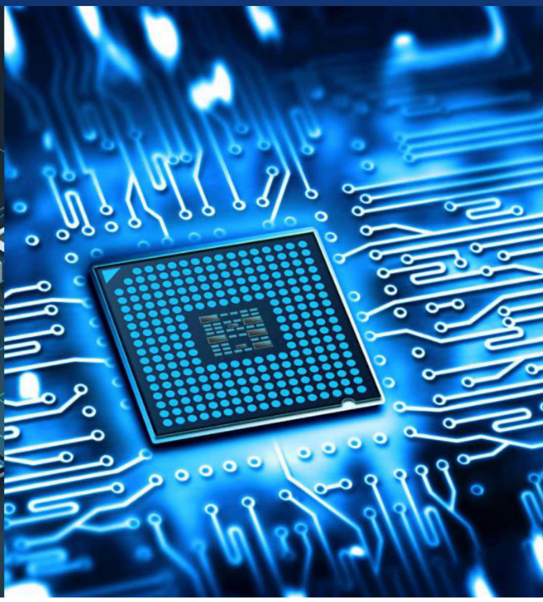
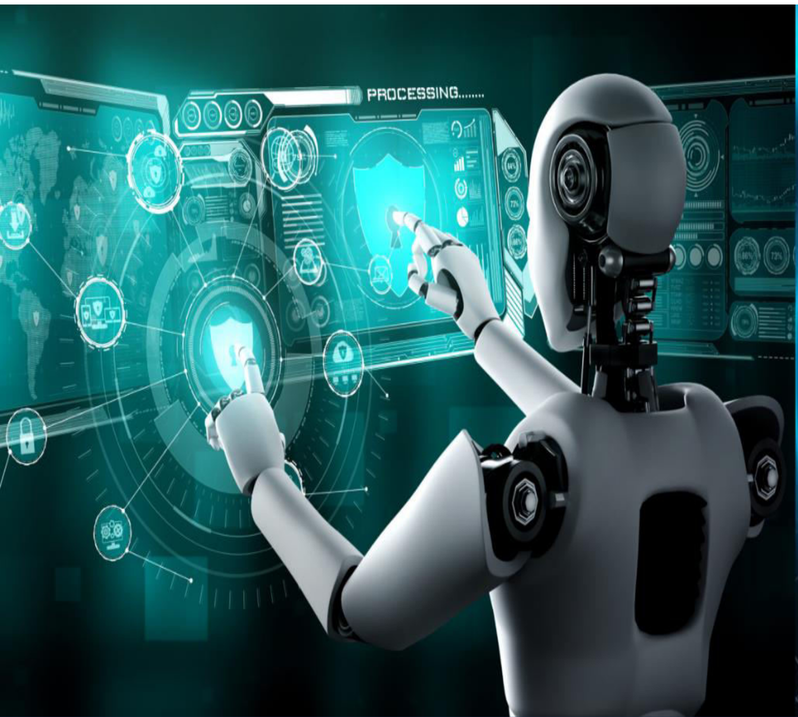
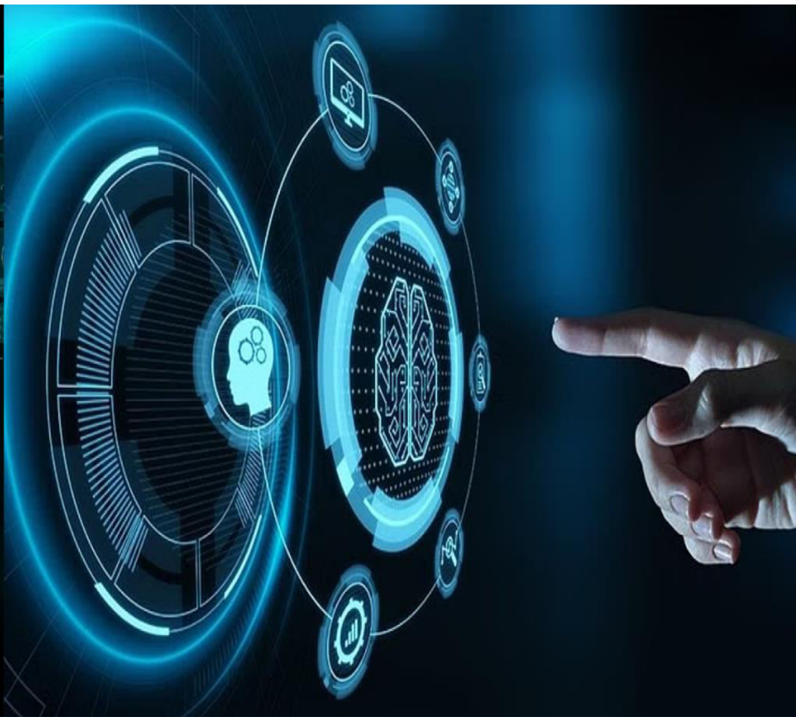
ISSN(O): 2320-9801
ISSN(P): 2320-9798


