

Hands-Free Human–Computer Interaction for Smarter Workspaces

Prerana Balip^{1,*}, Dnyaneshwar Sonavane², Shaila Pawar³, Omkar Shinde⁴, Ayush Deoghare⁵, Poonam Potraje⁶

Abstract

In current digital work environments, frequent interaction with keyboards and mouse devices can interrupt workflow, especially during presentations, collaborative sessions, and multitasking activities. This study introduces a touchless human–computer interaction system that allows users to operate computer functions through hand gestures and voice instructions, minimizing the need for traditional input devices such as keyboards and mice. The proposed system integrates gesture recognition, speech processing, and optical character recognition (OCR) technologies to support a variety of interactive tasks such as slide navigation, virtual drawing, handwriting capture, and text extraction from visual input. A voice-based assistant is incorporated to execute system-level operations and respond to user commands efficiently, while a vision-based module performs real-time hand tracking and gesture detection for seamless interaction. The system is designed to improve usability, accessibility, and interaction efficiency in environments where touch-free operation is beneficial, including smart classrooms, meetings, and presentation settings. By combining computer vision and voice interaction techniques, the proposed model offers a more intuitive and flexible method of human–computer communication. Experimental results indicate that the system operates efficiently in controlled environments and can carry out various functions with high accuracy while requiring very little physical effort from the user. The project also provides scope for future enhancement in dynamic environments through improved gesture robustness, adaptive learning, and advanced AI-based interaction capabilities.

Keywords: Gesture recognition, voice assistant, human–computer interaction, virtual assistant, speech recognition

INTRODUCTION

Human–computer interaction has evolved significantly with the integration of natural input methods, such as voice and gestures [1, 2]. Traditional interactions using keyboards and mouse devices can be restrictive, particularly in dynamic environments such as presentations and meetings [3, 4]. Recent studies have explored gesture-based systems and virtual assistants to improve interaction efficiency [5–

7]. The proposed system focuses on enabling hands-free interaction by combining gesture recognition and voice commands. It allows users to perform tasks such as controlling presentations, interacting with a virtual whiteboard, and executing system commands without physical contact [3, 6, 8]. This approach improves accessibility and supports more intuitive interactions in modern smart workspace environments [9, 10].

*Author for Correspondence

Prerana Balip
E-mail: preranasbalip@acpce.ac.in

^{1, 2, 4, 5}Student, Department of Information Technology, A.C. Patil College of Engineering, Kharghar, Maharashtra, India

^{3,6}Assistant Professor, Department of Information Technology, A.C. Patil College of Engineering, Kharghar, Maharashtra, India

Received Date: May 08, 2026

Accepted Date: June 04, 2026

Published Date: June 09, 2026

Citation: Prerana Balip, Dnyaneshwar Sonavane, Shaila Pawar, Omkar Shinde, Ayush Deoghare, Poonam Potraje. Hands-Free Human-Computer Interaction for Smarter Workspaces. Journal of Advancements in Robotics. 2026; 13(2):50–56p.

PROBLEM STATEMENT AND SIGNIFICANCE

Problem Statement

In traditional work environments, users depend heavily on physical input devices such as

keyboards, mouse devices, and remotes to operate digital systems [1]. These tools can become limiting, disruptive, or inaccessible during real-time presentations, collaborative sessions, or hands-busy situations [3, 7], leading to reduced efficiency and friction in productivity when quick and intuitive interactions are required. In addition, the continuous dependence on manual input methods reduces interaction flexibility and limits the development of seamless, touchless communication systems suitable for modern smart workspace environments [5, 6, 10]. Existing interaction methods also provide limited support for multimodal communication, where gesture recognition and voice-based control can work together to improve accessibility, user convenience, and real-time task management in intelligent workspace applications [5, 7, 8].

Significance

- It improves accessibility, especially for people with mobility issues or in situations where using hands is difficult [2, 10].
- Boosts productivity by enabling faster, more intuitive interaction and reducing dependency on continuous switching between input devices and workspaces [5, 3, 8].
- This aligns with current trends in smart workspaces and intelligent meeting environments that emphasize automated and seamless interactions [6, 2].
- It enhances hygiene and safety by reducing physical contact with shared surfaces in public or health-sensitive environments [9, 10].

