The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40−47

Controlled Presentation and Interactive Whiteboard System

[1] Y.Suresh Gopi, [2],S.Padma Priya, [3],P.Chaitanya, [4] T.Sandeep, [5] P.Sai Kumar,
[1] Professor Department of Computer Science & Engineering
Nadimpalli Satyanarayana Raju Institute of Technology, Sontyam,
Vishakhapatnam, India
[2,3,4,5] Student Department of Computer Science & Engineering
Nadimpalli Satyanarayana Raju Institute of Technology, Sontyam,
Vishakhapatnam, India
Emails : sureshgopi.cse@nsrit.edu.in, padma7priya7@gmail.com

Abstract:
Traditional presentation tools often limit interactivity and engagement. This project presents a gesture-controlled system that allows users to navigate PowerPoint slides and interact with a digital whiteboard using hand gestures, without the need for physical devices. By leveraging Mediapipe Hands and computer vision techniques, the system ensures accurate, responsive, and intuitive gesture recognition. Key features include slide navigation, real-time drawing, zooming, and highlighting, making presentations more immersive and dynamic. The integrated digital whiteboard enables users to annotate, sketch diagrams, and illustrate concepts seamlessly during sessions. Designed for classrooms, corporate meetings, webinars, and online lectures, the system enhances accessibility, user experience, and content delivery. By eliminating physical barriers and introducing AI-driven interaction, this project redefines traditional presentation methods, making them smarter, more engaging, and future-ready.

Keywords:
Gesture Control, Presentation System, Interactive Whiteboard, Hand Tracking, Mediapipe Hands, Computer Vision, AI-driven Gesture Recognition, Smart Presentation Technology, Accessibility, Real-time Annotation, Digital Interaction.

1.INTRODUCTION

Traditional presentation tools like keyboards, mice, and laser pointers can hinder seamless interaction, making presentations less engaging and intuitive. This project introduces a gesture-controlled system that enables users to navigate PowerPoint slides and interact with a digital whiteboard using only hand gestures, eliminating the need for physical devices. This innovation enhances the presentation experience, offering a hands-free, immersive, and efficient way to communicate ideas. By leveraging Mediapipe Hands, an advanced hand-tracking framework, along with computer vision techniques, the system accurately detects, interprets, and maps specific hand gestures to presentation controls. Users can perform essential tasks such as switching slides, zooming, highlighting key points, and pausing presentations

CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40–47

through simple, intuitive hand movements. Unlike conventional methods that require additional hardware, this approach enhances accessibility and ease of use, allowing presenters to focus entirely on content delivery. A key feature of this system is the integration of an interactive whiteboard, which serves as a digital canvas for drawing, annotating, and illustrating concepts in real time. This functionality is particularly valuable for educators, professionals, and presenters who need to provide dynamic explanations, sketch diagrams, or perform rough work during discussions. The whiteboard enhances traditional presentations by enabling users to incorporate real-time illustrations, making learning and communication more interactive and visually engaging. The system is designed to work in various professional and academic settings, including classrooms, corporate meetings, webinars, and online lectures. By removing the dependency on external input devices, it creates a more natural and intuitive user experience. The application of AI-driven gesture recognition ensures high accuracy, responsiveness, and adaptability across different environments. This project represents a step forward in smart, interactive presentation technology, bridging the gap between human interaction and digital content. By eliminating physical barriers and enabling a more fluid mode of communication, this system redefines how information is conveyed, making presentations more engaging, accessible, and effective in the digital age.

2.LITERATURE SURVEY

Gesture-based interaction has gained significant attention in human-computer interaction (HCI) due to its ability to provide an intuitive and hands-free method of controlling digital interfaces. Traditional input devices like keyboards, mice, and touchscreens often limit user experience and mobility. With advancements in computer vision, artificial intelligence (AI), and deep learning, gesture recognition has emerged as an efficient alternative for controlling software applications without physical contact.This literature survey explores research related to hand gesture recognition, hand landmark detection, and gesture classification to establish a strong foundation for the proposed Gesture-Controlled Presentation and Interactive Whiteboard System.