# International Journal of Innovative Research in Computer and Communication Engineering

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





Impact Factor: 9.927

Volume 14, Issue 6, June 2026

www.ijircce.com | ijircce@gmail.com | +91-9940572462 | +91 63819 07438

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

# Hand Gesture Control Presentation System: A Computer Vision Framework for Touchless Slide Navigation

**Dr. T.V.S Sriram[1], D. Appalaraju[2], S. Tirupathi Rao[3]**

Sr. Asst. Professor, Dept. of MCA, NSRIT, Visakhapatnam, AP, India[1]

Asst. Professor, Dept. Of MCA, NSRIT, Visakhapatnam, AP, India[2]

Dept. of MCA, NSRIT, Visakhapatnam, AP, India[3]

**ABSTRACT:** Conventional presentation systems depend on physical input devices such as keyboards, mice, or remote clickers, which can restrict a presenter's mobility and reduce natural interaction with an audience [1][2]. Hand gesture recognition offers a touchless and intuitive alternative by allowing presenters to control slides directly through hand movements captured by a standard webcam. This paper presents a Hand Gesture Control Presentation System that applies computer vision and machine learning techniques to detect, track, and classify hand gestures in real time. The system uses a hand-landmark detection pipeline [3][4] to extract twenty-one key points per hand, from which geometric features such as finger angles and inter-joint distances are computed. These features are passed to a trained classifier that recognizes gestures including swipe right, swipe left, pointer mode, and session-end, which are mapped to corresponding presentation commands such as next slide, previous slide, cursor control, and session termination. The proposed framework removes the dependency on physical remotes, improves accessibility, and provides a smooth, interactive presentation experience. Experimental evaluation shows that the system achieves reliable real-time gesture recognition with low latency under normal lighting conditions.

**KEYWORDS:** Hand Gesture Recognition, Computer Vision, Human-Computer Interaction, Presentation Control, Machine Learning, Gesture Classification, Touchless Interface.

## I. INTRODUCTION

Presentations are a central part of academic, professional, and corporate communication. Traditional control mechanisms such as keyboards, computer mice, and handheld remote clickers, while functional, tie the presenter to a fixed location or a specific device and can disrupt the natural flow of a talk. Hand gesture recognition introduces a more natural and engaging mode of human-computer interaction, enabling presenters to move freely while controlling their slides [5][6].

The Hand Gesture Control Presentation System proposed in this paper allows users to control slide navigation and an on-screen pointer using predefined hand gestures captured through a standard webcam, without requiring any specialized hardware such as gloves, depth sensors, or infrared cameras.

Recent advances in real-time hand-tracking frameworks [3][7] and lightweight machine learning classifiers [8][9][10] have made it possible to build accurate, low-latency gesture recognition systems that run on conventional consumer hardware. This paper proposes a complete pipeline that captures webcam video, detects and tracks hand landmarks, extracts discriminative features, classifies the performed gesture, and maps the recognized gesture to a corresponding presentation action.

