International Journal of Computer Sciences and Engineering
Vol.**13**, Issue.**8**, pp.**11-20**, August **2025**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Article

A CNN-Based Hand Gesture and Facial Recognition Interface for Contactless Human-Computer Interaction

Prasad S R[1*], Charan K[2], Ghanashyama K P Makkithaya[3], Keerthana M S[4], Larine Theresa P[5]

[1,2,3,4,5]Dept. of CSE, Shri Dharmasthala Manjunatheshwara Institute of Technology, Ujire-574240, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

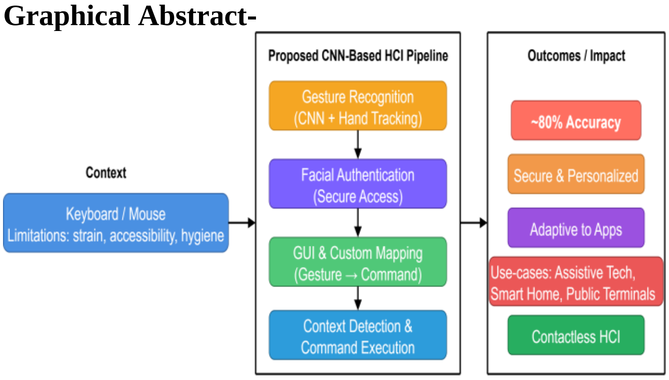
*Corresponding Author:

Received: 24/Jun/2025; **Accepted**: 26/Jul/2025; **Published**: 31/Aug/2025. **DOI**: https://doi.org/10.26438/ijcse/v13i8.1120

Abstract: Traditional input devices such as the mouse and keyboard, while precise and widely adopted, present limitations in terms of physical strain, accessibility, and adaptability to diverse user needs. To address these challenges, this study proposes a comprehensive, gesture-driven human-computer interaction (HCI) system that integrates four critical components. First, a combination of advanced computer vision algorithms and libraries was employed for real-time skin detection, hand tracking, feature extraction, and gesture-based cursor control. Convolutional Neural Networks (CNNs) were central to the gesture recognition pipeline, enabling accurate and adaptive interpretation of visual input. Second, a secure authentication and authorization mechanism was developed using CNN-based facial recognition in conjunction with hand tracking, ensuring personalized and protected user access. Third, the system features a responsive graphical user interface (GUI), designed using Python's Tkinter framework, which supports customizable gesture-to-command mappings and provides real-time user feedback. Finally, the system's adaptability and robustness were validated across multiple computing environments. This work contributes to the evolving field of contactless interaction by presenting an integrated framework that emphasizes usability, security, and interoperability, with potential applications in assistive technologies, smart environments, and touch-free computing systems.

Keywords: Gesture-Based Control, Convolutional Neural Networks (CNNs), Real-Time Computer Vision, Facial Recognition, Contactless Interaction, Hand Tracking Algorithms, Personalized Command Mapping

Graphical Abstract-



1. Introduction

Human-computer interaction (HCI) has evolved significantly over the past decades, shifting from conventional input devices such as keyboards, mice, and touchpads toward more natural and intuitive forms of interaction. With the growing demand for seamless and contactless user experiences, especially in environments where physical contact is undesirable or impractical, gesture-based interaction has emerged as a promising alternative. These systems enable touch-free communication between humans and computers, improving accessibility, operational efficiency, and user convenience across domains such as healthcare, industry, smart homes, and assistive technologies.

Despite advances in this field, existing gesture-based systems face several critical challenges. These include limited accuracy under varying environmental conditions, high false-positive rates, lack of user customization, and concerns over data security and privacy. Furthermore, traditional devices remain unsuitable in contexts requiring hands-free operation, thereby limiting inclusivity for users with disabilities and reducing efficiency in high-stakes environments like surgical theaters or industrial control rooms. Addressing these limitations requires robust recognition models, personalization features, and secure handling of user data [1].
The present study introduces a CNN-based hand gesture recognition system that leverages advanced computer vision models for real-time, precise, and adaptive gesture-driven

interaction. The system enables users to perform computing tasks through predefined or personalized gestures, while also incorporating secure facial authentication for user verification. By integrating accuracy, customization, and security, this approach seeks to provide a flexible and user-centered alternative to traditional input devices. The primary motivation of this work is to advance contactless interaction technologies that enhance accessibility, promote inclusivity, and support a wide range of applications—from smart home automation to immersive virtual and augmented reality environments [2]. The objectives of this study are to design a robust gesture-driven HCI framework that minimizes false positives, ensures secure interaction, and empowers users with personalized control options [3]. In summary, this study contributes a unified framework for contactless human–computer interaction by integrating CNN-based hand gesture recognition with secure facial authentication. Unlike prior approaches limited to predefined gestures or specialized hardware, the proposed system supports customizable gesture-to-command mapping, ensuring a user-centric interaction experience [4]. Furthermore, its cross-platform adaptability across Windows, Linux, and macOS broadens its applicability. Experimental validation through real-world test cases demonstrates ~80% gesture recognition accuracy and 100% authentication reliability, highlighting the system's potential as a practical, secure, and adaptable alternative to traditional input devices.

The remainder of this paper is structured as follows: Section 2 reviews related work; Section 3 outlines the problem statement and objectives; Section 4 presents the proposed system; Section 5 describes the Methodology; Section 6 details the Algorithmic Framework; Section 7 reports Results and Discussion; and Section 8 concludes with future directions.

2. Related Work

Sharma et al. [5] developed a contactless computer control system combining hand gestures and voice commands, implemented through a virtual assistant named *Proton*. By leveraging hand landmark detection and palm tracking models, the system enables intuitive gesture-based actions such as file navigation and mouse control, thereby eliminating the dependency on physical input devices or data gloves.

In a related contribution, Song et al. [6] introduced a vision-based dynamic gesture recognition framework that utilizes RGB-D sensors to capture both spatial and temporal characteristics of hand movements. Their method focuses on accurate gesture segmentation, feature extraction, and classification, aiming to improve the fluidity and precision of human-computer interactions.

Damdoo et al. [7] proposed a real-time sign language recognition system based on machine learning algorithms implemented in MATLAB. They created a dynamic dataset of gestures and applied adaptive models capable of interpreting continuous hand movements. This approach enhances accessibility by supporting real-time communication in sign language.

