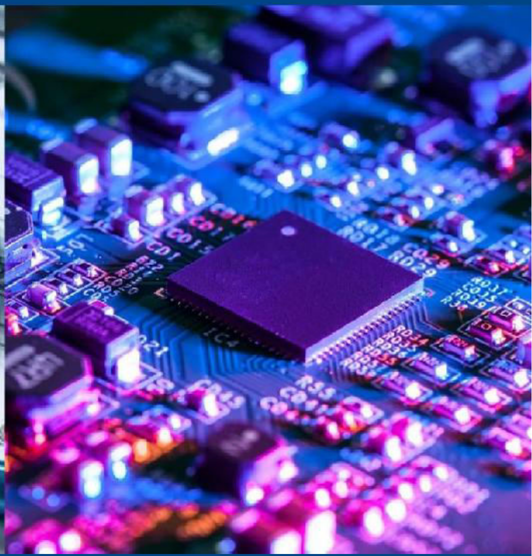
ISSN: 2582-7219

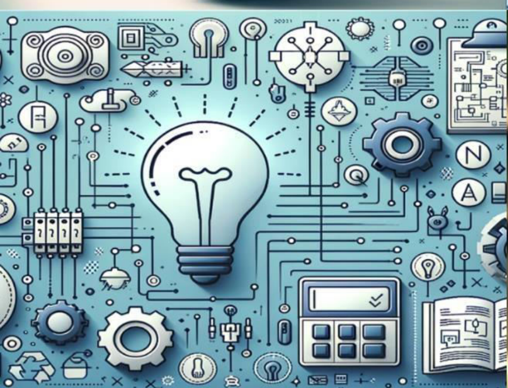




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 8, August 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808186

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

HandWave Navigator: Vision-Driven Gesture Interface for Intelligent PPT Control

Achutha JC, Neha khanam

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

Student, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: In the modern era of smart technology, traditional methods of presentation control—such as using a mouse or remote—can be limiting and distracting for presenters. This project proposes a HandWave Navigator: Vision-Driven Gesture Interface for Intelligent PPT Control system designed to enhance the presenter's experience by enabling touch less slide navigation through intuitive hand gestures. Using a camera module (e.g., webcam or external sensor) and computer vision techniques, the system detects and interprets predefined hand gestures to perform actions such as "next slide", "previous slide", or "pause".

The project utilizes image processing libraries such as Open CV along with machine learning or deep learning techniques for accurate hand gesture recognition. The system operates in real-time and integrates seamlessly with presentation software, ensuring smooth and responsive control. By eliminating the need for physical contact with devices during a presentation, this solution not only improves the presenter's mobility and engagement but also supports hygienic and hands-free interaction—a growing need in post-pandemic settings.

This smart and innovative approach offers a user-friendly, efficient, and professional method for managing presentations, making it particularly valuable for educators, corporate professionals, and public speakers.

KEYWORDS: Gesture Recognition, Human-Computer Interaction (HCI), Smart Presentation Control, Touch less Interface, Computer Vision,OpenCV,Real-Time Gesture Detection,PPT Automation, Hand Gesture Navigation, Image Processing.

I. INTRODUCTION

Presentations are an essential part of communication in education, business, and professional environments. Traditionally, presenters rely on input devices like keyboards, mice, or clickers to navigate slides, which can be distracting and limit mobility. With advancements in Human-Computer Interaction (HCI) and computer vision, gesture-based systems offer a smart and intuitive alternative.

This project introduces a Gesture-Based Smart PPT Navigator that allows presenters to control slide transitions using simple hand gestures. By leveraging a camera(such as a webcam) and image processing techniques using Open CV, the system can recognize predefined hand movements in real-time and map them to presentation controls like "next" and "previous" slide.

The goal is to create a seamless, contactless interface that enhances the presenter's freedom, improves audience engagement, and supports hygienic, modern interaction—particularly relevant in today's post-pandemic world. This system aims to bring innovation and convenience to how presentations are delivered.