## II. LITERATURE REVIEW

Hand gesture-based human-computer interaction has been an active research area for over two decades, driven by the goal of replacing physical input devices with natural, touchless interaction systems relied on colored markers or

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

instrumented gloves to simplify hand localization, but these approaches limited usability outside controlled environments [13].

With the growth of vision-based methods, researchers have applied classical machine learning algorithms including Support Vector Machine (SVM), Random Forest, and Artificial Neural Networks (ANN) [8][9][14] to classify static and dynamic hand gestures from camera input. Convolutional Neural Networks (CNN) have further improved recognition accuracy for complex gesture vocabularies by learning spatial features directly from image data [10][15].

A significant development in this field has been the introduction of real-time, pre-trained hand-landmark detection frameworks [3][7], which estimate twenty-one three-dimensional key points per hand directly from RGB video without requiring depth sensors. Such frameworks substantially reduce the engineering effort needed to build gesture-controlled applications and have been adopted in several presentation-control and human-computer interaction studies [16][17][18].

Despite these advances, many existing systems remain sensitive to variable lighting conditions, cluttered backgrounds, and gesture misclassification, and some require significant computational resources that limit real-time responsiveness [19][20]. There is therefore a continued need for lightweight, accurate, and user-friendly gesture-controlled presentation systems suitable for everyday use on standard laptops and webcams. The present work builds upon these findings to develop a practical hand gesture control system optimized for accuracy, responsiveness, and ease of use.

## III. SYSTEM OVERVIEW

The Hand Gesture Control Presentation System is a real-time, vision-based application designed to allow presenters to control slide navigation and an on-screen pointer using natural hand gestures captured by a webcam. The system continuously processes incoming video frames to detect the presenter's hand, track relevant landmarks, and classify the performed gesture.

The framework consists of five major stages: input acquisition, hand detection and tracking, feature extraction and classification, action mapping, and presentation output. After a gesture is recognized with sufficient confidence, the system translates it into a corresponding control action such as advancing to the next slide, returning to the previous slide, activating a virtual pointer, or ending the presentation session.

The system is designed to operate reliably under normal indoor lighting using a standard laptop webcam, without requiring specialized sensors, gloves, or markers, making it accessible and easy to deploy in classroom, seminar, and corporate presentation settings.

## IV. METHODOLOGY

The development of the Hand Gesture Control Presentation System follows a structured pipeline to ensure accurate and responsive gesture-based control. The methodology consists of the following phases.

### 4.1 Video Input Acquisition

Live video frames are captured continuously from a standard webcam using a video capture interface. Frame resolution and frame rate are configured to balance recognition accuracy with real-time responsiveness.

### 4.2 Hand Detection and Landmark Tracking

Each captured frame is processed using a real-time hand-tracking pipeline [3][7] that detects the presence of a hand and extracts twenty-one landmark coordinates corresponding to the wrist and finger joints. These landmarks are tracked across consecutive frames to maintain stable detection during motion.

### 4.3 Feature Extraction

Geometric features such as inter-finger distances, joint angles, and the relative position and orientation of the hand are computed from the landmark coordinates. These features form a compact representation that is invariant to minor changes in hand position and camera distance.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

4.4 Gesture Classification

The extracted feature vector is passed to a trained classifier. Multiple machine learning algorithms, including Artificial Neural Network (ANN), Support Vector Machine (SVM), and Random Forest [8][9][14], are evaluated, and the best-performing model is selected to classify the gesture into one of the predefined categories.

4.5 Action Mapping

Each recognized gesture is mapped to a specific presentation command using a rule-based lookup table. Recognized gestures are smoothed and debounced across consecutive frames to avoid unintended repeated triggers from a single gesture.