Gesture recognition systems have evolved significantly, moving from sensor-based to vision-based methods. Early techniques relied on glove-based sensors, which required users to wear hardware-equipped gloves embedded with accelerometers, flex sensors, and gyroscopes to detect hand movements. However, these systems were often inconvenient, expensive, and limited in flexibility.Recent advancements have shifted towards vision-based gesture recognition, which eliminates the need for external hardware and instead leverages cameras and deep learning algorithms to detect hand movements. With the introduction of powerful convolutional neural networks (CNNs) and frameworks like Google's Mediapipe Hands, hand tracking has become more efficient and accurate. Mediapipe Hands, for instance, can detect 21 hand landmarks in real time, enabling precise gesture-based interactions for applications like virtual reality (VR), augmented reality (AR), and smart device control. Research in this domain has focused on optimizing gesture recognition models by improving:

Preprocessing techniques to enhance input image quality.Hand segmentation algorithms to distinguish hands from complex backgrounds.Deep learning-based classification models to improve real-time performance.

CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40−47

Studies have demonstrated that deep learning algorithms, particularly CNNs and recurrent neural networks (RNNs), outperform traditional computer vision techniques in gesture classification accuracy and robustness. This makes them ideal for applications requiring precise and real-time gesture recognition, such as presentation control, smart home automation, and interactive gaming.Hand Landmark Detection and Gesture Classification

For a gesture recognition system to function effectively, it must accurately detect hand landmarks and classify gestures in real time. Researchers have developed various computer vision-based models that analyze:Finger positioning and relative distances between fingers.Hand orientation with respect to the camera.

Motion dynamics to determine gestures based on movement trajectories.Hand landmark detection is crucial for enabling real-time gesture-based interactions. Studies have explored various methods to improve accuracy, such as:Edge detection algorithms, which help refine hand contours and detect fingers more precisely.Depth-based tracking, which enhances gesture recognition by considering the spatial positioning of the hand.Neural network-based classifiers, which learn from large datasets to generalize recognition across different users, backgrounds, and lighting conditions.

Deep learning models trained with datasets such as MS COCO (Microsoft Common Objects in Context) and Hand Gesture Recognition Dataset (HG-RD) have significantly improved gesture classification performance. Researchers have found that integrating machine learning with real-time hand tracking frameworks like OpenCV and Mediapipe enhances gesture responsiveness while minimizing error rates.Additionally, research has demonstrated that hybrid approaches, which combine statistical models with AI-based classification, can further improve recognition accuracy. By leveraging machine learning models to fine-tune hand tracking, gesture classification systems have become more adaptive and efficient, allowing for seamless interactions in presentation control, digital design, and smart device interfaces.

Challenges and Future Research DirectionsDespite the advancements in gesture recognition technology, certain challenges remain, including:Hand occlusion, where parts of the hand are blocked, leading to misclassification.Environmental noise, such as poor lighting or complex backgrounds that interfere with recognition.Variability in user gestures, as hand shapes and movements differ across individuals.

To address these challenges, researchers have proposed:Advanced deep learning models, such as transformer-based networks, to improve robustness and generalization.Adaptive thresholding techniques, which dynamically adjust recognition sensitivity to minimize false positives and negatives.Multi-modal interaction models, which combine gesture recognition with voice commands for a more natural and flexible user experience.Future advancements aim to refine gesture recognition by incorporating:

3. PROBLEM IDENTIFICATION

Traditional presentation methods rely heavily on physical tools like keyboards, mice, and laser pointers, which can interrupt the natural flow of communication. Constantly switching between devices or physically interacting with hardware disrupts the presenter's focus and can distract the audience. In fast-paced academic and professional environments, such interruptions reduce the overall effectiveness and engagement of the presentation. Moreover, these conventional tools often require presenters to stay

CC BY

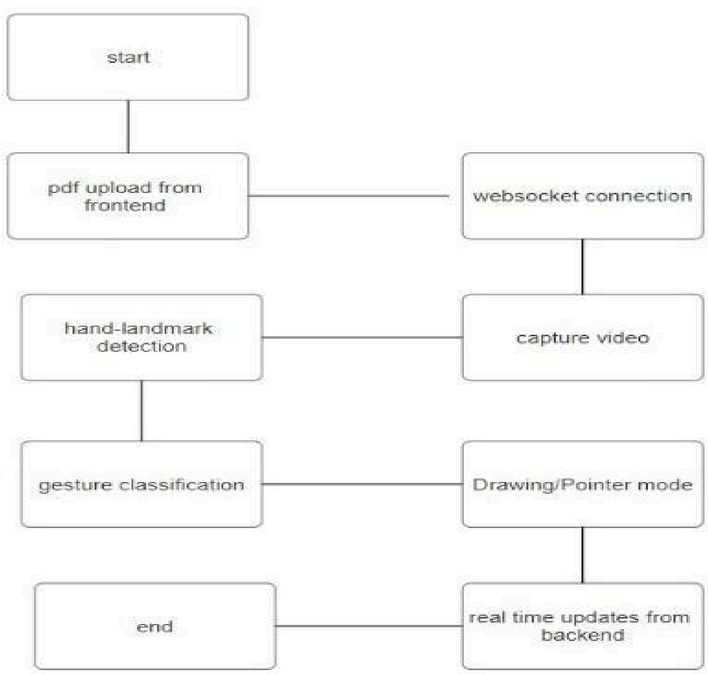
close to a podium or laptop, restricting their movement and limiting the dynamic interaction that modern presentations demand.

