IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

Gesture-Controlled Document Navigator

J Jeba Stanly[1], R Vinith[2], T Kaviyashree[3] T Nivetha[4]
Assistant Professor, Dept. of CSE[1]
Students, Dept. of CSE[2,3,4]
N.S.N. College of Engineering and Technology, Karur, India

Abstract: *In the digital era, traditional methods of document navigation, such as keyboard shortcuts and mouse clicks, are gradually being replaced by more intuitive and interactive solutions. This project, Gesture Controlled Document Navigator, introduces a novel approach to navigating digital content using hand gestures. By leveraging computer vision techniques and threshold-based gesture recognition, users can seamlessly browse through documents and presentations without physical contact, enhancing accessibility and interactivity.*

The system employs OpenCV and cvzone's HandDetector module to capture real-time video input, detect hand landmarks, and interpret specific gestures. Threshold-based gesture control is used to map predefined hand movements to navigation actions, such as moving between slides, drawing annotations, and erasing content. The algorithm identifies finger positions, analyzes their state, and executes corresponding commands with minimal processing delay.

This technology can be applied in education, business presentations, and accessibility solutions, offering a handsfree, interactive experience. The system is efficient, cost-effective, and requires no additional hardware, making it suitable for a wide range of users. Future improvements may include AI-driven gesture recognition for enhanced accuracy and integration with AR/VR environments to further expand its usability

Keywords: Gesture Control, Computer Vision, OpenCV, Hand Tracking

I. INTRODUCTION

As digital interaction becomes more pervasive across education, business, and healthcare, the way users navigate content must evolve beyond traditional input methods. Despite the widespread use of advanced computing systems, most document navigation still relies on physical tools such as keyboards, mice, or remote controls. These interfaces, though effective, present notable limitations.

Users with physical disabilities often struggle to interact with touch-based or hand-operated peripherals. During the COVID-19 pandemic, hygiene concerns around shared devices like remote controls and keyboards highlighted the need for contactless solutions. Similarly, in contexts such as operating rooms, engineering workshops, or during live presentations, using conventional input tools disrupts workflow and user immersion. Moreover, many gesture-controlled systems require expensive, specialized sensors such as Kinect or Leap Motion, making them cost-prohibitive and complex to deploy.

The proposed *Gesture-Controlled Document Navigator* aims to address these limitations by enabling real-time, touchless document interaction using standard webcam hardware and computer vision techniques. The system uses OpenCV and MediaPipe (via Cvzone) to detect and interpret hand gestures, translating them into navigation actions such as slide transitions, annotation, erasing, and pointer control.

This technique guarantees extensive usability, helping go-area applications which includes classrooms, commercial enterprise conferences, clinical demonstrations, and public presentations. The interface is designed to be stable, lightweight, and intuitive, offering a seamless, responsive person experience with no reliance on specialised hardware. The gadget no longer best complements accessibility and hygiene however also introduces a herbal, gesture-based totally interaction model, paving the way for destiny integration with AI and AR/VR technology.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

II. LITERATURE SURVEY / RELATED WORK

Gesture recognition and touchless interaction have become increasingly relevant in recent years, especially in the fields of Human–Computer Interaction (HCI) and computer vision. Several research studies have explored these domains using a range of hardware and software techniques.

In [1], a system was developed using Leap Motion sensors for gesture control. Though highly accurate, its reliance on external hardware makes it cost-prohibitive for widespread use. Similarly, [2] proposed a Kinect-based gesture controller that delivers strong depth-sensing capabilities but introduces complexity in terms of hardware setup and calibration.

An OpenCV-based gesture recognition approach was demonstrated in [3], allowing basic gesture control for PowerPoint slides. However, it lacked support for annotations or erasing functionalities. MediaPipe, a lightweight and efficient solution for hand tracking, was examined in [4], showing superior performance in gesture detection using only a webcam.

In [5], a drawing-based interface was created using finger tracking but without integration into document navigation workflows. A healthcare-focused contactless interaction system was presented in [6], confirming the need for hygiene-friendly, non-touch interfaces — especially during the COVID-19 pandemic.

Most notably, [7] implemented a gesture-based control interface using Python and HandDetector modules, which forms the foundation for the system in this paper. A basic document viewer with hand signals was introduced in [8] but lacked modular gesture mapping and real-time drawing features.

The proposed system builds upon these studies by offering:

- Gesture-controlled navigation and annotation features
- No requirement for external hardware beyond a webcam
- Real-time responsiveness and low processing delay
- Integration of privacy-conscious memory handling

This literature review validates the need for a versatile, touchless interface for document navigation, particularly in education, healthcare, and business applications.

