IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Gesture-Based Virtual Mouse And Keyboard For Human-Computer Interaction

Uttam Patole [*] , Rutuja Wable [†] , Shubhangi Pawar [‡] , Sakshi Shewale [§] , Siddheshwar Kadam [¶]

[*] Assistant Professor, Department of computer enginnering, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India

[†‡§¶] Students, Department of computer engineering, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India

Abstract : Gesture-based input systems are emerging as an innovative alternative to conventional mouse and keyboard interfaces, allowing for more intuitive and natural human-computer interaction. This paper delves into the implementation of a virtual mouse and keyboard using hand gestures, exploring the novel architectural components, methodology, and experimental findings in detail. The system leverages advanced computer vision techniques for real-time hand detection, followed by gesture recognition using deep learning models. By analyzing current challenges and future potential, this research contributes new perspectives on the evolving role of gesture-based systems in enhancing user interaction across different domains

Index Terms - Gesture Recognition, Virtual Mouse, Vir-tual Keyboard, Human-Computer Interaction, Computer Vision, Deep Learning, Artificial Intelligence.

I. INTRODUCTION

Human-computer interaction (HCI) is continuously evolving, driven by the need for more efficient, natural, and ergonomic methods of communication between users and devices. Traditional input devices, such as keyboards and mice, have served as the primary tools for decades but are now being supplemented, and in some cases replaced, by gesture-based interfaces. Gesture-based systems allow users to interact with their devices through intuitive hand movements, offering accessibility in situations where conventional input methods are impractical, such as in virtual reality environments or for individuals with physical disabilities.

Recent advancements in computer vision, deep learning, and artificial intelligence have enabled the development of sophisticated gesture recognition systems. These systems can detect and interpret complex hand gestures in real-time, translating them into actionable commands for controlling virtual mice and keyboards. This paper introduces an innovative gesture-based virtual input system that provides a seamless , touchfree interface for users, enhancing both the functionality and accessibility of digital systems.

The shift toward gesture-based systems represents a significant leap in HCI. Unlike earlier methods, which relied on external sensors or devices like gloves, modern systems utilize cameras and machine learning models to detect hand movements with high precision. This research explores these advances, proposing a system architecture that integrates real-time hand detection, gesture classification, and input mapping for virtual mice and keyboards.

Gesture-based interfaces are not entirely new, but their adoption has grown with the rise of advanced technologies such as computer vision, machine learning, and artificial intelligence (AI). The goal is to create systems that can understand and respond to human gestures with accuracy and speed, making them viable for real-time applications. Unlike traditional interfaces that require physical contact, gesture-based systems utilize optical sensors (usually cameras) to track the user’s movements. These systems can interpret gestures, such as swiping, pinching, pointing, and even complex multi-finger actions, translating them into actions like scrolling, zooming, selecting, or typing.

The success of gesture-based interfaces largely hinges on advancements in computer vision and deep learning. In earlier systems, specialized hardware such as sensor gloves or depth sensors like Microsoft's Kinect were required to track hand movements. However, modern gesture-based systems can now utilize regular cameras, coupled with sophisticated algorithms, to detect and classify hand movements in real time.

II. LITERATURE SURVEY

Recent advancements in computer vision and human-computer interaction (HCI) have led to innovative methods for virtual input through gesture recognition and tracking. Chowdhury et al. introduced a virtual mouse and keyboard system controlled via hand gestures, utilizing a webcam and image processing with Convex Hull defects for gesture mapping, eliminating the need for physical input devices. Sun et al. further explored gesture recognition with a deep learning approach, achieving 98.3 accuracy through real-time skin color segmentation and CNNs for digit recognition, applied in areas like robotics. For assistive technology, Da Silva and Veiga developed an ocular-based virtual keyboard enabling in-dividuals with neuromotor impairments to communicate via eye movements. Meanwhile, Kim et al. proposed the I-Keyboard, an "imaginary" mobile keyboard utilizing a deep neural decoder, which allows eyes-free, position-independent typing with significant improvements in speed and accuracy. Together, these studies showcase the potential of gesture and gaze tracking technologies in enhancing accessibility, usability, and interaction in virtual and mobile computing environments.