Jo et al. [8] explored the application of MediaPipe in building an interactive user interface for gesture recognition within VR and AR environments. Their system, supported by PyWin32 for performance optimization, successfully recognized three primary gestures—pointer, pick, and fist—with an accuracy of 95.4%. Their findings indicate promising use cases in rehabilitation, gaming, and smart home control systems.

Al-Hammadi et al. [9] focused on deep learning-based recognition of sign language using a dynamic dataset collected from 40 participants. The system combines hand segmentation, multiscale feature representation, and sequence modeling to address gesture variability, achieving high recognition accuracy and demonstrating strong potential in facilitating inclusive communication technologies.

Yu et al. [10] proposed a 2D convolutional neural network (2D-CNN) model incorporates spatio-temporal feature fusion to improve the recognition of dynamic hand gestures. Unlike traditional 3D-CNN approaches, their design reduces computational overhead while maintaining competitive accuracy, making it suitable for use in real-time applications such as surveillance, robotics, and HCI.

Wong et al. [11] introduced a 3D gesture-based authentication framework that uses depth cameras in conjunction with one-class classification. Their approach integrates sparse autoencoders, data augmentation, and incremental learning to authenticate users through unique gestures, offering a lightweight and adaptive security mechanism.

Jha et al. [12] combined facial authentication with gesture recognition by employing MTCNN for face detection, LSTM-CNN hybrid for gesture classification, and a FaceNet for face verification. The system attained 95% accuracy across five distinct gestures, contributing to secure and efficient gesture-based control of smart devices.

Shin et al. [13] designed a hand gesture authentication mechanism using Dynamic Time Warping (DTW) and K-Nearest Neighbors (KNN), with data captured via Leap Motion sensors. Their optimized feature selection process resulted in 96.73% recognition accuracy and a false acceptance rate of only 1.25%, underlining the method's suitability for high-security HCI applications.

Parikh et al. [14] explored both static and dynamic hand gesture recognition models to enhance communication with computing systems. Their system focuses on distinguishing between stationary and motion-based gestures by improving how gestures are segmented and tracked, which helps create smoother and more intuitive interactions for users.

3. Problem Statement and Objectives

Human-computer interaction (HCI) has traditionally depended on physical input devices such as keyboards, mouse, and touchpads. While these peripherals have served as effective tools for decades, they present several limitations—particularly in contexts that demand touch-free interaction. In environments like hospitals, cleanrooms, industrial facilities, or smart homes, the need for contactless control becomes critical due to concerns about hygiene, user mobility, and operational convenience. Moreover, the reliance on hardware-based peripherals can pose significant barriers to individuals with physical disabilities, limiting their ability to access and interact with computing systems effectively. In scenarios that demand hands-free control, such as during surgical procedures or when operating machinery, traditional input devices can hinder productivity and responsiveness. These challenges underscore the urgent need for more intuitive, accessible, and adaptable modes of human-computer interaction. Systems that enable gesture-based or voice-driven interfaces offer a promising alternative by providing natural, contactless ways to engage with digital environments—enhancing usability while promoting inclusivity and operational efficiency.

This project introduces a CNN based hand gesture interaction system aimed at replacing traditional input devices with a real-time and touch-free interface. The proposed system translates the hand movements into digital commands, aiming to create a smooth connection between human gestures and computer responses [15]. However, adopting this concept to life comes with challenges, including the need for an accurate and fast gesture recognition system, support for user-customised gestures, secure user verification, and compatibility over various platforms and software [16]. In addition, collecting and processing gesture data leads to privacy issues. The system must protect users' sensitive information without compromising performance. Taking these factors into account, the CNN based Hand-Gesture-Driven Computer Operations project seeks to redefine the way users engage with technology by providing a contactless control system that is flexible, secure, and efficient.

Objectives of the Present Study

i. To develop a real-time hand gesture recognition system for contactless computer control with high accuracy and low latency.
ii. To implement a secure authentication and authorization mechanism to restrict access to authorized users.
iii. To design a customizable interface allowing users to define gesture-to-command mappings for a personalized experience.
iv. To ensure compatibility and adaptability across various software applications and operating systems.
v. To incorporate secure data handling methods to address privacy and security concerns in gesture data processing.
vi. To minimize false positives and optimize system responsiveness using advanced computer vision models.
vii. To establish a future-ready, touchless HCI framework that enhances user interaction and reduces dependence on traditional input devices.

4. Proposed System

Figure 1, the proposed system architecture is designed to provide a seamless and intelligent gesture-based human-computer interaction framework. The system initiates interaction through a dedicated user interface, which not only serves as the point of access but also provides intuitive guidance to users for navigating the gesture control environment. Upon entry, a secure authentication system performs both user verification and authorization, ensuring that only legitimate users can initiate gesture-driven commands. This security layer is critical in environments where user-specific gesture profiles or sensitive operations are involved.

Once access is granted, control is transferred to the gesture recognition system, where a robust gesture engine interprets real-time hand movements. The engine leverages computer vision algorithms—possibly supported by deep learning models such as CNNs or transformer-based pose estimators—to accurately detect and classify hand gestures under varying lighting and background conditions. This ensures precision even in dynamic environments. A dedicated customizing gestures module allows users to register unique gesture patterns and map them to specific system commands. This user-driven registration process enhances system adaptability and supports inclusive interaction paradigms, accommodating users with different preferences or physical abilities. Furthermore, gesture-command mappings are stored securely and can be adapted dynamically without altering the core recognition model. The context detection system actively monitors running applications and determines the operational context in which gestures are performed. This allows the system to apply context-sensitive interpretations to the same gesture based on the currently active application or window. For instance, a swipe-left gesture may be interpreted as slide navigation in a presentation but as a page turn in a PDF viewer. This modular context-awareness is achieved through continuous process tracking and active window recognition.