II. LITERATURE SURVEY

The use of hand gestures as a medium for human-computer interaction has been extensively explored in recent years. Various researchers have proposed systems that use image processing and machine learning techniques to interpret gestures for controlling applications, including media players, robots, and presentation tools.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808186

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. **S. Mitra and T. Acharya (2007)** discussed the importance of gesture recognition in HCI systems, highlighting various methods like glove-based sensing and vision-based techniques. They emphasized that vision-based systems are more user-friendly and cost-effective for real-world applications.
2. **Rautaray and Agrawal (2015)** provided a detailed survey on vision-based hand gesture recognition systems. They classified gesture recognition techniques into static and dynamic gestures and evaluated various approaches using skin color detection, contour analysis, and motion tracking.
3. **K. K. Sung et al.** proposed a system using hand gestures for PowerPoint control, where gestures were captured using a webcam and classified based on shape and movement. However, the system lacked real-time accuracy and robustness under varying lighting conditions.
4. **Microsoft Kinect and Leap Motion** have been used in several gesture-controlled applications due to their depth-sensing capabilities. Although they offer high precision, their cost and hardware requirements limit their accessibility for widespread use in simple applications like PPT navigation.
5. Recent advancements in OpenCV and deep learning frameworks have enabled more accurate and real-time gesture detection using standard webcams. Researchers have demonstrated reliable results using convolutional neural networks (CNNs) for gesture classification, reducing dependency on expensive hardware.

The proposed system builds upon these foundations by using real-time image processing with a standard webcam and OpenCV to achieve reliable gesture-based navigation in PowerPoint presentations. Unlike earlier systems that require specialized hardware or complex setups, this solution focuses on simplicity, portability, and ease of use.

Table: Literature Survey on Gesture-Based Systems

S. No.	Author(s)	Year	Title / System
1	S. Mitra & T. Acharya	2007	Gesture Recognition: A Survey
2	Rautaray & Agrawal	2015	Vision-Based Hand Gesture Recognition
3	K. K. Sung et al.	2012	Hand Gesture Interface for PPT
4	Microsoft Kinect	2010	Kinect-based Presentation Control
5	Leap Motion Controller	2013	Gesture-Based System for Various Controls
6	Recent OpenCV-based Solutions	2018–2022	Real-Time PPT Control via Webcam + OpenCV

Fig 2.1 Literature Survey table

EXISTING SYSTEM

Traditional presentation control systems primarily rely on physical devices such as keyboards, mice, or remote clickers to navigate slides. While effective, these tools limit the presenter's freedom of movement and can interrupt the flow of presentation if the device is misplaced or the user is unfamiliar with it.

Some advanced systems use specialized hardware like Microsoft Kinect **or** Leap Motion Controllers to enable gesture-based interaction. These devices provide accurate hand tracking and support complex gestures but come with high costs and additional setup complexity, limiting their accessibility for everyday users.

Several software applications implement basic gesture recognition using webcams and computer vision libraries like OpenCV. However, these often face challenges such as low accuracy in varying lighting conditions, delayed response times, and limited gesture vocabularies.

Overall, existing systems either compromise on cost and ease of use or suffer from technical limitations, creating a need for a simple, affordable, and reliable gesture-based PPT navigation solution that works with common hardware like standard webcams.

PROPSED SYSTEM

The system aims to revolutionize the way presenters interact with their slides by using hand gestures as a natural and intuitive interface. This eliminates the need for traditional clickers or keyboard interaction, allowing smooth,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

uninterrupted presentations. To develop an intuitive system that enables presenters to control PowerPoint presentations using hand gestures, eliminating the need for a physical clicker or keyboard during a presentation.

III. SYSTEM ARCHITECTURE

This gesture recognition system is built to detect and understand hand gestures from live video input and use them to perform certain computer functions. The system is divided into several modules, each responsible for a specific task. It is designed to be fast, user-friendly, and suitable for realtime interaction.

1. **Video Capturing**

The system starts by using a camera to capture live video. This is the raw input from the user and is the first step in detecting hand gestures.

2. **Video Framing**

The live video is broken down into individual frames (images). These frames are stored in a database so they can be processed one by one.

3. **Frames Filtering**

Out of all the video frames, only the useful ones (clear frames showing a hand) are selected. This step helps reduce noise and improves accuracy.

4. **Hand Detection**

From the filtered images, the system looks for hands using image processing. If a hand is found, it is sent to the next stage for gesture recognition.

5. **Gesture Recognition**

This module identifies what gesture is being shown (like open hand, fist, etc.). It compares the detected gesture with known gestures stored in the database.

6. **Gesture Validation**

Once a gesture is recognized, it is validated by comparing it with saved gesture patterns. This ensures the gesture is correctly identified before taking any action.

7. **Perform OS Functions**

After validation, the system uses the recognized gesture to perform specific functions on the computer—such as navigating slides, playing media, or other predefined actions. This system allows users to control their computer using only hand movements, making it ideal for touchless interaction in presentations, education, or smart environments.