RELATED WORK

Gesture Recognition System [1]

This study focused on the development of a gesture and posture recognition framework intended to improve nonverbal human–computer interaction. The authors implemented an SVD–PCA-based feature extraction approach combined with a feed-forward neural network for classifying six gesture and six posture categories. The system demonstrated satisfactory recognition accuracy under controlled conditions and highlighted the feasibility of real-time gesture interpretation using computer-vision techniques. However, this work was limited by a small gesture dataset and lacked integration with application-level controls or multimodal interaction mechanisms. The paper is relevant to the proposed research as it establishes a foundational approach for gesture-based interaction systems used in hand-free workspace environments.

Desktop's Virtual Assistant Using Python [5]

A Python-based desktop virtual assistant was designed to automate routine computer operations using speech-driven interactions. The authors employed speech recognition, text-to-speech conversion, and natural language processing modules to execute commands such as opening applications, retrieving information, and performing system-level tasks. The implementation emphasized low-cost software integration using Python libraries, including pyttsx3, speech recognition, and web automation modules. Experimental observations indicated improved task automation and user convenience for desktop environments. Nevertheless, the assistant primarily relies on voice input and lacks contextual gesture interaction or adaptive learning capability. The work directly supports the proposed research by demonstrating how voice-controlled assistance can be integrated into a broader hands-free human–computer interaction framework.

A Gesture-Controlled Whiteboard System with OpenCV, MediaPipe and NumPy [6]

This study aimed to design a virtual whiteboard system capable of enabling touchless drawing and writing through hand gesture recognition. The proposed framework utilized OpenCV for image processing, MediaPipe for real-time hand landmark detection, and NumPy for coordinate computation and drawing operations. The system successfully allows users to perform writing, drawing, and gesture-based interactions using a webcam interface without requiring specialized hardware. This study demonstrated the effectiveness of computer vision techniques in educational and collaborative applications. However, performance limitations were observed under varying lighting conditions and

complex backgrounds. This paper is highly relevant to the proposed research because it validates the implementation of virtual drawing and gesture-controlled interaction within a hands-free workspace environment.

A Review of the Hand Gesture Recognition System: Current Progress and Future Directions [2]

This review examines the evolution of hand gesture recognition systems and analyzes existing methodologies, sensing technologies, and application domains. The authors discussed vision-based and sensor-based gesture recognition approaches, including machine learning and deep learning techniques for feature extraction and classification. This study highlights the increasing adoption of gesture recognition in healthcare, robotics, education, and assistive technologies. Additionally, this study identified major challenges, such as environmental sensitivity, computational complexity, and limited robustness under dynamic conditions. Although the work was theoretical and lacked implementation-level validation, it provided a comprehensive understanding of the current research directions. The survey is relevant to the proposed work because it establishes technical background and identifies research gaps associated with real-time gesture-based interaction systems.

Virtual Assistant in Native Language [7]

This study proposed a virtual assistant capable of processing commands in a native language environment to improve accessibility and user interaction. The system incorporates speech recognition, language processing, and voice response mechanisms to support localized communication between users and computing systems. The authors focused on improving usability for non-English-speaking users through multilingual interaction and speech-based automation. Experimental analysis demonstrated effective command recognition and response generation in the selected language domain. However, its implementation is limited in terms of scalability and lacks integration with gesture-based interfaces or multimodal interaction techniques. The study is relevant to the proposed research because it highlights the significance of voice-driven interaction and accessibility enhancement in developing intelligent hands-free human–computer interaction systems.

METHODOLOGY

System Architecture

The Nova AI Assistant processes user input through preprocessing and intent recognition using natural language processing (NLP) techniques [5, 7]. Based on the identified intent, appropriate responses are generated and displayed, enabling voice-based system interaction, as shown in Figure 1. The assistant also supports task automation, application control, and interactive command execution, thereby improving user convenience and reducing dependency on traditional manual input methods in workspace environments [5, 7].

The gesture-controlled whiteboard processes hand movements using real-time tracking to detect shapes and user inputs through computer vision and hand landmark detection techniques [6, 2]. Recognized gestures are rendered on the canvas, enabling intuitive drawing and interaction, as shown in Figure 2. The system also supports touchless writing and dynamic user interaction, improving collaboration efficiency and reducing dependence on conventional input devices in smart workspace environments [6, 10].