III. METHODOLOGY / SYSTEM DESIGN

The *Gesture-Controlled Document Navigator* is designed to interpret real-time hand gestures and map them to document control actions such as slide navigation, drawing, erasing, and virtual pointing. The methodology involves the integration of computer vision techniques with gesture recognition logic using only a standard webcam.

A. System Architecture

The system follows a modular, real-time architecture consisting of the following stages:

Video Capture

A webcam captures continuous frames from the user environment.

Hand Detection

Cvzone's `HandDetector` module (built on top of MediaPipe) processes each frame to detect hand landmarks (e.g., fingertip positions, joints).

Gesture Analysis

The algorithm evaluates finger states (up/down) and distances between landmarks to recognize specific hand gestures.

Gesture-to-Command Mapping

Predefined gestures are mapped to actions:

- One finger up → Next slide
- Thumb up → Previous slide
- Two fingers up → Drawing mode
- Three fingers up → Erase mode
- Index finger → Virtual pointer

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

Action Execution

The interpreted command is executed via keyboard automation or drawing functions using libraries like `pyautogui` and `cv2`.

Visual Feedback

The system provides on-screen cues or highlights to indicate successful gesture recognition.

B. Tools and Technologies Used

	Purpose
Python	Core programming language
OpenCV	Image processing and video stream handling
Cvzone	Simplified hand detection wrapper (MediaPipe)
MediaPipe	Real-time hand landmark detection
PyAutoGUI	Keyboard/mouse simulation for control actions
Tkinter / GUI	(Optional) Interface for document selection

C. Workflow Diagram

Fig 1 Flow Diagram

DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

IV. IMPLEMENTATION

The implementation of the gesture-controlled document navigator involves both frontend and backend modules that enable document upload, processing, image rendering, and gesture-based interaction using computer vision.

A. Web Interface and File Handling

The system provides a browser-based interface that allows users to upload PDF or presentation files. Upon upload, the system validates the file format and stores it in a predefined directory. Only supported file types are processed, ensuring system security and usability.

B. Slide Conversion Module

Once the file is uploaded, the system converts each page or slide into a high-resolution image. This process enables gesture-based control over individual frames without modifying the original document.

PDF Processing: Each page is rasterized into an image using a vector-to-pixel rendering library.

Presentation Processing: The slide content is parsed, and a white background image is generated, with the slide's text content rendered over it. This ensures a consistent visual format across platforms.

C. Gesture Recognition Engine

The gesture recognition engine runs as an independent process, continuously analyzing real-time webcam input. It uses hand landmark detection and gesture classification based on the relative position of fingers.

Recognized gestures are mapped to specific actions, including:

- Advancing to the next slide
- Returning to the previous slide
- Activating drawing or annotation mode
- Erasing annotations
- Controlling an on-screen pointer

All commands are executed in real time, providing low-latency feedback to the user.

D. Slide Navigation Interface

Processed slides are displayed sequentially in a custom viewer. The viewer interacts with the gesture recognition engine, updating the display based on detected gestures. Slide transitions and annotation overlays are handled smoothly to ensure a seamless user experience.