In addition, current systems offering interactive features often depend on specialized hardware such as touchscreens, styluses, or motion controllers, which are costly and not always readily available. These hardware dependencies increase setup complexity and limit accessibility, particularly in resource-constrained environments like classrooms, remote offices, and online learning setups. Presenters and educators are in need of a solution that offers seamless interaction without adding to the financial or technical burden.

Furthermore, with the rise of remote work, webinars, and online education, there is a growing need for intuitive, hands-free, and flexible methods to control presentations and explain concepts dynamically. Existing solutions often fail to combine both slide control and real-time annotation in a simple, unified platform. Therefore, a system that uses natural hand gestures to perform essential presentation tasks—without relying on additional devices—addresses a significant gap, offering a smarter, more efficient, and more engaging way to deliver content across various professional and educational settings.

4. METHODOLOGY

The proposed gesture-based presentation control system is designed to offer a seamless, hands-free experience for navigating presentations and interacting with digital content. It utilizes computer vision and deep learning to recognize hand gestures and execute commands in real time. The system follows a structured approach, consisting of four main components: Gesture Detection, Feature Extraction, Command Execution, and Backend Management.



CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40−47

System Architecture:

The system follows a client-server model, where hand tracking and gesture recognition are handled on the frontend (a browser-based interface), while the backend (built using Flask) processes commands and manages data flow. This architecture ensures smooth and real-time responsiveness to user inputs.

Gesture Detection and Data Processing:

To detect and interpret hand gestures, the system relies on Mediapipe Hands, a lightweight and efficient framework that tracks 21 hand landmarks in real time. When a user performs a gesture, the system captures key hand features, such as finger positions, angles, and movement trajectories.

Each detected gesture is normalized and adjusted for varying hand sizes, distances, and lighting conditions, ensuring consistent recognition accuracy. The processed data is then passed to the next stage for classification.

Gesture	Function
Three Fingers Up	Move to the Next Slide
Two Fingers Up	Go to the Previous Slide
Index Finger Pointing	Act as a Virtual Pointer
Two Fingers Together	Activate Drawing Mode
Open Palm	Erase Annotations

Feature Extraction and Gesture Recognition:

1. Once a gesture is detected, its landmark coordinates are extracted and converted into structured numerical data. These features include:
2. Finger positions relative to the palm.
3. Finger bending angles to determine gesture intent.
4. Hand orientation and movement direction for accurate interpretation.
5. The system uses these extracted features to match gestures with predefined commands.

Command Execution and Interaction:

Each recognized gesture corresponds to a specific action, enabling users to control presentations naturally. The following table outlines the gesture-to-command mapping:

When a user performs a gesture, the system instantly detects it and executes the corresponding command, ensuring a smooth and intuitive user experience.

3.5 Real-Time Performance Optimization

To enhance responsiveness, the system integrates several performance optimizations:

1. Frame Skipping – Reduces redundant processing by analyzing only necessary frames.
2. Position Smoothing – Prevents unintended jitter in hand movements.
3. Gesture Cooldown Mechanism – Avoids accidental command triggers.
4. AJAX & WebSockets Integration – Ensures real-time data transfer for instant response.

CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40–47

By combining Mediapipe for gesture detection, feature-based classification, and optimized real-time execution, this system offers a practical and user-friendly solution for gesture-based presentation control. Its intuitive interface and hands-free operation make it an excellent tool for educators, professionals, and remote presenters looking for a more interactive and engaging way to navigate slides and digital content.