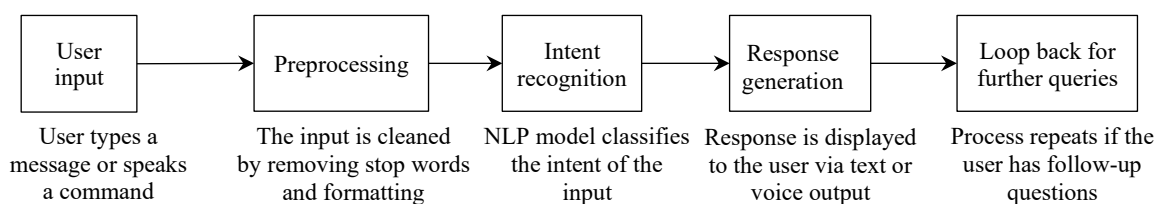


Figure 1. Nova AI assistant workflow.

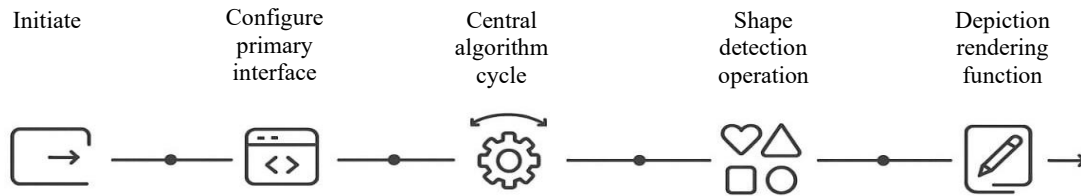


Figure 2. Gesture-controlled whiteboard workflow.

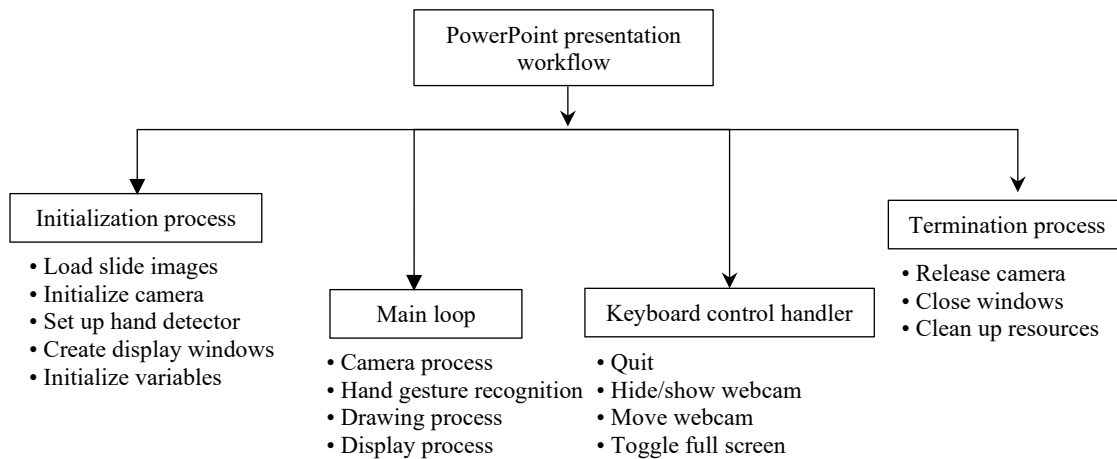


Figure 3. PowerPoint presentation workflow.