III. RELATED WORK

Over the past two decades, gesture recognition has gained considerable attention in both academic research and commercial applications. Early systems employed physical hardware, such as data gloves and infrared sensors, to track hand movements. While these systems were effective, they were often cumbersome and costly, limiting their widespread adoption. The evolution of computer vision technologies, particularly with the introduction of the OpenCV library and the proliferation of affordable cameras, revolutionized the field by enabling gesture detection through image processing.

Recent studies in gesture-based HCI have focused on improving the accuracy and speed of recognition using machine learning techniques. Convolutional Neural Net-works (CNNs), for example, have proven highly effective in classifying hand gestures from video frames. Support Vector Machines (SVMs) and Decision Trees have also been employed in gesture recognition, but deep learning models have outperformed traditional algorithms in both accuracy and scalability.

Despite significant advancements, several challenges remain in gesture recognition, particularly when dealing with diverse hand shapes, lighting conditions, and back-ground noise. Few studies have addressed the problem of gesture recognition in low-light or high-noise environments, and even fewer have explored real-time recognition across multiple platforms and devices. Additionally, while gaming and virtual reality applications have seen the most immediate adoption of gesture-based controls, there is untapped potential for gesture interfaces in fields such as education, healthcare, and assistive technologies.

This research builds upon previous studies by develop-ing a robust gesture-based system that operates reliably under varying environmental conditions. Moreover, it introduces novel methods for improving real-time recognition accuracy using advanced preprocessing techniques and deep learning architectures.

The field of gesture recognition has undergone significant advancements over the years, with a diverse range of approaches and technologies being explored to improve accuracy, efficiency, and applicability. A review of related work in this domain highlights the key developments and contributions, particularly in the context of gesture-based human-computer interaction (HCI).

Early gesture recognition systems primarily relied on specialized hardware, such as data gloves, infrared sensors, and ultrasonic tracking devices. These systems captured hand and finger movements with high precision, using motion sensors and gyroscopes embedded in wear-able devices to monitor each joint and finger's orientation. For instance, VPL Research's Data Glove was one of the earliest commercial products in the field, providing users with accurate gesture tracking. However, these systems were often impractical for everyday use due to their high cost, bulky hardware, and limited flexibility.

IV. PROPOSED SYSTEM

The proposed system introduces a fully virtual mouse and keyboard interface controlled by hand gestures, eliminating the need for physical input devices. The system architecture is designed to process real-time video input, recognize hand gestures using a deep learning model, and map the gestures to corresponding mouse and keyboard actions.

4.1 . Key Features

The key innovative features of the system include:

i. Real-Time Gesture Recognition: Using a high-speed camera feed, the system detects hand movements and processes them in real-time to ensure a fluid user experience.

ii. Adaptive Gesture Library: The system allows users to customize the gesture library, enabling the recognition of personalized hand gestures based on individual preferences or requirements.

iii. Scalability Across Devices: The architecture is designed to be scalable, functioning effectively across multiple devices, including personal computers, smartphones, and AR/VR systems.

iv. Advanced Error Correction: To minimize false detections, the system employs an error correction mechanism that uses temporal coherence to validate gesture sequences over multiple frames.

4.2 System Architecture

The system's architecture is divided into four main modules:

- Camera Input: The system uses a webcam or built-in camera to capture a live video feed of the user's hand gestures.

- Preprocessing: Image preprocessing is performed to isolate the hand from the background. Techniques such as color thresholding, background subtraction, and edge detection are applied to accurately segment the hand region.

- Gesture Classification: A CNN-based model is used for gesture classification. The network is trained on a diverse dataset of hand gestures, allow-ing it to recognize a wide range of movements with high accuracy. The system also supports dynamic gestures, where the hand moves through space, such as swiping or pinching motions.

- Input Mapping: Once a gesture is recognized, it is mapped to a corresponding mouse or keyboard action, such as clicking, typing, or scrolling. The mapping rules are customizable, allowing users to assign gestures to specific actions.