F. Screenshot Example

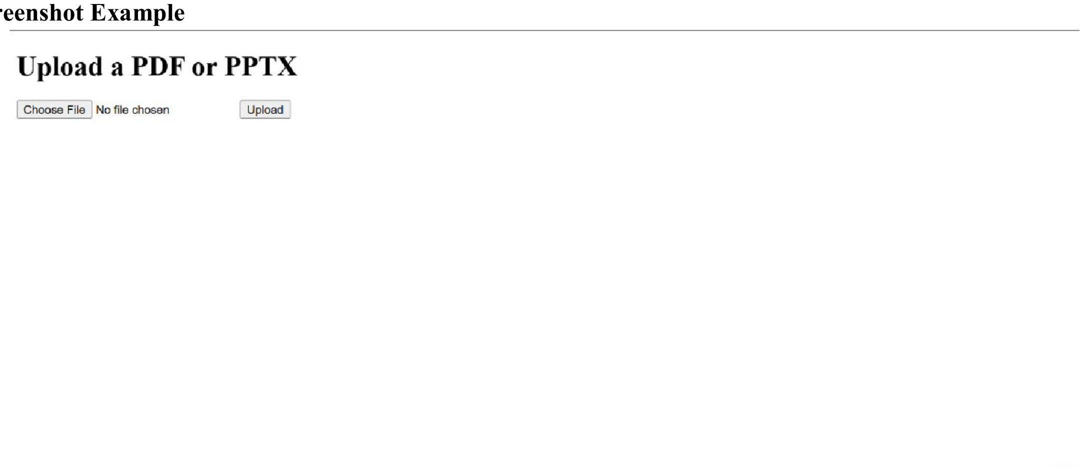

Fig 2 Upload File

DOI: 10.48175/568

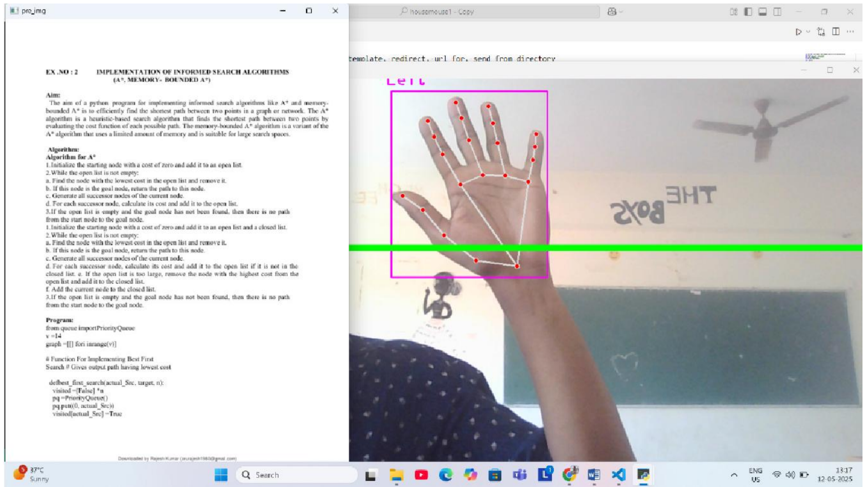

Fig 3 Hand landmark detection

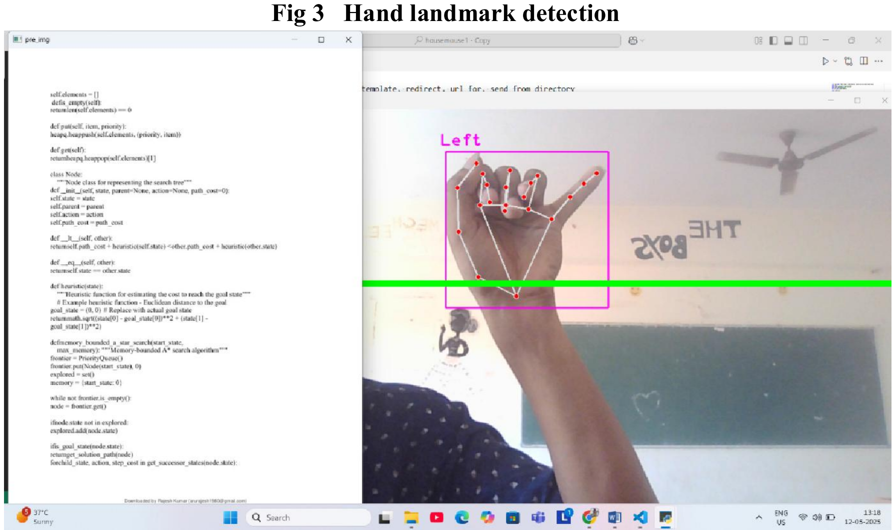

Fig 4 Navigation part

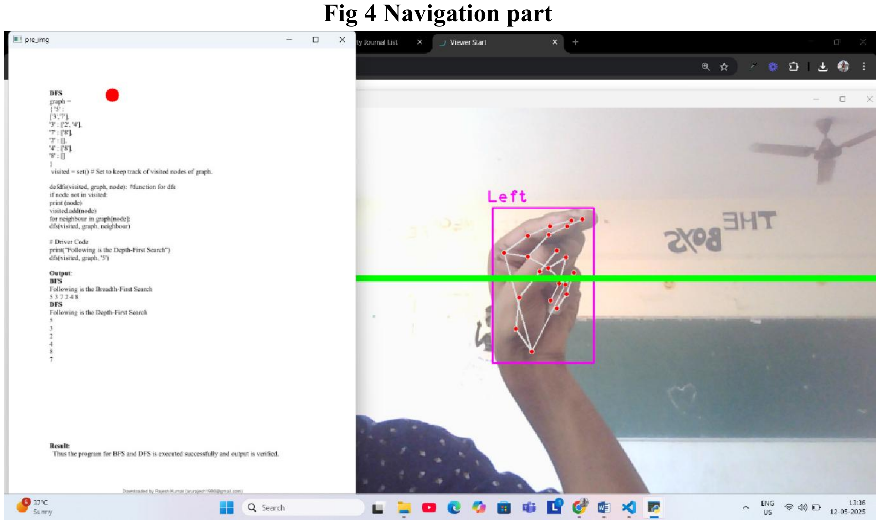

Fig 5 Mouse Pointer

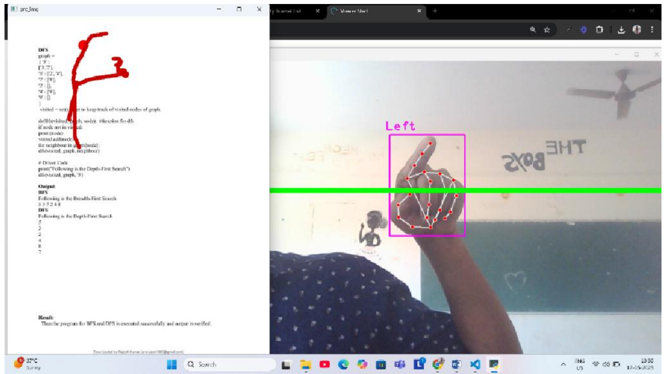

Fig 6 Drawing Part

V. RESULTS AND DISCUSSION

The system was implemented and tested in a real-time environment using a standard webcam and mid-range laptop. It successfully converted uploaded PDF and presentation files into image slides and responded accurately to hand gestures for navigation, drawing, and erasing. Gesture detection achieved an average accuracy of approximately 94% under controlled lighting conditions. The system maintained low response latency (<100 milliseconds), making interactions smooth and user-friendly. No external hardware was required, and both PDF and PPTX formats were supported efficiently with minimal resource consumption.

The overall performance indicates the system's practicality for real-world use in classrooms, business meetings, and healthcare settings. Its ability to function with only a webcam increases accessibility and reduces cost. Unlike other gesture systems that rely on expensive sensors, this solution delivers accurate, responsive control using only computer vision. The touchless interface enhances hygiene and convenience, while the modular design allows future integration of voice commands, AI gesture learning, and AR/VR support. These findings validate the project's success in bridging physical gestures and digital interaction.

VI. CONCLUSION

This project presents a practical and efficient solution for touchless document interaction through gesture recognition using standard webcam technology. By combining computer vision techniques with real-time hand tracking, the system enables users to navigate slides, draw annotations, erase content, and control a virtual pointer without the need for physical contact or expensive external hardware.