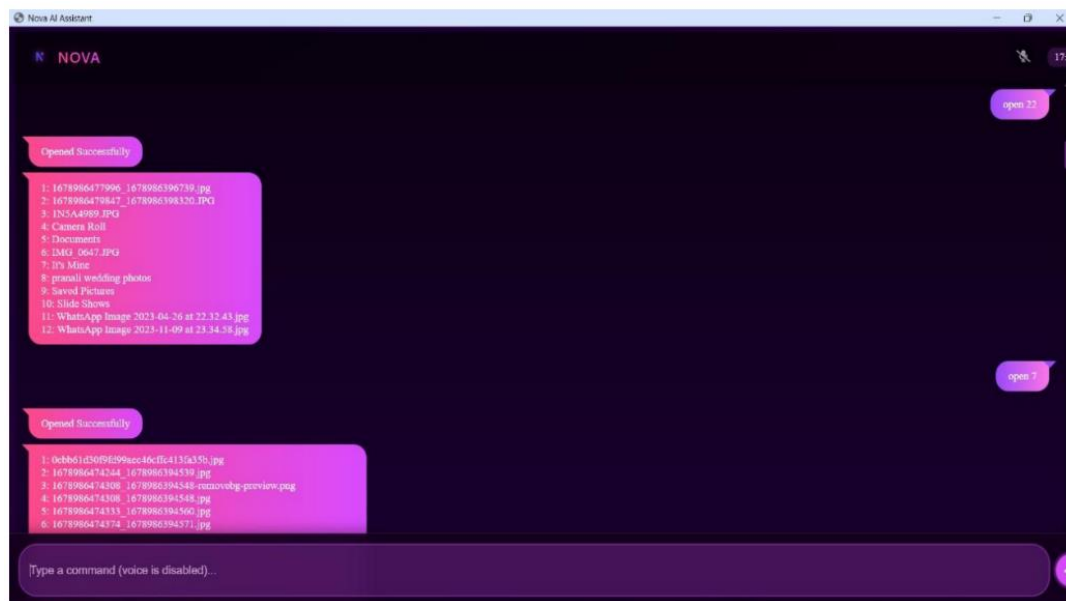


Figure 4. Voice-based file control using Nova AI assistant.

RESULT AND DISCUSSION

The presentation module processes camera input to detect hand movements for slide navigation and control using computer vision and gesture recognition techniques [3, 4, 8]. The identified actions were mapped to system commands to enable hands-free presentation interaction, as shown in Figure 3.

The Nova AI Assistant processes voice commands to retrieve and interact with the system files. It first generates an indexed list of files and directories, enabling navigation using spoken input. Based on the selected command, the system maps the input to the corresponding file and executes system-level operations to open it, allowing for efficient hands-free file control, as shown in Figure 4.

The assistant interprets voice input as a search or location-based query and dynamically generates the corresponding results [11]. It constructs appropriate URLs to perform web searches or retrieve map-based information, enabling real-time, hands-free access to information and geolocation services, as shown in Figure 5.

The system utilizes camera-based hand tracking to detect finger movements and interprets them as alphanumeric inputs. Hand landmarks are analyzed in real time to convert gestures into corresponding characters, enabling intuitive and hands-free text input, as shown in Figure 6.

The virtual painter captures hand movements to identify predefined geometric shapes using contour detection and gesture-tracking techniques. The recognized shapes are rendered digitally on the canvas, providing immediate visual feedback for user interaction, as shown in Figure 7.

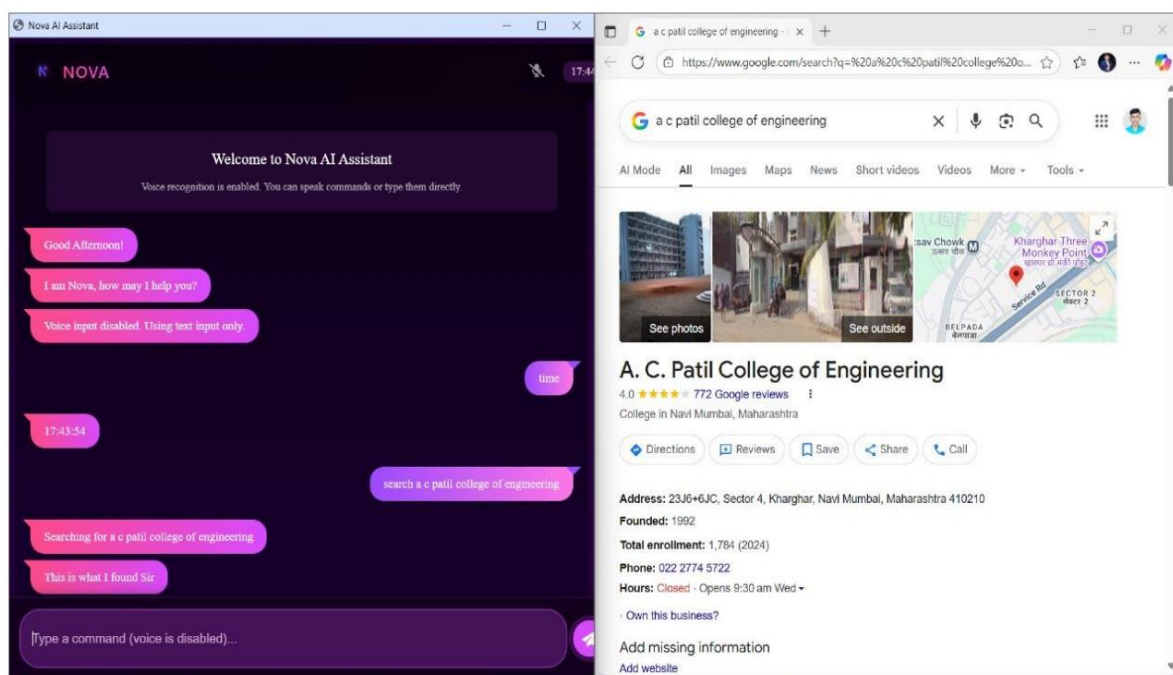


Figure 5. Voice-based search and location using Nova AI assistant.