4.6 Performance Evaluation

The system's performance is evaluated using metrics such as recognition accuracy, precision, recall, F1-score, and response latency, to determine the reliability and real-time suitability of the proposed framework.

## V. SYSTEM ARCHITECTURE

The architecture of the Hand Gesture Control Presentation System follows a layered design that enables efficient, real-time processing of webcam video into presentation control commands.

The system consists of five major layers: the Input Acquisition Layer, the Hand Detection and Tracking Layer, the Feature Extraction and Classification Layer, the Action Mapping and Control Layer, and the Presentation Output Layer. The Input Acquisition Layer captures and normalizes live video frames from the webcam. The Hand Detection and Tracking Layer identifies the hand region and extracts landmark coordinates. The Feature Extraction and Classification Layer converts landmarks into feature vectors and classifies the performed gesture. The Action Mapping and Control Layer translates the recognized gesture into an operating-system-level command. Finally, the Presentation Output Layer reflects the resulting action on the presentation software and provides live feedback to the user.

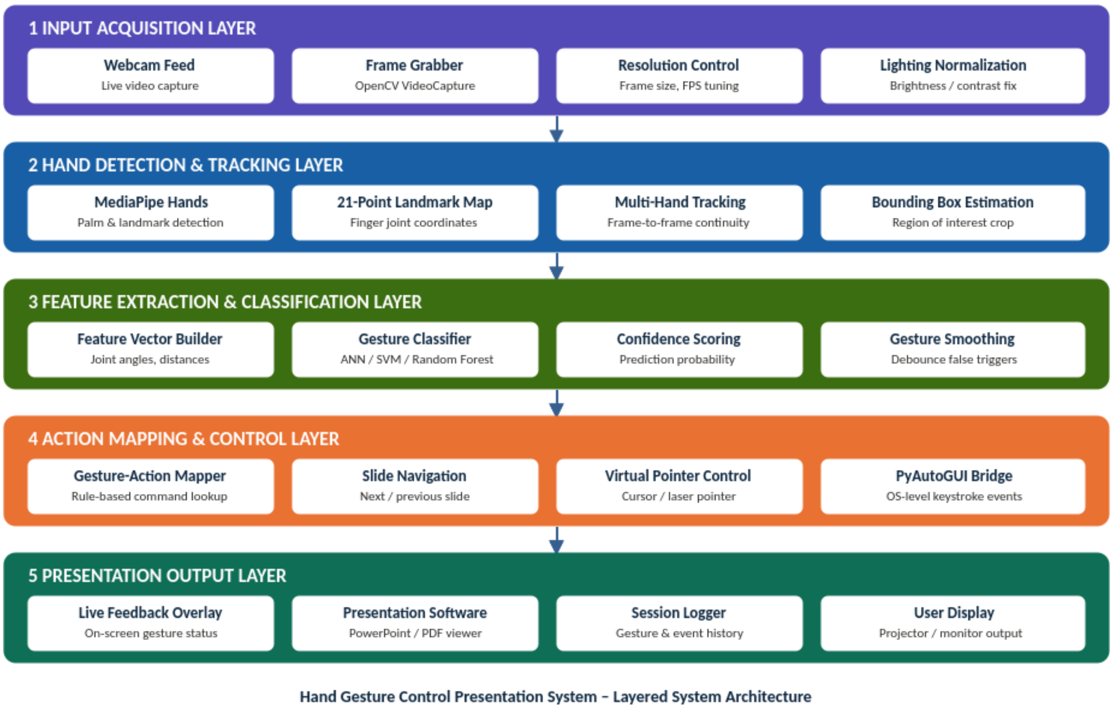



**Fig. 1: Hand Gesture Control Presentation System – Layered System Architecture**

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406097

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

## VI. MODULES

**User Module:** In this module
User can register and login into the system User can start and control presentations
User can use hand gestures to perform actions User can view the live hand gesture detection feed User can navigate slides (next / previous)
User can control the on-screen pointer position User can view system feedback and gesture status

