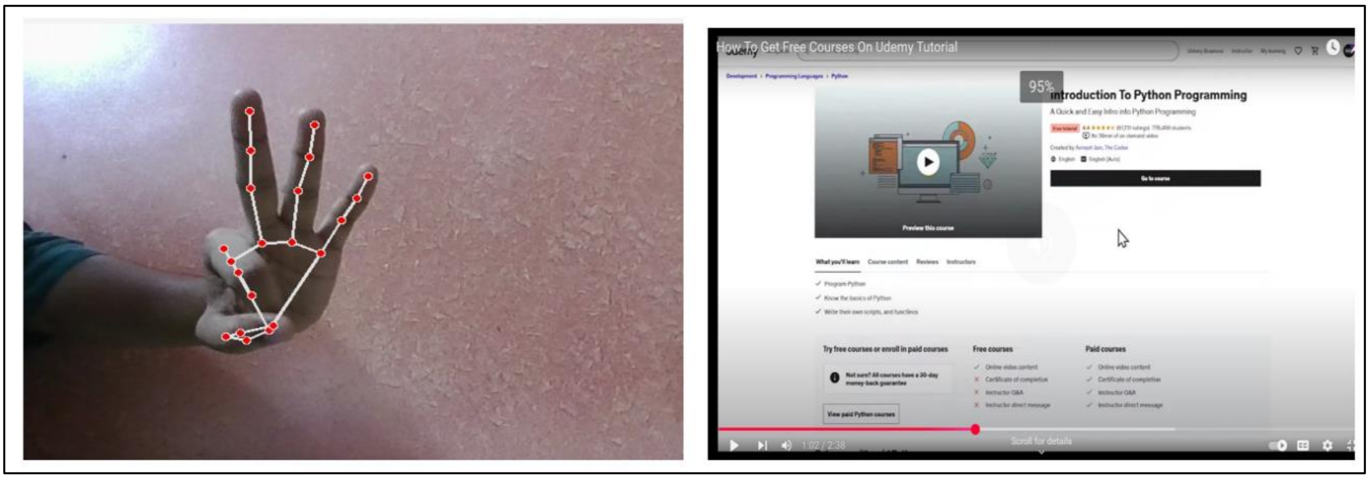

Fig 5 Volume Decrease

The system recognizes the hand gesture and decreases the video volume accordingly. This enables smooth and efficient volume control using intuitive hand movements (Figure 5).

Fig 6 Play/Pause

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

The system detects the hand gesture and pauses the video instantly. This provides a hands-free and seamless way to control video playback (Figure 6).

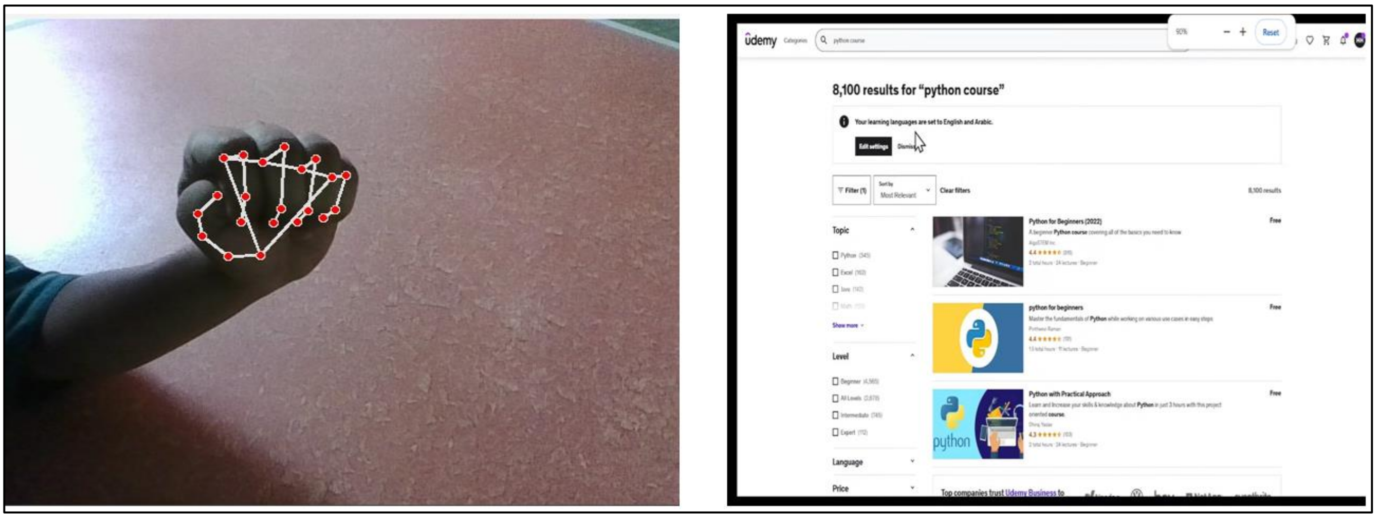

Fig 7 Zoom in

The system recognizes the fist gesture and triggers the zoom-in function. This reduces the content size, allowing a wider view of the screen (Figure 7).