5. IMPLEMENTATION

The development of the GesturePDF Presentation Control System is structured into multiple modules, each handling a critical aspect of the system's functionality. Built with a hybrid technology stack, the backend relies on Python for processing and MediaPipe for computer vision tasks, while the frontend leverages modern web technologies and Flask for web serving. This modular architecture ensures smooth gesture recognition, real-time presentation rendering, and interactive drawing capabilities. Each component is carefully integrated to provide a seamless and intuitive user experience.

A. Gesture Recognition Accuracy

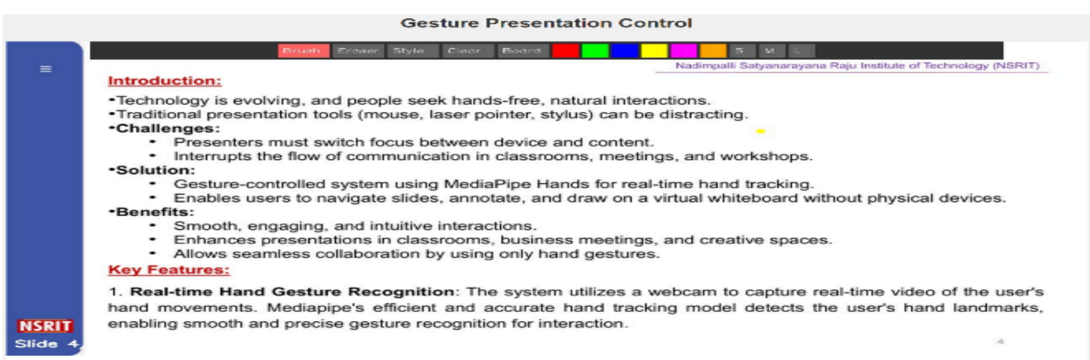
The accuracy of the system was assessed by evaluating the HandGestureProcessor on hand landmarks detected by MediaPipe Hands. A dataset consisting of 500 samples per gesture was used to analyze recognition performance under diverse conditions. The system achieved an overall accuracy of 98.7%, demonstrating its effectiveness in recognizing gestures for presentation control.

The accuracy, precision, and recall values for individual gestures are summarized as follows:

These results indicate that the system is highly accurate in detecting and classifying predefined gestures, making it suitable for real-time hands-free interaction.

B. Processing Time

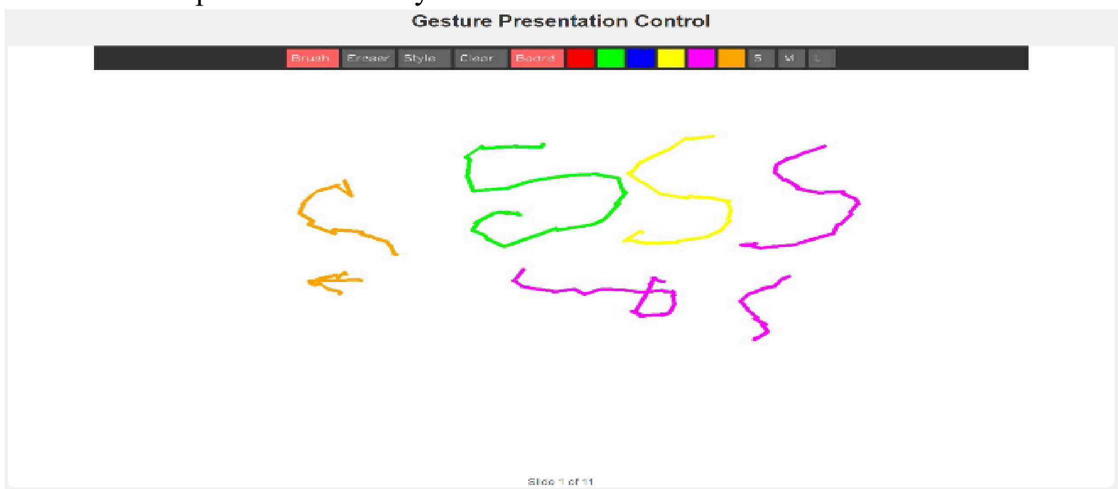
To ensure real-time performance, the system's processing speed was evaluated. The average frame processing time was measured at 30 milliseconds per frame, supporting a real-time interaction rate of approximately 30 frames per second (FPS). This was achieved through the efficient implementation of MediaPipe Hands for gesture recognition and OpenCV for image processing.