**Feature Selection Module:** In this module:
Gesture capture and detection from webcam frames Feature extraction and processing (joint angles, distances) Selection of relevant features for gesture classification

**System Module:** In this module:
Secure storage of user and session data Logging of gesture and presentation activity
Interface management with presentation software

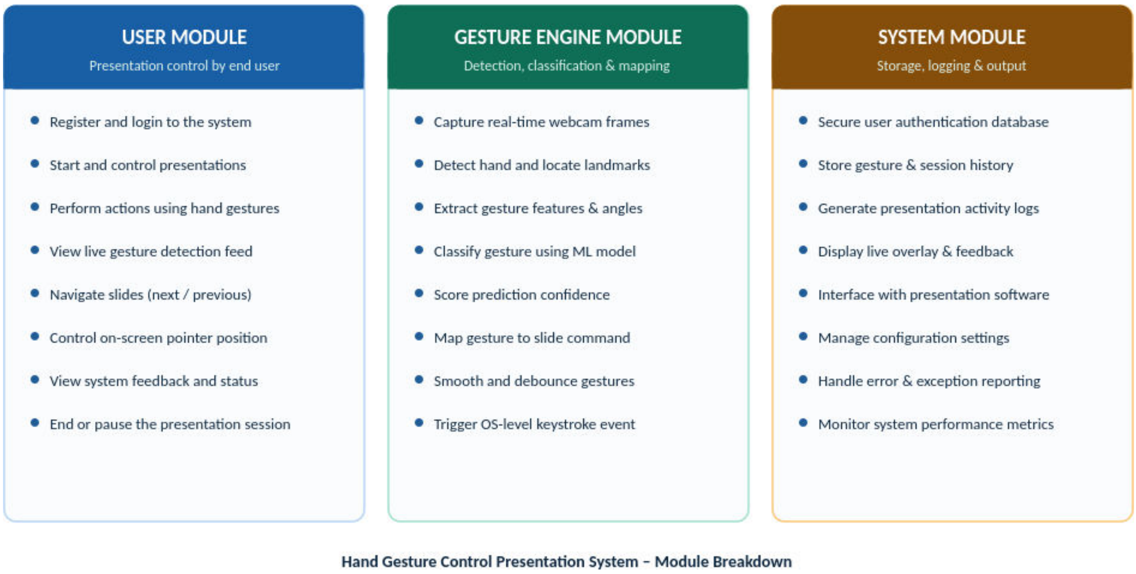



**Fig. 2: Hand Gesture Control Presentation System – Module Breakdown**

## VII. IMPLEMENTATION DETAILS

7.1 Technology Stack
The proposed system is developed using Python as the programming language. Hand detection and landmark tracking are implemented using a real-time hand-tracking framework [3][7], while video capture and image processing are performed using OpenCV [21]. Machine learning classifiers are implemented using Scikit-learn [22], and NumPy [23] is used for numerical computation. PyAutoGUI is used to translate recognized gestures into operating-system-level keystroke events.

7.2 Frontend Implementation
The user interface displays the live webcam feed with an overlay showing detected hand landmarks and the currently recognized gesture, providing the presenter with immediate visual feedback during operation.

7.3 Backend Implementation
The backend manages frame capture, landmark extraction, feature computation, gesture classification, and the mapping of recognized gestures to operating-system-level commands that control the presentation software.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

7.4 Gesture Vocabulary Design

A compact set of gestures was designed to minimize misclassification: an open palm swiped right triggers the next slide, an open palm swiped left triggers the previous slide, a single extended index finger activates pointer mode, and a closed fist held briefly ends the session.

7.5 Model Implementation

Machine learning algorithms such as ANN, SVM, and Random Forest [8][9][14] are trained on the extracted gesture feature vectors. The best-performing model is selected for real-time gesture classification based on validation accuracy and inference speed.