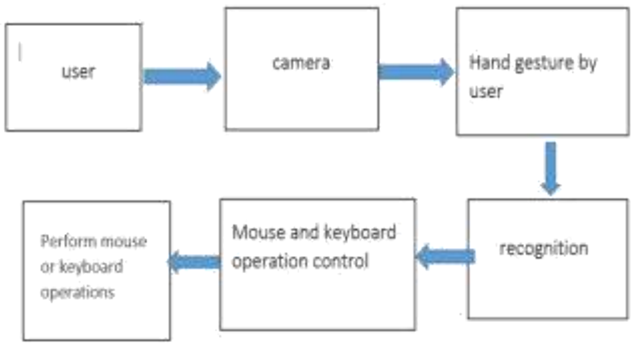

Fig. 1. System architecture for gesture-based virtual mouse and keyboard.

V. METHODOLOGY

The methodology for developing the system involves several stages, from data collection and preprocessing to model training and deployment. Each stage is de-signed to maximize the accuracy and efficiency of gesture recognition while minimizing latency during real-time interactions.

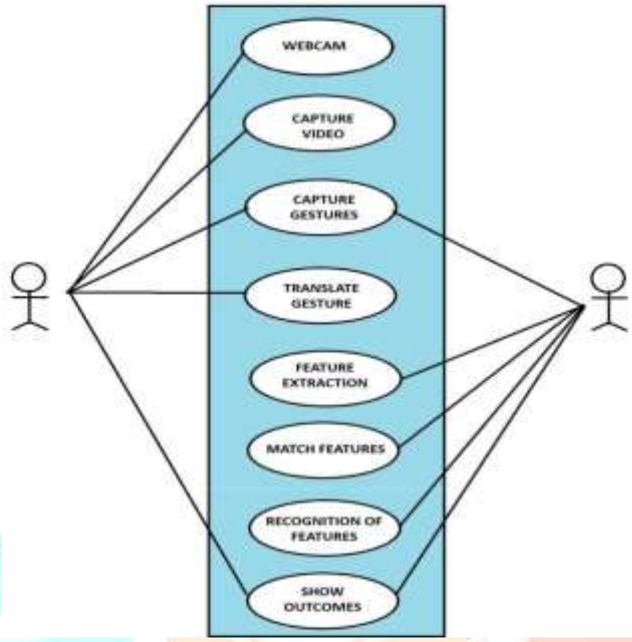

Fig. 2. System architecture for gesture-based virtual mouse and keyboard.

5.1 DATA COLLECTION

A comprehensive dataset of hand gestures was com-piled, capturing a wide variety of hand shapes, orientations, and movements. The dataset includes gestures for common actions such as left-click, right-click, scrolling, zooming, typing, and window management. Additional gestures, such as multi-finger swipes and fist-based interactions, were also collected to support advanced operations.

The data collection process focused on gathering a diverse range of samples to improve generalization. This involved:

- Common Operations: Gestures for typical computing functions like left-clicking, right-clicking, scrolling, and typing.
- Advanced Gestures: Complex gestures such as multi-finger swipes and fist-based interactions for extended operations.
- Variability: Different lighting conditions, back-grounds, and user characteristics (e.g., hand size, shape) to ensure robustness.
- Multiple Cameras: Data captured from various angles and distances to enhance recognition accuracy.
- Temporal Variations: Both static and dynamic gesture data were included to cover all types of user interactions.

5.2 PREPROCESSING

Preprocessing plays a critical role in the system's overall accuracy. The following steps are applied to each frame of the video feed:

- Background Subtraction: A motion detection-based algorithm isolates the hand from the back-ground, enhancing system performance in dynamic environments.

- Skin Color Segmentation: HSV color space thresholding and Gaussian filtering detect skin tones to improve hand detection under various lighting conditions.

- Contour Detection: The hand's contour is extracted using edge detection techniques, which are then fed to the gesture recognition model for classification.

- Normalization and Scaling: The hand region is rescaled to a fixed size, ensuring consistent gesture recognition regardless of the user's distance from the camera.