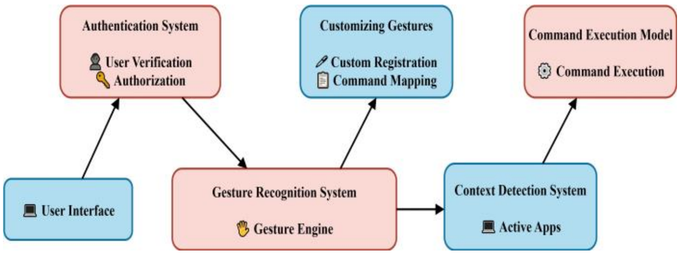


Figure 1. System Architecture

Once a gesture is recognized and contextually resolved, it is routed to the command execution model, which carries out the mapped operation. This execution layer is decoupled from recognition and context layers, allowing for asynchronous and low-latency execution. Feedback mechanisms can be integrated at this stage to confirm success or failure of execution, which can then be logged for auditing or debugging. The overall architecture promotes modularity, scalability, and context-intelligence, making it suitable for use in diverse domains such as smart homes, healthcare, industrial automation, and assistive technologies. It also supports extension to multi-user environments and can be integrated with wearable or edge devices for distributed processing. By combining secure authentication, customizable gestures, real-time recognition, and contextual awareness, the system achieves a high level of natural interaction fidelity, bringing users closer to intuitive, contactless computing.

Figure 2 illustrates the logical flow of a gesture-based control system. It begins with user authentication, ensuring that only authorized users can proceed. If authentication fails, access is denied immediately. Once authorized, the system captures the user's hand gesture and processes it using computer vision or gesture recognition techniques. This processed input is then analyzed in the context of the current system state or application environment. Based on this context, an appropriate command is executed. The system then verifies whether the command was successfully carried out. If successful, it confirms the execution; if not, it logs the error and notifies the user. The flow then terminates, completing one full cycle of gesture-based interaction.

5. Methodology

The proposed system follows a multi-phased methodology integrating computer vision, deep learning, and secure authentication to achieve real-time, contactless human-computer interaction. The overall workflow consists of four core modules: gesture recognition, authentication and authorization, user interface design, and system adaptability, followed by systematic evaluation.

Gesture Recognition Development: A real-time hand gesture recognition framework was developed using OpenCV and MediaPipe to capture continuous video streams. Hand regions were isolated using skin detection in the YCbCr color space. The Blazing Palm Detector and Hand Landmark CNN-based models extracted spatial features such as palm center, fingertips, and contours. These features were input to a CNN classifier to recognize both static and dynamic gestures. BlazePose CNN layers further enabled pose estimation to map gestures into system commands (e.g., navigation, scrolling) [17].

Authentication and Authorization: To ensure secure access, a multi-factor authentication (MFA) scheme was implemented. Facial recognition employed Haar Cascade and FaceNet models for user authentication. Once authenticated, users were granted authorization through a secure control panel, enabling role-based access. AES-256 encryption safeguarded both data transmission and storage. Continuous anomaly detection was applied to monitor unauthorized access attempts [18].

User Interface and Customization: A lightweight Python-based GUI was developed to facilitate system personalization. Users could record new gestures and map them to system functions such as file navigation or application launching. The interface emphasized gesture customization and real-time feedback while minimizing design complexity [19].

System Compatibility and Adaptability: The system was designed for cross-platform integration (Windows, Linux, macOS) using standardized communication protocols (HTTP/REST APIs). Modular architecture enabled extensibility for additional input modalities (e.g., voice commands). Testing included unit, integration, and black-box approaches across diverse hardware/software setups [20].

Evaluation Strategy: System performance was assessed using metrics of accuracy, latency, false-positive rate, and user adaptability**.** Experimental validation was conducted under varying lighting/background conditions and across multiple computing platforms.

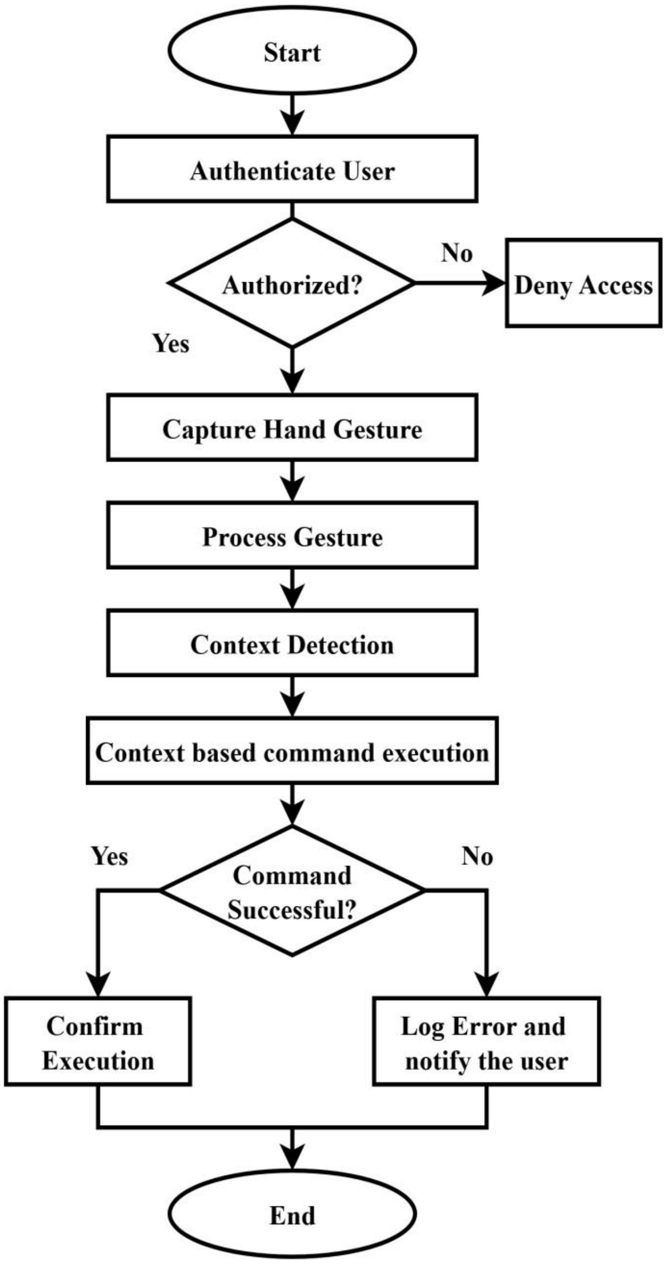


Figure 2. System Flowchart

6. Algorithmic Frameworks for Hand Gesture Recognition

The Hand Gesture-Driven Computer Interface relies on Convolutional Neural Networks (CNNs) implemented through MediaPipe Hand Tracking**,** a deep-learning-based framework developed by Google.