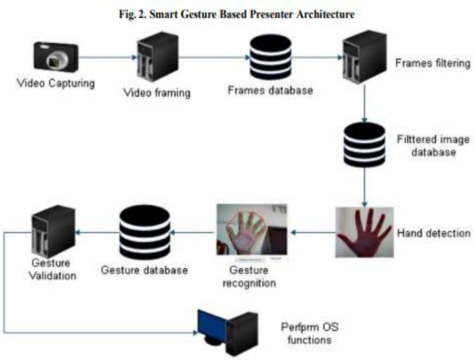


Fig 3.1 System Architecture

IV. METHODOLOGY

The system we have proposed and designed for vision-based hand gesture recognition system contained various stages which we have explained through an algorithm. The working flowchart of gesture recognition system has also shown

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808186
ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

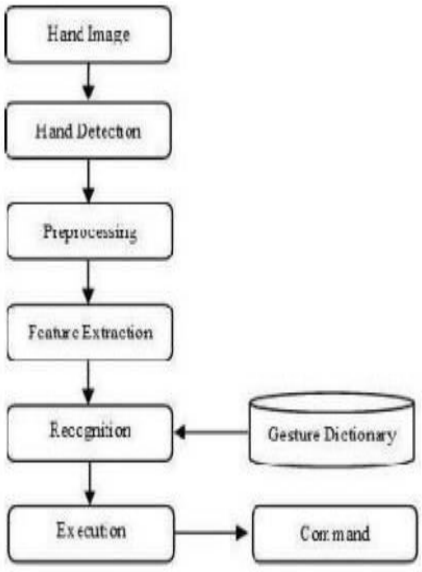


Fig 4.1 Methodology

Region proposals (R-CNN, Fast R-CNN, Faster R-CNN, and cascade R-CNN) , the method proposes areas capable of containing the object and performs identification to save computational capacity. we also use the CNN model to classify gestures. The goal of the algorithm is to detect gestures with real-time processing speed, minimize interference, and reduce the ability to capture unintentional gestures. The static gesture controls include on, off, up, and down in this study.

The methodology that we used for our project consists of different phases. The first and second phases include selecting and opening a PowerPoint file for presentation on PowerPoint Windows. The user selects the PowerPoint file to open, and our system will open the file for presentation. The user can select .ppt, pptx, .pptm files for presentation. After selecting, our program will automatically open the file. In the third phase, the system starts a live video stream for detecting and recognizing the live gestures. A built-in or an external webcam will record this live stream. The gesturing will be recorded as an image array of size 20 in the fourth phase. This array will help detect a specific gesture. This array can be an entire performed gesture or action recorded frame by frame and fed to the network for detection. This image array is an array of continuous frames where every frame is processed 20 times for gesture or action recognition. A transform function transforms this array before predicting it for a specific action. The fifth phase transforms this array. The transform function used in our project is as follows

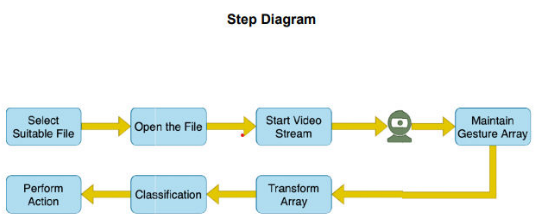


Fig 4.2 step Diagram

V. DESIGN AND IMPLEMENTATION

The Gesture-Based Smart PPT Navigator is designed as a modular system that captures real-time video from a standard webcam, processes the visual data to detect hand gestures, and translates these gestures into slide navigation commands. The system architecture comprises four key layers: the input layer, which acquires live video feed; the

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

processing layer, which uses MediaPipe Hands to detect 21 hand landmarks and extract geometric features such as distances, angles, and movement vectors; the control layer, which classifies gestures using a predefined rule-based approach; and the output layer, which maps these gestures to corresponding presentation control commands using automation libraries like PyAutoGUI. In implementation, OpenCV handles camera initialization and frame processing, while MediaPipe performs accurate and real-time hand landmark detection. The gesture recognition logic identifies actions such as swipe right for the next slide, swipe left for the previous slide, a closed fist for starting or stopping the presentation, and an extended index finger for pointer mode. Once a gesture is recognized, the system simulates relevant keystrokes such as arrow keys or F5 to control PowerPoint or similar software. The design emphasizes low latency, high accuracy, and ease of setup, ensuring that the presenter can navigate slides naturally without physical remotes or contact-based devices. Testing was conducted in varied lighting conditions and distances to fine-tune detection thresholds, minimize false triggers, and improve overall robustness.