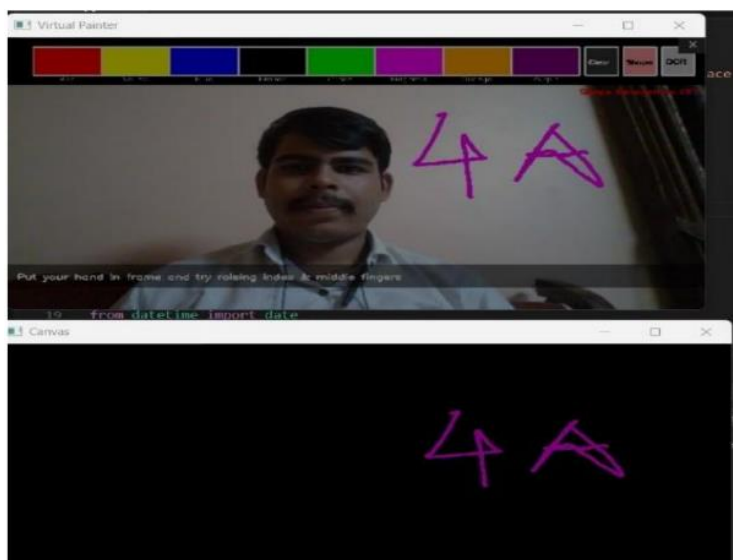


Figure 6. Alphanumeric recognition using virtual painters.

The system enables a virtual pointer using finger tracking, where fingertip coordinates are mapped to the screen positions. This allows real-time cursor control to highlight content during presentations without requiring physical input devices, as shown in Figure 8.

The system detects predefined hand gestures using finger landmark analysis to control the slide transitions. Recognized gestures are mapped to keyboard events, enabling hands-free navigation of presentation slides and improving interaction flow, as illustrated in Figure 9.

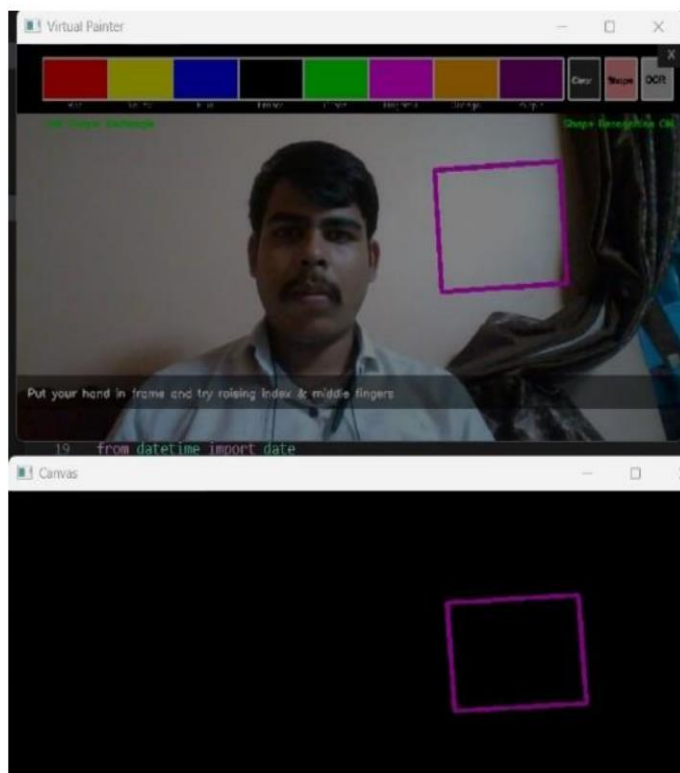


Figure 7. Shape recognition using virtual painter.

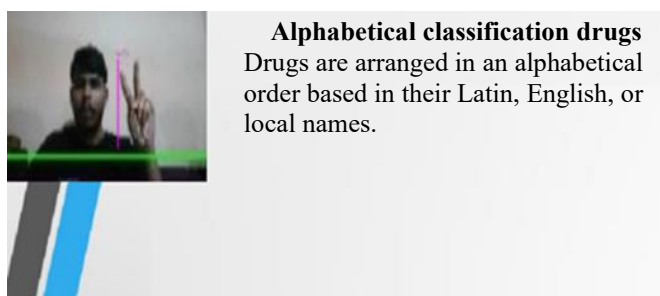


Figure 8. Pointer gesture.