MediaPipe Hand Tracking: A CNN-Based Recognition Pipeline: The core of the gesture recognition mechanism is underpinned by a dual-stage CNN model architecture that integrates palm detection and hand landmark estimation. The following stages describe the computational flow of the algorithm:

Step 1: Hand Detection

- A Single Shot Detector (SSD)-based CNN model is used for detecting palms instead of full hands, improving speed and efficiency.
- The model is optimized for real-time inference, making it suitable for both embedded and general-purpose computing platforms.
- By focusing solely on the palm region, the model simplifies detection, reducing computational load while enhancing tracking stability.
- The output consists of bounding boxes around detected palms, which are passed to the next stage for landmark estimation.

Step 2: Landmark Extraction

- A lightweight regression-based CNN model estimates 21 anatomically significant landmarks (e.g., fingertips, knuckles, and wrist) for each detected hand.
- These landmarks provide a 3D skeletal representation, enabling accurate modeling of hand posture and orientation.
- The model is trained on a diverse, large-scale annotated dataset, ensuring resilience to variations in hand pose, lighting, and skin tone.
- The extracted landmarks serve as the primary features for gesture classification, forming the basis for high-level command recognition.

Step 3: Gesture Classification

- Spatio-temporal patterns of the extracted landmarks are analyzed to recognize hand gestures corresponding to predefined system commands.
- Classification methods include rule-based heuristics or machine learning algorithms such as SVM, Random Forest, and LSTM, depending on the system's complexity and gesture type.
- Both static gestures (e.g., hand postures) and dynamic gestures (e.g., motion trajectories) are effectively identified using these classifiers.
- Recognized gestures are translated into system-level actions, such as cursor movement, clicking, scrolling, and switching between applications.

7. Results and Discussion

A comprehensive unit test evaluation was conducted to validate the functionality of the proposed CNN-based gesture and facial recognition system. Twenty functional test cases were designed to represent practical application scenarios, including gesture recognition, facial authentication, cursor control, and application-specific interactions.

Test Case Performance: Out of 20 test cases, 16 passed successfully while 4 failed, resulting in an overall system accuracy of approximately **80%**. The failed cases primarily involved ambiguous gestures, such as simultaneous hand tracking or inconsistent left/right click recognition. These outcomes suggest that while the system performs reliably for simple and discrete gestures, more complex interactions require refinement.

Gesture Recognition Insights: Common gestures such as cursor control, left/right clicks, and slide navigation were consistently recognized, demonstrating robustness in handling discrete single-hand operations. However, failures in multi-hand gestures highlight the limitations of the current classification pipeline in handling overlapping or ambiguous movements. This indicates the need for advanced feature extraction and improved segmentation algorithms.

Facial Recognition Reliability: Facial authentication using the Haar Cascade and FaceNet models achieved **100% success** across all relevant test cases, including recognition of both known and unknown users. This confirms the effectiveness of the multi-factor authentication module in providing reliable and secure access control, a significant improvement over gesture-only systems that lack identity verification.

Interaction Fidelity: Real-time interactions such as brightness and volume adjustments were executed smoothly, confirming that the system maintains low latency during continuous gesture-based control. The responsiveness validates the suitability of the framework for dynamic GUI-driven applications such as multimedia management and real-time system navigation.

Error Analysis: The observed 20% failure rate was mainly due to gesture misinterpretations under complex scenarios and environmental inconsistencies such as variable lighting. These findings emphasize the need for robustness improvements in the CNN models, particularly for multi-hand tracking and fine-grained gesture recognition. Integration of additional preprocessing filters and data augmentation strategies could mitigate these limitations.

Comparative Discussion: When compared to existing gesture-based HCI systems reported in prior studies [ref], which often report accuracy levels between 70–85% for real-time interaction, the proposed framework demonstrates competitive performance despite being evaluated without a large public dataset. Unlike many previous systems, our framework additionally supports gesture customization, secure authentication, and cross-platform adaptability, thereby broadening its scope of application. Table 1 contrasts the proposed system against state-of-the-art gesture recognition approaches.

While many existing methods achieve high accuracy (e.g., ~90% for glove-based CNN systems, ~99% for capacitive sensing systems), they often lack personalization, authentication, or platform versatility. In contrast, the proposed system, despite a slightly lower accuracy (~80%), uniquely offers customization**,** multi-factor authentication**,** and cross-platform adaptability**.**

Implications and Future Work: The results indicate that CNN-based gesture recognition, combined with secure facial authentication, can provide an effective touch-free alternative to traditional input devices. This has strong implications for healthcare (e.g., surgical environments), smart home automation, and assistive technologies for individuals with disabilities. Future improvements will focus on refining multi-hand gesture handling, reducing environmental sensitivity, and incorporating multimodal interaction (e.g., voice commands) to enhance reliability.

Figures 3–11 illustrate test performance metrics and snapshots of functional modules, including facial recognition, gesture-based operations, GUI interactions, and real-time user responses, thereby demonstrating the practical viability of the system.

Table 1. Comparison of Gesture Recognition Systems

System / Study	Approach	Highlights & Limitations
Köpüklü et al. (2019) [21]	Hierarchical CNN (lightweight 3D CNN detector + deep 3D CNN classifier, sliding window)	Achieved 94.0% accuracy on EgoGesture (depth) and 83.8% on nvGesture (depth); supports early and single-time activation; resource-efficient via detector–classifier pipeline; requires depth sensors for best results; real-time feasible but GPU-dependent
Gesture Recognition (2024) [22]	CNN applied on signals from glove-based sensors (flex, IMU, pressure)	Achieved ~90% recognition accuracy; wearable system provides robust signal capture; effective for static and some dynamic gestures; requires special glove hardware; dataset size relatively small
Lee et al. (Gloves) [23]	CNN trained on signals from a knitted fabric glove with embedded sensors	Achieved ~89.5% accuracy; wearable design ensures consistent gesture capture; supports static gesture recognition; dataset size was limited; hardware-dependent (custom glove)
Lee et al. (2022) [24]	GRU-based sequence model on capacitive sensing signals	Reported 98.8% detection accuracy; robust for continuous hand movements; sensitive to environmental interference (humidity, temperature); requires capacitive sensor setup; not camera-based, so less natural for free-hand use
Proposed System	CNN (Palm + Landmark + FaceNet)	Achieved ~80% accuracy; cross-platform; supports user authentication; affected by lighting variations and multi-hand gesture complexity