7.6 Reliability Measures

Gesture smoothing, temporal debouncing, and confidence thresholding are implemented to reduce false triggers and ensure that presentation commands are issued only when a gesture is recognized consistently across multiple consecutive frames.

## VIII. EXPERIMENTAL DETAILS AND RESULTS

8.1 Experimental Setup

The proposed system was implemented and tested using Python, OpenCV, and Scikit-learn. Experiments were conducted on a system with an Intel Core i5 processor, 8GB RAM, a standard laptop webcam, and the Windows operating system, under normal indoor lighting conditions.

8.2 Test Results

Several test cases were executed to validate the functional correctness and reliability of the system. The results are summarized in Table I.

| Test Case | Description | Result | Status |
|---|---|---|---|
| User Authentication | Login / Registration with valid and invalid credentials | Passed | ✓ |
| Webcam Initialization | Detect and initialize webcam video stream | Passed | ✓ |
| Hand Detection | Detect presence of hand and extract landmarks | Passed | ✓ |
| Gesture Classification | Classify swipe right, swipe left, point, fist gestures | Passed | ✓ |
| Next Slide Action | Trigger next-slide command on swipe-right gesture | Passed | ✓ |
| Previous Slide Action | Trigger previous-slide command on swipe-left gesture | Passed | ✓ |
| Pointer Control | Activate and move on-screen pointer using index gesture | Passed | ✓ |
| Session End | End presentation session on closed-fist gesture | Passed | ✓ |
| Response Time | Gesture-to-action latency within acceptable range | Passed | ✓ |
| Low-Light Performance | Gesture recognition under dim lighting conditions | Partial | ~ |

**TABLE I. TEST CASE RESULTS FOR HAND GESTURE CONTROL PRESENTATION SYSTEM**

8.3 Gesture Recognition Pipeline

Figure 3 illustrates the end-to-end gesture recognition pipeline, from webcam frame capture through landmark detection, classification, and final action execution, along with the gesture vocabulary recognized by the system.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

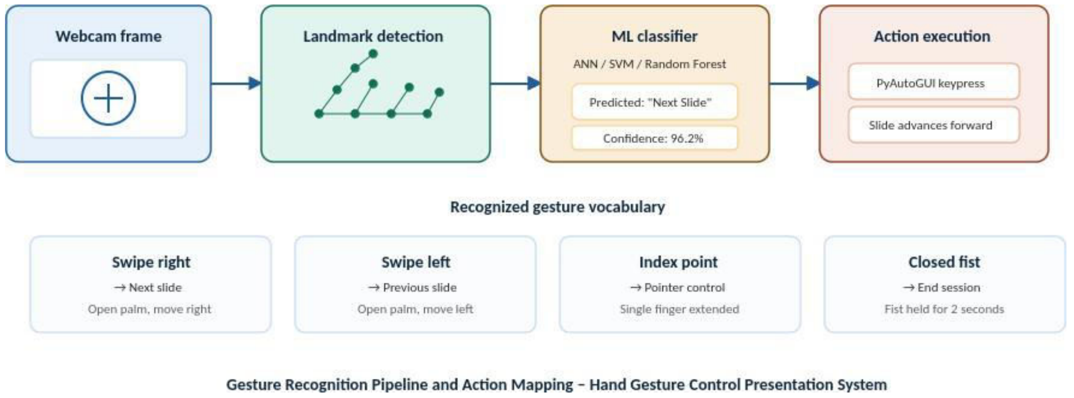



**Fig. 3: Gesture Recognition Pipeline and Action Mapping**

8.4 Performance Evaluation

System performance was evaluated using accuracy, precision, recall, and F1-score, in addition to gesture-to-action response latency. Experimental results indicate that the proposed framework recognizes the defined gesture vocabulary reliably under normal lighting, with response times suitable for live presentation use.