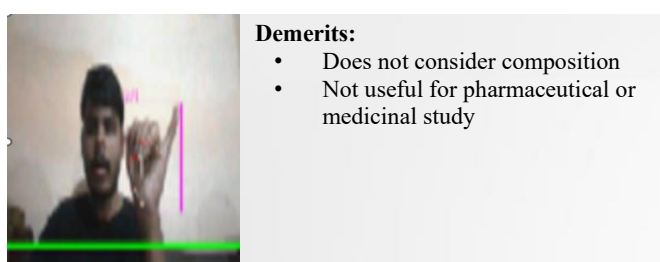


Figure 9. Gesture to go to the next slide.

CONCLUSION

The Hands-Free human-computer interaction framework that combines gesture recognition, voice commands, and handwriting recognition delivers more natural and accessible control of digital tools, including slide presentations, virtual whiteboarding, and text capture. Experimental results demonstrate promising accuracy and responsiveness in controlled settings, with challenges remaining in handling varied lighting, background clutter, and simultaneous multimodal inputs. Further improvements can focus on expanding gesture and voice vocabularies, optimizing algorithms for lower latency, and enhancing robustness for diverse user environments. This study contributes to smarter and more inclusive workspaces where interaction adapts to human behavior rather than requiring users to adapt.

Acknowledgment

We thank our college and our Principal, Dr. S.R. Deore, for providing the required resources for the development of the project. We would also like to thank the HOD, Prof. Dr. Varsha Bhole, for suggesting such a great project topic for our departmental purpose. Our sincere thanks to our Project Guide, Prof. Shaila Pawar, for helping, suggesting new ideas, and guiding us throughout the semester. We are also grateful to all faculty members for their support and encouragement.

REFERENCES

1. Sharma P, Sharma N. Gesture recognition system. In: 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU); 2019; Ghaziabad, India. p. 1-3. doi:10.1109/IoT-SIU.2019.8777487.
2. Dias PM, Jayakody K. Virtual assistant in native language. In: 2020 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS); 2020. p. 16-18. doi:10.1109/AGERS51788.2020.9452751.
3. Pustode B, Pawar V, Pawar V, Pawar T, Pokale S. Smart presentation system using hand gestures [Preprint]. Research Square. 2023. doi:10.21203/rs.3.rs-2549833/v1.
4. Umaphathi N, Karthick G, Venkateswaran N, Jegadeesan R, Srinivas D. Desktop's virtual assistant using Python. Eur Chem Bull. 2023;12(S3):5975-5984. doi:10.31838/ecb/2023.12.s3.667.
5. Soroni F, Sajid SA, Bhuiyan MNH, Iqbal J, Khan MM. Hand gesture based virtual blackboard using webcam. In: 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON); 2021. p. 134-140. doi:10.1109/IEMCON53756.2021.9623181.
6. Singh S, Shukla H, Sharma K, Tyagi H, Prasad J. Digitized interaction: a gesture-controlled whiteboard system with OpenCV, MediaPipe and NumPy. IPEC J Sci Technol. 2023;2(1):43-49.
7. Idrees M, Ahmad A, Butt MA, Danish HM. Controlling PowerPoint using hand gestures in Python. Webology. 2021;18(6):1372-1388.
8. Powar S, Kadam S, Malage S, Shingane P. Automated digital presentation control using hand gesture technique. ITM Web Conf. 2022;44:03031. doi:10.1051/itmconf/20224403031.
9. Abid Siddique RV, Naik S. A survey on gesture control techniques for smart object interaction in disability support. Electronics. 2023;12:512.
10. Mohamed N, Mustafa MB, Jomhari N. A review of the hand gesture recognition system: current progress and future directions. IEEE Access. 2021;9:157422-157436. doi:10.1109/ACCESS.2021.3129650.
11. Prasetya DD, Ashar M. Design of interactive whiteboard to support e-learning. In: Proceedings of the 1st International Conference on Vocational Education and Training (ICOVET 2017); 2017 Nov 4-5; Malang, Indonesia. Paris: Atlantis Press; 2017. p. 121-124. doi:10.2991/icovet-17.2017.26.