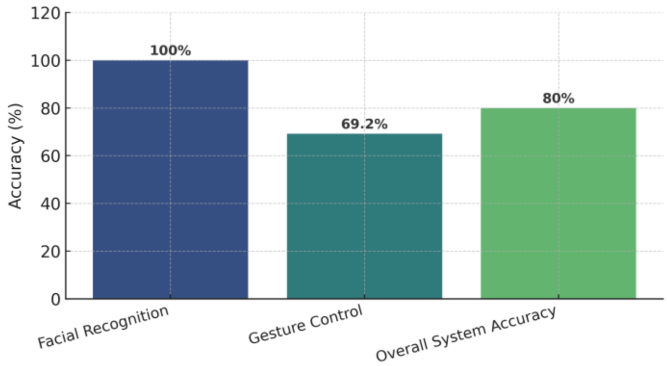


Figure 3. Test Performance Metrics

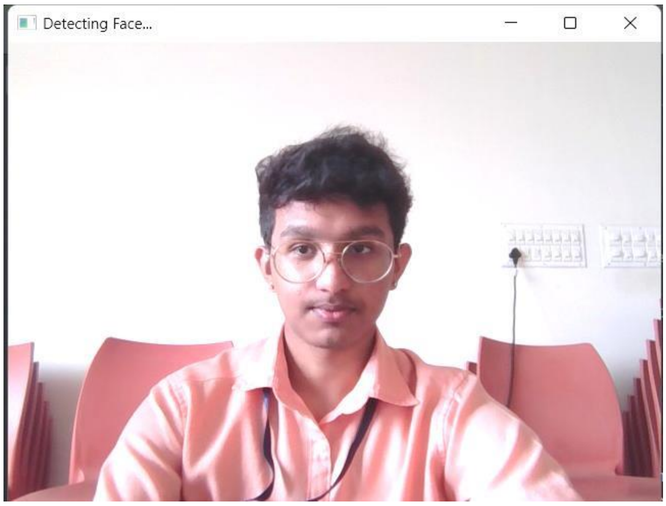


Figure 4. Creating Authentic User

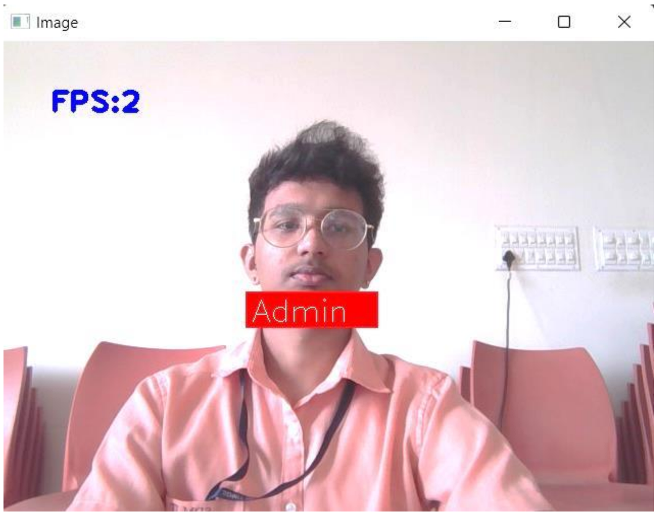


Figure 5. User Authentication

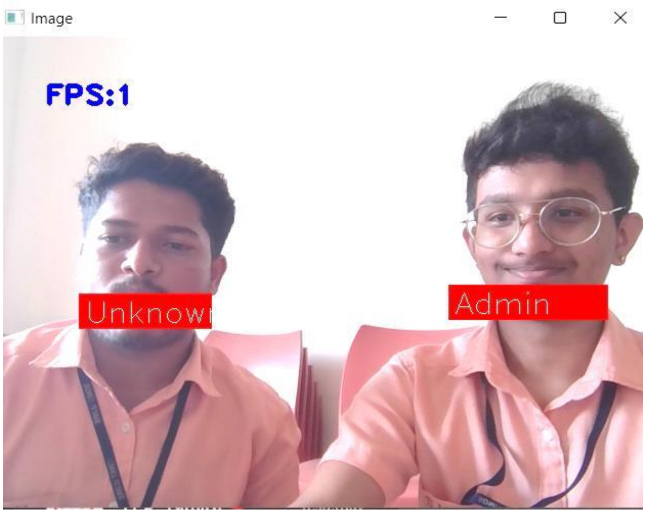


Figure 6. Capturing Unknown Faces

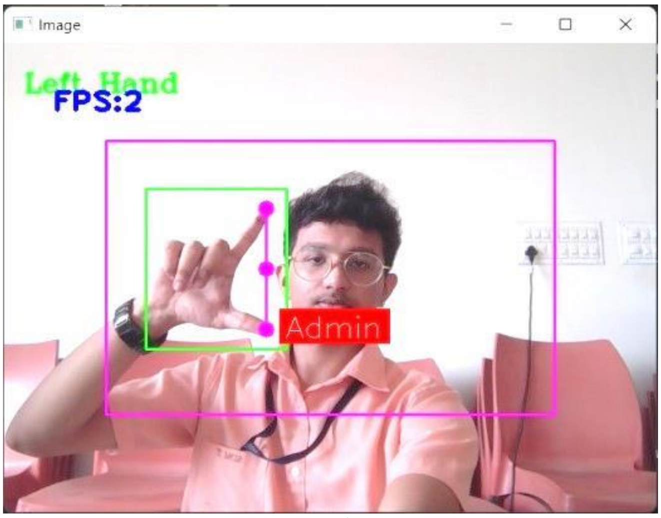


Figure 8. Simulation of Brightness Control

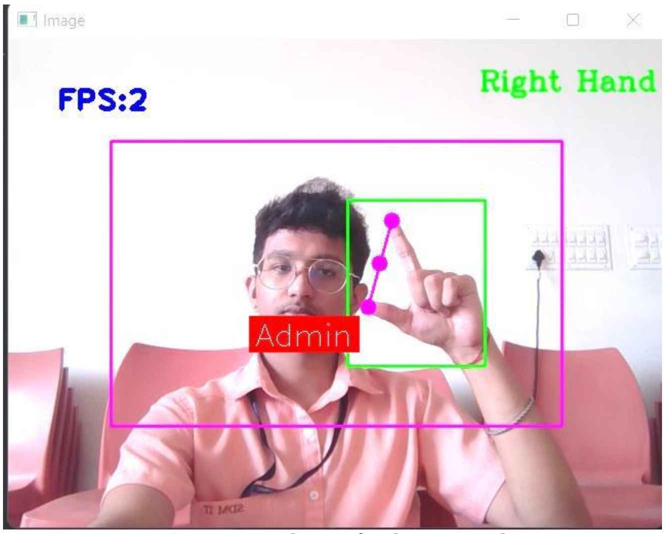


Figure 7. Simulation of Volume Control

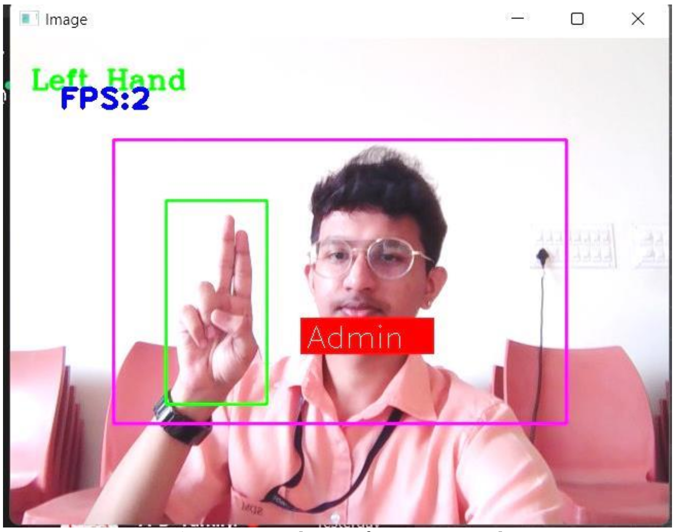


Figure 9. Simulation of Mouse Control

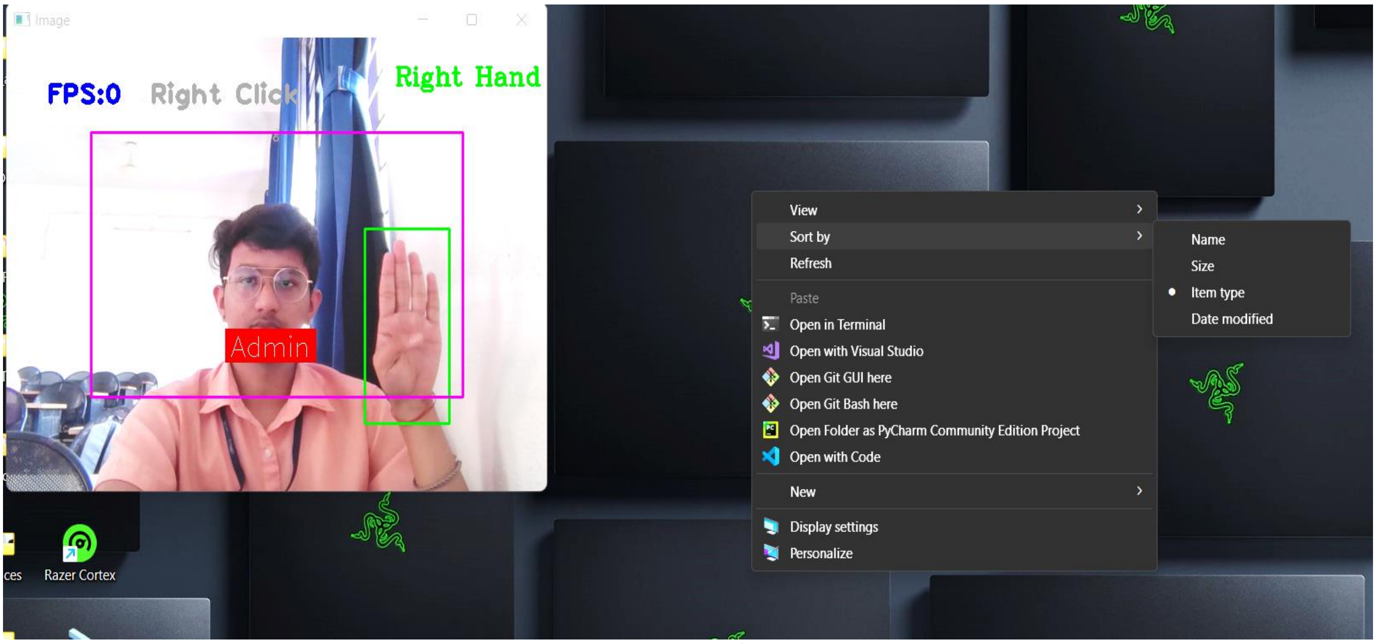


Figure 10. Simulation of Right Click

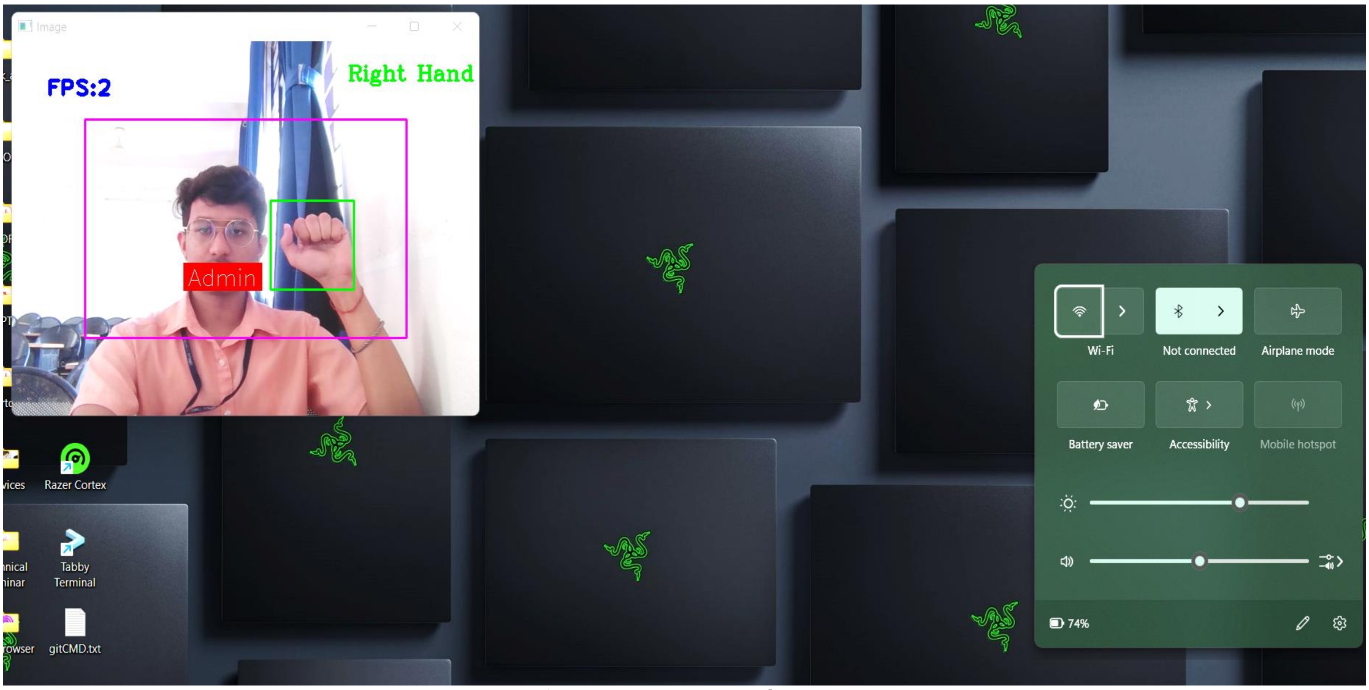


Figure 11. Opening Control Centre

Result Summary

The results confirm that the proposed CNN-based gesture-driven system offers a reliable, touch-free, and customizable human–computer interaction framework. Despite some challenges in multi-hand gesture recognition, the system demonstrates strong performance in common tasks and secure authentication. Its combination of adaptability, personalization, and accuracy highlights its potential for broader adoption in next-generation HCI applications.

8. Conclusion

This study proposed a CNN-based framework for contactless human–computer interaction by combining real-time hand gesture recognition with facial authentication. Experimental evaluations across 20 functional test cases achieved an overall accuracy of ~80%, with gesture recognition performing reliably for discrete single-hand operations and facial authentication achieving 