8.5 Observations

The system successfully classified all four target gestures and translated them into the correct presentation commands during testing. Recognition accuracy decreased slightly under low-light conditions and at greater distances from the camera, consistent with findings reported in related vision-based gesture recognition studies.

## IX. APPLICATIONS

- **Touchless Presentation Control:**

The system allows presenters to navigate slides and control a pointer without any physical device, improving mobility and audience engagement.
- **Classroom and Seminar Use:**

Teachers and speakers can control lecture slides naturally while moving around the room or facing the audience.
- **Accessibility Support:**

Gesture-based control can assist users who find physical remotes or keyboards difficult to operate.
- **Remote and Hybrid Presentations:**

The system can be integrated into video conferencing setups to support touchless slide control during remote talks.
- **Smart Classroom and Meeting Rooms:**

The framework can be extended to control other in-room systems such as displays and lighting using gestures.
- **Interactive Kiosks and Exhibits:**

The same gesture vocabulary can be adapted for public information kiosks and museum exhibits requiring touchless interaction.

## X. LIMITATIONS

- **Lighting Sensitivity:**

Recognition accuracy decreases under poor or uneven lighting conditions, as hand landmark detection relies on visible contrast.
- **Background Interference:**

Cluttered or high-motion backgrounds may occasionally cause false hand detections.
- **Limited Gesture Vocabulary:**

The current system supports a small, fixed set of gestures, which may limit functionality for more complex control needs.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

- **Camera Distance Constraints:**

Recognition reliability decreases significantly when the presenter moves far from the webcam.

- **Single-User Assumption:**

The system is designed primarily for single-presenter scenarios and may behave unpredictably if multiple hands appear in frame.

- **Hardware Dependency:**

Performance depends on webcam quality and processor speed, which may vary across different presentation setups.

## XI. FUTURE ENHANCEMENT

The future scope of the Hand Gesture Control Presentation System focuses on improving robustness, expanding functionality, and enhancing the overall user experience. Advanced deep learning architectures such as Convolutional Neural Networks and Long Short-Term Memory networks [10][24] can be integrated to recognize a larger and more dynamic gesture vocabulary. Adaptive lighting normalization and background subtraction techniques can be applied to improve performance in challenging environments. Multi-hand and multi-user support can be added to extend the system to collaborative presentation scenarios. The framework can also be integrated with voice commands for a multimodal control interface, and extended to support additional applications such as document scrolling, video playback control, and smart-room automation.

## XII. CONCLUSION

The proposed Hand Gesture Control Presentation System provides an effective, touchless approach to controlling presentations using computer vision and machine learning techniques. By detecting hand landmarks, extracting geometric features, and classifying gestures in real time, the system enables presenters to navigate slides and control an on-screen pointer without relying on physical input devices. Compared with traditional remote-based presentation control, the proposed system offers a more natural, accessible, and engaging method of interaction. Experimental results demonstrate that the framework reliably recognizes the defined gesture vocabulary under normal conditions with response times suitable for live use. With further improvements in robustness, gesture vocabulary, and multimodal integration, the Hand Gesture Control Presentation System can become a practical and widely deployable tool for touchless human-computer interaction.

## REFERENCES

[1] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 37, no. 3, pp. 311-324, 2007.
[2] R. S. Pressman and B. R. Maxim, Software Engineering: A Practitioner's Approach, 8th ed., McGraw-Hill Education, 2015.
[3] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.
[4] F. Zhang et al., "MediaPipe Hands: On-device Real-time Hand Tracking," arXiv preprint arXiv:2006.10214, 2020.
[5] M. Turk, "Gesture Recognition," in Handbook of Virtual Environment Technology, Lawrence Erlbaum Associates, 2002.
[6] S. Rautaray and A. Agrawal, "Vision Based Hand Gesture Recognition for Human Computer Interaction: A Survey," Artificial Intelligence Review, vol. 43, no. 1, pp. 1-54, 2015.
[7] Google AI, "Hand Landmarks Detection Guide," MediaPipe Solutions Documentation, 2023.
[8] C. Cortes and V. Vapnik, "Support Vector Networks," Machine Learning, vol. 20, pp. 273-297, 1995.
[9] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.
[11] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. MacDorman, "Review of Constraints on Vision-Based Gesture Recognition for Human-Computer Interaction," IET Computer Vision, vol. 12, no. 1, pp. 3-15, 2018.
[12] M. Bhuiyan and R. Picking, "A Gesture Controlled User Interface for Inclusive Design and Evaluative Study of Its Usability," Journal of Software Engineering and Applications, vol. 4, no. 9, pp. 513-521, 2011.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