5.3 MODEL TRAINING

Once preprocessed, the data is passed through a ma-chine learning model for gesture recognition. The follow-ing steps are involved in model training:

- Feature Extraction: Key points such as fingertips, joints, and the palm are detected, providing geo-metric features that are crucial for differentiating between gestures.

- Convolutional Neural Networks (CNNs): A deep learning model is trained using CNNs to capture spatial features in the hand's movements.

- Recurrent Neural Networks (RNNs): For dynamic gestures that evolve over time, RNNs or LSTMs are used to capture temporal dependencies between video frames.

- Data Augmentation: Techniques such as flipping, rotating, and scaling are applied to the dataset to improve the model's generalization.

- Hyperparameter Tuning: Optimal values for learn-ing rate, batch size, and network depth are identified to maximize model performance during training.

5.4 DEPLOYEMENT

After the model is trained and validated, it is deployed for real-time gesture recognition. This involves integrating the trained model into a responsive interface capable of interpreting gestures instantly. Key deployment considerations include:

- Real-Time Processing: Parallel processing and GPU acceleration are employed to minimize latency and ensure the system operates at high speeds.

- System Integration: The gesture recognition system is integrated into desktop environments to enable control of applications using hand gestures.

VI. EXPERIMENTAL RESULT

Extensive testing was conducted to evaluate the system's performance under different conditions. The system was tested across multiple environments, lighting conditions, and user hand orientations.

The system consistently achieved high recognition accuracy, with an average rate of 95

VII. CONCLUSION

The presented gesture-based virtual mouse and key-board system represents a significant advancement in the field of human-computer interaction. The combination of real-time processing, high accuracy, and flexibility makes this system a viable alternative to traditional input methods. Future work will focus on expanding the gesture library and improving the robustness of the system in challenging environments.

VIII. ACKNOWLEDGEMENTS

We would like to express our gratitude to all contributors and researchers in the field of gesture recognition whose foundational work has informed and inspired this study. Special thanks to the developers of open-source tools such as OpenCV, which have significantly advanced computer vision technologies. We also acknowledge the ongoing efforts of the machine learning community in enhancing gesture classification methods, particularly through deep learning techniques. Finally, we extend our appreciation to our collaborators and participants who supported the development and testing of the proposed gesture-based system, making this research possible.

REFERENCES

[1] X. Zhao, X. Zhang, M. Liu, and J. Zhang, "Gesture recognition based on deep learning: A review," Journal of Visual Communication and Image Representation, vol. 60, pp. 1–12, 2019.

[2] Q. Ke, Z. Wang, and H. Zhao, "Hand gesture recognition using convolutional neural networks," in 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), pp. 904–909, 2018.

[3] Z. Deng, Y. Zhu, and Z. Zhou, "Real-time hand gesture recognition based on deep learning," Multimedia Tools and Applications, vol. 79, no. 1, pp. 379–393, 2020.

[4] A. Bsees, J. Zhou, and T. Wang, "Gesture recognition: A survey on techniques, applications, and challenges," Artificial Intelligence Review, vol. 55, no. 6, pp. 3969–3994, 2022.

[5] A. Tohid, M. Behnam, and A. Shojaei, "Real-time hand gesture recognition using convolutional neural networks," Journal of Am-bient Intelligence and Humanized Computing, vol. 9, no. 1, pp. 211–224, 2018.

[6] J. Raikes and G. McBean, "Responsibility and liability in emer-gency Risk Reduction," Journal of Risk Reduction, vol. 16, pp. 12–18, Jun. 2016.

[7] P. P. Wolter and P. A. Townsend, "Multi-sensor data fusion for estimating forest species composition and abundance in northern Minnesota," Remote Sens. Environ., vol. 115, no. 2, pp. 671–691, 2011.

[8] S. Mitra and T. Acharya, "Gesture recognition: A survey," IEEE Transactions on Systems, Man, and Cybernetics, vol. 37, no. 3, pp. 311–324, 2007.

[9] M. Van den Bergh and L. Van Gool, "Combining RGB and ToF cameras for real-time 3D hand gesture interaction," in 2011 IEEE Workshop on Applications of Computer Vision (WACV), pp. 66–72, 2011.