100% success. These outcomes demonstrate that the system provides a secure, customizable, and low-latency alternative to traditional input devices. Despite its limitations in multi-hand gesture recognition and sensitivity to environmental conditions, the framework outperforms many existing approaches by offering multi-factor authentication, gesture customization, and cross-platform adaptability. The integration of these features supports its applicability in healthcare, smart homes, and assistive technologies. Future enhancements will focus on improving robustness under variable conditions, expanding the gesture vocabulary, and incorporating multimodal interaction (e.g., gestures combined with voice commands). By addressing these challenges, the proposed system can evolve into a scalable and reliable next-generation human–computer interaction platform.

Author's statements

Disclosures-The authors declare that there are no financial, personal, or professional disclosures that could be perceived as influencing the research reported in this paper.

Acknowledgements-The authors are grateful to the reviewers for their valuable comments and suggestions, which helped improve the quality of the manuscript.

Funding Source-This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions-Author 1: Conceptualized the research idea, performed the literature review, and contributed to the study design. Author 2: Collected and preprocessed the data, and compiled the initial results. Author 3: Developed and implemented the methodology, including algorithmic design and experimental validation. Author 4: Conducted statistical analysis and contributed to interpretation of the findings. Author 5: Drafted the initial manuscript and was involved in critical revision for intellectual content. All authors: Reviewed, edited, and approved the final version of the manuscript.

Conflict of Interest-The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability-The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

[1] P. Niu, "Convolutional neural network for gesture recognition human-computer interaction system design", PLoS One, Vol.**20**, No.**2**, pp.**e0311941, 2025.** http://dx.doi.org/10.1371/journal.pone.0311941

[2] H. Kim, J. Lee, and J. Park, "Dynamic hand gesture recognition using a CNN model with 3D receptive fields", Proc. Int. Conf. Neural Netw. Signal Process., Nanjing, China, pp.**14–19, 2008.** IEEE.

[3] P. K. Pisharady, P. Vadakkepat, and A. P. Loh, "Attention based detection and recognition of hand postures against complex backgrounds", International Journal of Computer Vision, Vol.**101**, pp.**403–419, 2013.**

[4] M. Bhat, N. Kumar, P. Y. Poojitha, S. Thulasi, and V. Arvind, "CNN based facial recognition with age invariance", International Journal of Research in Applied Science and Engineering Technology, Vol.**11**, pp.**1061–1065, 2023.** http://dx.doi.org/10.22214/ijraset.2023.56680

[5] M. Sharma, G. Akilesh, V. V. Vishwaa, S. F. P. Sharon, and A. Kala, "Virtually controlling computers using hand gesture and voice commands", Journal of Current Research in Engineering and Science, Vol.**5**, No.**17**, **2022**.