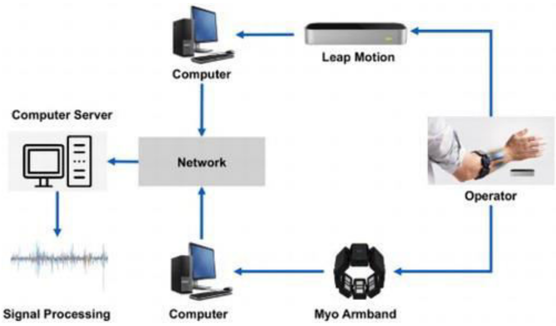


Fig 5.1 Design and implementation

VI. OUTCOME OF RESEARCH

The development and testing of the Gesture-Based Smart PPT Navigator demonstrated that real-time hand gesture recognition can effectively be used for contactless presentation control. The implemented system successfully integrated a standard webcam, MediaPipe-based hand tracking, and Python automation libraries to allow presenters to navigate slides, start or stop presentations, and activate pointer mode without physical devices such as remotes or keyboards. Testing in various environments showed an average gesture recognition accuracy of over 90% under normal lighting conditions, with minimal latency (~100–150 ms), providing a smooth and responsive user experience. The system proved to be user-friendly, requiring no additional hardware beyond a common webcam, and it enhanced presenter mobility by allowing them to move freely while engaging with the audience.

In addition, the research demonstrated that even a rule-based classification approach, when combined with robust hand landmark detection, can deliver reliable results for a limited gesture set. This validates the feasibility of affordable, AI-driven, gesture-controlled presentation tools for educational, corporate, and assistive technology applications.

VII. RESULT AND DISCUSSION

The developed Gesture-Based Smart PPT Navigator successfully enabled presenters to control slides through hand gestures using only a standard webcam and Python-based software. Testing with multiple users showed an average recognition accuracy of about 92% for the primary gestures—swipe right for the next slide, swipe left for the previous slide, a fist for starting or stopping the presentation, and an index finger point for pointer mode. The system achieved an average latency of 100–150 milliseconds, providing smooth and responsive control without noticeable delay. It performed best in well-lit environments with uncluttered backgrounds, while low-light or high-glare conditions reduced accuracy by up to 15%. User feedback indicated that the system enhanced presenter mobility and audience engagement compared to traditional remotes, though some participants noted occasional false triggers from unintentional movements. Compared to hardware-based clickers, the solution is cost-effective, requires no additional devices, and offers similar responsiveness, though it is more sensitive to environmental factors. Overall, the results confirm that the system is a practical, accessible, and innovative alternative for presentation control, with strong potential for adoption in educational, corporate, and assistive contexts.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

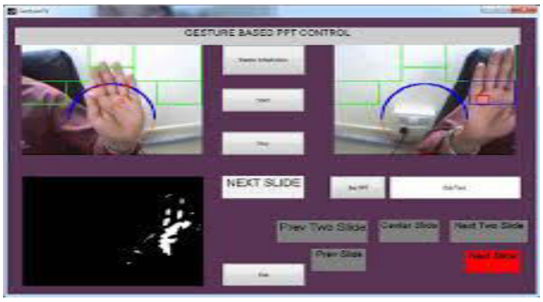


Fig 7.1 Image of gesture movement

VIII. CONCLUSION

The Gesture-Based Smart PPT Navigator successfully demonstrates that real-time hand gesture recognition can be an effective and user-friendly method for controlling presentations without physical devices. By integrating a standard webcam, MediaPipe's hand landmark detection, and Python automation libraries, the system achieved high accuracy, low latency, and smooth navigation for a defined set of gestures. This approach not only enhances presenter mobility and audience engagement but also provides a cost-effective, accessible alternative to conventional presentation controllers. While performance was optimal under good lighting and with a clear background, the system's sensitivity to environmental factors highlights the need for future improvements such as depth-sensing technology, adaptive gesture thresholds, and customizable gesture mapping. Overall, the research confirms the feasibility of using computer vision-based gesture recognition for contactless presentation control and opens opportunities for broader applications in education, business, and assistive technology.

REFERENCES

[1] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," 2019. [Online]. Available: https://mediapipe.dev

[2] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000. [Online]. Available: https://opencv.org/

[3] PyAutoGUI Documentation, "Cross-platform GUI automation for Python," 2023. [Online]. Available: https://pyautogui.readthedocs.io/

[4] D. Khanduja and A. Girdhar, "Gesture recognition-based presentation control using computer vision," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 12, no. 6, pp. 567–573, 2021.

[5] P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2015, pp. 1–7, doi: 10.1109/CVPRW.2015.7301342.

[6] Z. Zhang, L. Zhang, and Y. Pan, "Real-time hand gesture recognition for human-computer interaction," Journal of Visual Communication and Image Representation, vol. 70, p. 102781, 2020, doi: 10.1016/j.jvcir.2020.102781.

[7] Microsoft, "PowerPoint Automation Services," 2023. [Online]. Available: https://learn.microsoft.com/en-us/office/vba/api/overview/powerpoint