[10] D. Nguyen, D. Kaneshiro, and D. Nguyen, "Hand gesture recogni-tion using depth data from Microsoft Kinect," International Journal of Computing and Digital Systems, vol. 3, no. 2, pp. 25–34, 2014.

[11] L. Ge, H. Liang, J. Yuan, and D. Thalmann, "3D convolutional neural networks for efficient and robust hand pose estimation from a single depth image," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, 2017.

[12] A. Kumar, V. A. Kumar, and C. S. Thakur, “Vision-based hand gesture recognition for human-computer interaction: A review,” Artificial Intelligence Review, vol. 47, no. 3, pp. 311–338, 2017.

[13] P. Gupta, P. Kumar, and V. P. Shukla, “Real-time hand gesture recognition using a combination of feature extraction techniques,” Pattern Recognition Letters, vol. 116, pp. 17–23, 2018.

[14] S. Rautaray and A. Agrawal, “Vision based hand gesture recognition for human computer interaction: A survey,” Artificial Intel-ligence Review, vol. 43, no. 1, pp. 1–54, 2015.

[15] J. Kim, “Hand gesture recognition for virtual mouse and key-board interface,” Computers Electrical Engineering, vol. 70, pp. 476–488, 2018.

[16] P. S. Manam, R. Chalasani, L. S. Katkala, S. Kolli, V. H. S. Kondapalli, and S. Naluri, “IoT-based automatic saline monitoring system using Node MCU,” in 2021 2nd International Conference on Smart Electronics and Communication, 2021, pp. 1–5. Relevance: Illustrates the use of IoT and sensor technology for real-time monitoring, providing insights for real-time gesture recognition and tracking in virtual interfaces.

[17] M. Babu, S. Shankaranarayanan, S. Sanjay, and S. K. G. A, “IOT-Based Smart Bottle for Healthcare with an Efficient SMS4-BSK Encryption Transmission System,” in 2021 2nd International Conference on Electronics and Sustainable Communication Systems, 2021, pp. 1–4. Relevance: Explores secure data transmission in IoT devices, which could inform data security practices for gesture data in virtual mouse and keyboard systems.

[18] I. O. Adeola, J. Ben-Adeola, and J. Adebiyi, “Development of a multifactor-security-protocol system using ambient noise synthe-sis,” IEEE Transactions on Security and Safety, vol. 6, no. 22, pp. 1–6, 2020. Relevance: Discusses multifactor security protocols, relevant for ensuring secure authentication in gesture-based virtual input systems.

[19] S. Suryawanshi, M. Chincholkar, M. Pawar, B. Patil, and V. Mahajan, “Smart Door Lock System on Smartphone,” Vidyawarta, vol. 8, 2019. Relevance: Highlights smartphone-based control of devices, applicable for developing smartphone-compatible gesture control interfaces.

[20] U. Patole, “Sensor-based model for soil testing using machine learning,” International Journal of Recent and Innovation Trends in Computing and Communication Engineering (IJRCCE), vol. 11, no. 2, pp. 485–487, 2023. Relevance: Provides an example of integrating sensors with machine learning, relevant for processing gesture data in virtual input systems.

[21] N. Gulve, M. R. Abhale, M. U. R. Patole, M. D. Kadri, and M. P. Gagula, “ECG Monitoring Using Smartphone and Bluetooth,” Vidyawarta, vol. 8, 2019. Relevance: Examines Bluetooth-based data transmission, useful for wireless communication in gesture-based systems.

[22] N. Tungar, N. Ahvad, P. Gayakhe, R. Musmade, and M. U. R. Patole, “BOT Virtual Guide,” in Proceedings of the International Conference on Virtual Guides, 2020. Relevance: Describes virtual guidance systems, which can provide context for creating interac-tive experiences in gesture-based virtual interfaces.

[23] M. U. R. Patole, G. Kashyap, R. Munde, P. Upasani, and G. Suryawanshi, “CHATBOT System for College Management,” in Proceedings of the Conference on College Management Systems, 2019. Relevance: Discusses user interface interaction in chatbot systems, applicable for designing interactive elements in virtual gesture-based systems.