[6] S. Song, D. Yan, and Y. Xie, "Design of control system based on hand gesture recognition", Proc. IEEE Int. Conf. Networking, Sensing and Control (ICNSC), **2018.**

[7] R. Damdoo, K. Kalyani, and J. Sanghavi, "Adaptive hand gesture recognition system using machine learning approach", Biosciences Biotechnology Research Communications, Vol.**13**, No.**14**, pp.**106–110, 2020.**

[8] J. J. Beom, S.-K. Kim, and S. Kim, "Enhancing virtual and augmented reality interactions with a MediaPipe–based hand gesture recognition user interface", Ingénierie des Systèmes d'Information, Vol.**28**, No.**3**, pp.**633, 2023.**

[9] M. Al-Hammadi et al., "Deep learning-based approach for sign language gesture recognition with efficient hand gesture representation", IEEE Access, Vol.**8**, pp.**192527–192542, 2020.**

[10] J. Yu, M. Qin, and S. Zhou, "Dynamic gesture recognition based on convolutional neural network and feature fusion", Scientific Reports, Vol.**12**, No.**1**, pp.**4345, 2022.**

[11] X. Wang and J. Tanaka, "GesID: 3D gesture authentication based on depth camera and one-class classification", Sensors, Vol.**18**, No.**10**, pp.**3265, 2018.**

[12] A. Jha et al., "Gessure: A robust face-authentic enabled dynamic gesture recognition GUI application", International Journal of Cybernetics and Informatics (IJCI), Vol.**11**, No.**11**, pp.**19, 2022.**

[13] J. Shin, M. A. M. Hasan, and M. Maniruzzaman, "Hand gesture authentication using optimal feature selection and dynamic time warping based K-nearest neighbor", Proc. Int. Conf. Electron., Commun. Control Eng., **2022.**

[14] S. Parikh, S. Banka, I. Lautrey, I. Gupta, and D. Yedurkar, "Human-computer interaction using dynamic hand gesture recognition to conveniently control the system", International Journal of Engineering and Applied Sciences Technology, Vol.**5**, No.**9, 2021**. ISSN: 2455-2143.

[15] S. K. Shareef, I. V. S. L. Haritha, Y. L. Prasanna, and G. K. Kumar, "Deep learning based hand gesture translation system", Proc. 5th Int. Conf. Trends Electron. Informatics (ICOEI), pp.**1531–1534, 2021.** http://dx.doi.org/10.1109/ICOEI51242.2021.9452947

[16] S. Shahi et al., "Vision-based hand gesture customization from a single demonstration", arXiv preprint, **2024**. https://doi.org/10.1145/3654777.3676378

[17] O. Köpüklü, A. Gunduz, N. Kose, and G. Rigoll, "Real-time hand gesture detection and classification using convolutional neural networks", Proc. 14th IEEE Int. Conf. Automatic Face & Gesture Recognition (FG), Lille, France, pp.**1–8, 2019.** http://dx.doi.org/10.1109/FG.2019.8756576

[18] H. Choi and H. Park, "A multimodal user authentication system using faces and gestures", BioMed Research International, Vol.**2015**, pp.**343475, 2015.** http://dx.doi.org/10.1155/2015/343475

[19] S. Krishna and N. Sinha, "Gestop: Customizable gesture control of computer systems", arXiv preprint, **2020**. https://doi.org/10.1145/3430984.3430993

[20] P. Ramanahally, S. Gilbert, T. Niedzielski, D. Velázquez, and C. Anagnost, "Sparsh UI: A multi-touch framework for collaboration and modular gesture recognition", Proc. ASME Conf. Virtual Environ. Human-Computer Interact. (WINVR), **2009**. http://dx.doi.org/10.1115/WINVR2009-740

[21] O. Köpüklü, A. Gunduz, N. Kose, and G. Rigoll, "Real-time hand gesture detection and classification using convolutional neural networks", arXiv preprint, **2019**.

[22] A. Filipowska, W. Filipowski, P. Raif, M. Pieniążek, J. Bodak, P. Ferst, K. Pilarski, S. Sieciński, R. J. Doniec, J. Mieszczanin, J. Skwarek, K. Bryzik, M. Henkel, and M. Grzegorzek, "Machine learning-based gesture recognition glove: Design and implementation", Sensors, Vol.**24**, No.**18**, pp.**6157, 2023.** http://dx.doi.org/10.3390/s24186157

[23] J. Wu, P. Ren, B. Song, R. Zhang, C. Zhao, and X. Zhang, "Data glove-based gesture recognition using CNN-BiLSTM model with attention mechanism", PLoS One, Vol.**18**, No.11, pp.**e0294174, 2023.** http://dx.doi.org/10.1371/journal.pone.0294174

[24] H. Lee, J. K. Mandivarapu, N. Ogbazghi, and Y. Li, "Real-time interface control with motion gesture recognition based on non-contact capacitive sensing", **2022**.

AUTHORS PROFILE

Mr. Prasad S. R received his B.E. degree in Computer Science and Engineering from VTU, Belagavi, in 2011 and his M.Tech degree from JNTU, Hyderabad, in 2015. He has over 10 years of teaching experience and is currently serving as an Assistant Professor in the Department of Computer Science and Engineering at SDM Institute of Technology. His areas of interest include Deep Learning, Internet of Things (IoT), and Artificial Intelligence (AI). He has guided several student projects in these domains and continues to contribute to academic research and innovative teaching practices.

Mr. Charan K received his B.E. degree in Computer Science and Engineering from Visvesvaraya Technological University (VTU), Belagavi. His academic interests span Deep Learning, the Internet of Things (IoT), and Artificial Intelligence (AI). He completed his undergraduate studies in the Department of Computer Science and Engineering at SDM Institute of Technology, Ujire, where he actively participated in project work and research-oriented activities in emerging technologies.

Mr. Ghanashyama K. P. Makkithaya received his B.E. degree in Computer Science and Engineering from Visvesvaraya Technological University (VTU), Belagavi. His areas of interest include Artificial Intelligence (AI), Deep Learning, and the Internet of Things (IoT). He pursued his undergraduate degree at the Department of Computer Science and Engineering, SDM Institute of Technology, Ujire, with a keen focus on practical applications of modern computational methods and innovation-driven learning.

Ms. Keerthana M. S completed her B.E. in Computer Science and Engineering from Visvesvaraya Technological University (VTU), Belagavi. Her academic interests lie in the domains of Deep Learning, Internet of Things (IoT), and Artificial Intelligence (AI). She pursued her undergraduate studies in the Department of Computer Science and Engineering at SDM Institute of Technology, Ujire.

Ms. Larine Theresa Pereira obtained her B.E. degree in Computer Science and Engineering under Visvesvaraya Technological University (VTU), Belagavi. She has a keen interest in areas such as Deep Learning, IoT, and Artificial Intelligence. She carried out her undergraduate studies in the Department of Computer Science and Engineering at SDM Institute of Technology, Ujire.