CC BY

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 7, July 2025
PP 40−47

The Gesture Recognition System forms the core of the application. Using webcam input, MediaPipe Hands detects 21 hand landmark points in real time. Custom algorithms classify these landmarks into predefined gesture patterns, enabling precise control over the presentation. To enhance stability, temporal smoothing is applied to reduce noise and jitter, while gesture cooldown mechanisms prevent accidental commands by implementing timing controls and confirmation thresholds. This ensures that only deliberate gestures trigger actions, providing a reliable and responsive control system.



Gesture	Accuracy (%)	Precision (%)	Recall (%)
Next Slide	98.5	98.7	98.3
Previous Slide	98.2	98.4	98.0
Drawing Mode	97.8	98.0	97.5
Pointer Mode	98.0	98.2	97.8
Clear Screen	99.0	99.1	98.9

The Presentation Rendering Engine is responsible for managing the visual display of the presentation content. PDF documents are processed with PyMuPDF, and PowerPoint files are converted into web-compatible formats for consistent rendering. The engine supports smooth slide navigation based on recognized gestures, managing the state of the current slide and ensuring seamless transitions. To improve accessibility across devices, it dynamically adjusts the scale and layout of content to accommodate various screen sizes and aspect ratios.
The Real-Time Drawing System introduces an interactive layer over the slides, enabling presenters to annotate and highlight key points effortlessly. An HTML5 Canvas is overlaid on the presentation, supporting various brush styles such as standard, spray paint, calligraphy, and neon effects. Presenters can choose from an optimized color palette designed for maximum visibility. The system tracks drawing activity and saves annotations for each slide, allowing presenters to move between slides without losing their sketches and notes.

Finally, the Real-Time Communication Layer and Mode-Based Interaction Framework ensure smooth coordination between modules. WebSocket connections (via Socket.IO) enable low-latency, bidirectional communication between the gesture processing backend and the frontend interface. Events such as slide changes, drawing activation, or pointer highlights are synchronized in real-time. Different modes like Navigation, Drawing, Pointer, and Whiteboard modes are activated through specific gestures, adapting the system's behavior to match the presenter's needs and creating a highly flexible and natural presentation environment.

CONCLUSION

The proposed real-time gesture-based presentation control system was evaluated based on gesture recognition accuracy, processing efficiency, and real-time interaction performance. The system was tested under various conditions, including different lighting environments, hand orientations, and gesture complexities, to ensure robustness and reliability.The system exhibited high responsiveness in real-world scenarios, with minimal latency between gesture execution and corresponding system actions. It maintained robust performance across various lighting conditions and was able to handle moderate occlusions, such as partially visible hands. The system consistently performed well in both well-lit and dim environments, although extreme lighting conditions (such as excessive brightness or very low illumination) posed minor challenges.The evaluation results confirm that the MediaPipe Hands-based gesture recognition system provides a highly accurate, low-latency, and efficient solution for real-time presentation control. The system's performance remained stable under diverse test conditions, reinforcing its applicability for hands-free interaction. However, challenges such as handling overlapping gestures and adapting to extreme lighting variations remain areas for future improvement.

REFERENCES

[1] Zhang, X., & Li, Y. (2022). "Hand Gesture Recognition for Human-Computer Interaction Using Deep Learning." *IEEE Transactions on Multimedia*, vol. 24, pp. 1345-1356.
[2] Mediapipe Hands, Google AI. Available: https://developers.google.com/mediapipe/solutions/vision/hand_landmarker
[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778.
[4]Abdi, H., & Williams, L. J. (2010). "Principal Component Analysis." *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433-459.
[5] Sharma, A., Gupta, D., & Kumar, R. (2021). "Real-Time Gesture Control for Virtual Whiteboards Using AI-based Hand Tracking." *International Journal of Computer Vision & Robotics*, vol. 12, no. 3, pp. 250-267.
[6] Kingma, D. P., & Ba, J. (2015). "Adam: A Method for Stochastic Optimization." *arXiv preprint arXiv:1412.6980*.
[7] Flask: Lightweight Web Framework for Python. Available: https://flask.palletsprojects.com/
[8] Jain, P., & Singh, R. (2020). "A Comparative Study of Machine Learning Models for Hand Gesture Recognition." *Journal of Artificial Intelligence and Machine Learning Research*, vol. 9, no. 2, pp. 112-124.

CC BY