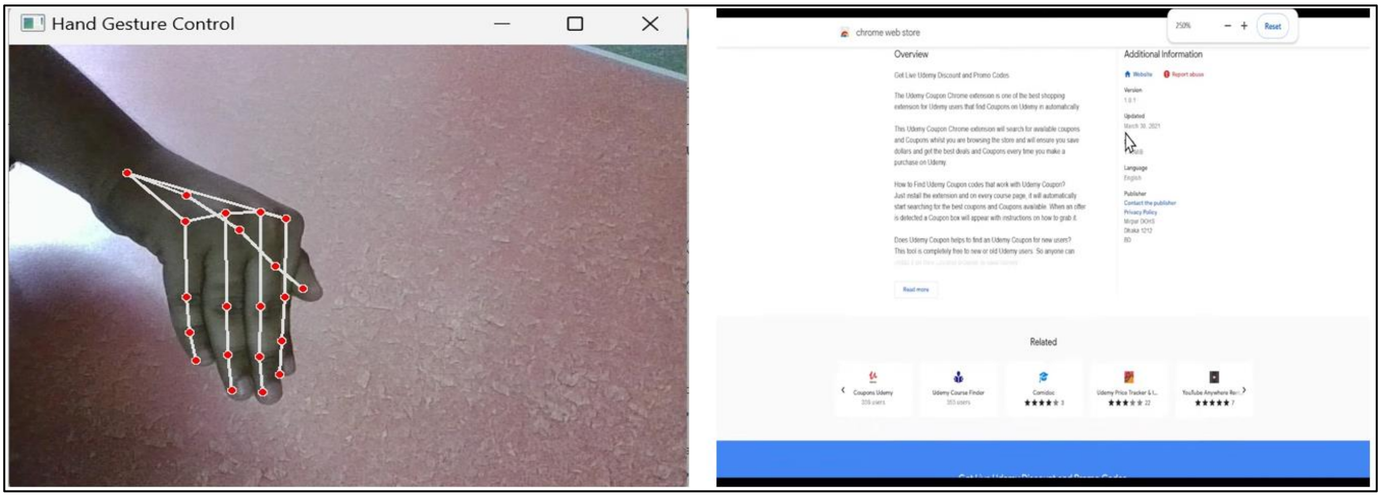

Fig 8 Zoom out

The system detects the five-finger backward movement gesture and activates the zoom-out function. This enhances the viewing experience by enlarging the content for better visibility (Figure 8).

VII. CONCLUSION

The hand gesture-based video navigation system offers a simple, touch-free method of controlling media playing, revolutionizing human-computer interaction. It tracks hand landmarks and recognizes gestures using OpenCV and MediaPipe, and PyAutoGUI maps these movements to media instructions. Simple hand gestures can be used by users to navigate, play, pause, and adjust volume. This method improves accessibility and provides those with mobility disabilities with an inclusive solution. By decreasing physical touch in public areas, it also acts as a hygienic substitute. Virtual reality, gaming, public displays, smart homes, and car infotainment systems can all benefit from the technology. For a seamless user experience, it guarantees real-time response and avoids unwanted behaviors. Digital media control is about to be redefined by advances in AI and gesture detection. Gesture-based navigation will be a common solution as contactless technology advances. Digital interactions are now more intelligent, effective, and futuristic thanks to this invention.

REFERENCES

[1] Anklesh G, Akash V, Prithivi Sakthi B , Kanthimathi.M – "Hand Gesture Recognition for Video Player" in 2024.
[2] Sakshi Shinde, Sarthak Mushrif, Aditya Pardeshi, Dhairyasheel Jagtap , Guide: Prof. Vandana Rupnar – "Gesture Based Media Player Controller" in 2022.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr992

[3] Shruti Tibhe, Ashwini Joshi, Aishwarya Warulkar, Aishwarya Sonawane , Miss. T.U.Ahirrao – “Media Controlling Using Hand Gestures” in 2023.
[4] Manjunath R Kounte, E Niveditha , A Sai Sudeshna , Kalaigar Afrose – “Video Based Hand Gesture Detection System Using Machine Learning” in 2020.
[5] Rishabh Agrawal, Nikita Gupta – “Real Time Hand Gesture Recognition for Human Computer Interaction” in 2016.
[6] Saransh Sharma, Samyak Jain, Khushboo – “A Static Hand Gesture and Face Recognition System For Blind People” in 2019.
[7] Siddharth Swarup Rautaray, Anupam Agrawal – “ A Vision based Hand Gesture Interface for Controlling VLC Media Player” in 2010.
[8] Serkan Genç, Muhammet Ba¸stan, Ugur Gudukbay ,Volkan Atalay, Ozgur Ulusoy – “HandVR: a hand-gesture-based interface to a video retrieval system” in 2014.
[9] Ahmad Puad Ismail, Farah Athirah Abd Aziz, Nazirah Mohamat Kasim, Kamarulazhar Daud – “Hand gesture recognition on python and opencv ” in 2020.
[10] Yuting Meng , Haibo Jiang, Nengquan Duan, Haijun Wen – “Real-Time Hand Gesture Monitoring Model Based on MediaPipe’s Registerable System” in 2024.