[13] J. L. Hernandez-Rebollar, R. W. Lindeman, and N. Kyriakopoulos, "A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation," Proc. IEEE ICMI, pp. 185-190, 2002.
[14] T. Mitchell, Machine Learning, McGraw-Hill, 1997.
[15] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.
[16] H. Khanum and H. B. Pramod, "Smart Presentation Control by Hand Gestures Using Computer Vision and Google's MediaPipe," International Research Journal of Engineering and Technology, vol. 9, no. 7, 2022.
[17] M. Paulson, N. P. R., S. Davis, and S. Varma, "Smart Presentation Using Gesture Recognition," International Journal for Research Trends and Innovation, vol. 2, no. 3, 2017.
[18] D. R. Jadhav and L. M. R. J. Lobo, "Navigation of PowerPoint Using Hand Gestures," International Journal of Science and Research, vol. 4, no. 1, 2015.
[19] S. Bhisikar et al., "Hand Gesture Recognition for Human-Computer Interaction," International Journal of Computer Applications, vol. 158, no. 3, 2017.
[20] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
[21] G. Bradski and A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, O'Reilly Media, 2008.
[22] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.
[23] C. R. Harris et al., "Array Programming with NumPy," Nature, vol. 585, pp. 357-362, 2020.
[24] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[25] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
[26] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[27] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2017.
[28] R. Y. Wang and J. Popovic, "Real-Time Hand-Tracking with a Color Glove," ACM Transactions on Graphics, vol. 28, no. 3, 2009.
[29] J. Suarez and R. R. Murphy, "Hand Gesture Recognition with Depth Images: A Review," Proc. IEEE RO-MAN, pp. 411-417, 2012.
[30] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor," IEEE Transactions on Multimedia, vol. 15, no. 5, pp. 1110-1120, 2013.
[31] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed., O'Reilly Media, 2022.
[32] S. Raschka and V. Mirjalili, Python Machine Learning, 3rd ed., Packt Publishing, 2019.
[33] W. McKinney, Python for Data Analysis, 3rd ed., O'Reilly Media, 2022.
[34] I. Sommerville, Software Engineering, 10th ed., Pearson Education, 2016.
[35] M. Z. Uddin, C. Boletsis, and P. Rudshavn, "Real-Time Norwegian Sign Language Recognition Using MediaPipe and LSTM," Multimodal Technologies and Interaction, vol. 9, no. 3, pp. 1-15, 2025.
[36] E. Fertl, E. Castillo, G. Stettinger, M. P. Cuéllar, and D. P. Morales, "Hand Gesture Recognition on Edge Devices: Sensor Technologies, Algorithms, and Processing Hardware," Sensors, vol. 25, no. 6, pp. 1-46, 2025.
[37] V. Bhat and B. S. Panchami, "Hand Gesture-Based Mouse Control System Using OpenCV and MediaPipe for Real-Time Interaction," IRE Journals, vol. 8, no. 6, pp. 862-867, 2024.
[38] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
[39] S. Haykin, Neural Networks and Learning Machines, 3rd ed., Pearson, 2009.
[40] A. Esteva et al., "A Guide to Deep Learning in Healthcare," Nature Medicine, vol. 25, no. 1, pp. 24-29, 2019.



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com