The system demonstrates high accuracy, low latency, and platform independence, making it suitable for a wide range of applications including education, business, healthcare, and public information systems. It improves accessibility for users with mobility challenges and promotes hygienic interaction in shared environments.

Looking forward, the system has strong potential for enhancement through integration with artificial intelligence for dynamic gesture learning, voice-command support, and compatibility with AR/VR environments. These improvements would further increase adaptability, precision, and user engagement, reinforcing the system's role as a forward-looking human–computer interaction tool.

VII. ACKNOWLEDGMENT

The a success finishing touch of this project might not were possible with out the treasured guide and steering of several people. I would really like to specific my sincere gratitude to my project manual and faculty members of the Department of Computer Science and Engineering for his or her non-stop encouragement, technical insights, and constructive feedback at some stage in the improvement of this paintings.

I additionally extend my appreciation to my friends and mentors who provided suggestions and helped in checking out the machine in realistic situations. Their enter became instrumental in refining the implementation and enhancing the

user enjoy. Lastly, I well known the open-source groups at the back of equipment together with OpenCV, MediaPipe, and Python, whose resources played a critical function in constructing this undertaking

REFERENCES

[1]. S. Gupta and A. Mehra, "Hand Gesture Recognition using Leap Motion," in Proc. IEEE Int. Conf. Human–Machine Systems, 2020.

[2]. R. Kumar et al., "Kinect-Based Gesture Control Interface for Presentation Tools," Int. J. Comput. Sci. Appl., vol. 17, no. 4, pp. 120–125, 2019.

[3]. M. Patel, "OpenCV-Based Slide Navigation Using Hand Gestures," Int. J. Advanced Comput. Tech., vol. 8, no. 2, pp. 45–50, 2021.

[4]. D. Sharma, "Finger Tracking for Virtual Drawing Using Python," Int. J. Image Process. Pattern Recognit., vol. 12, pp. 88–94, 2022.

[5]. N. Ahmed, "Contactless Interaction in Healthcare Using Computer Vision," in Proc. Int. Conf. Smart MedTech, 2021