Proceeding Paper

Multimodal Fusion for Enhanced Human–Computer Interaction †

Ajay Sharma [1,*], **Isha Batra** [1] , **Shamneesh Sharma** [2] **and Anggy Pradiftha Junfithrana** [3]

1   Department of Computer Science and Engineering, Lovely Professional University, Jalandhar 144411, Punjab, India; isha.batra2487@gmail.com
2   Customer Success and Quality Control, byteXL TechEd Private Limited, Hyderabad 500081, Telangana, India; shamneesh.sharma@gmail.com
3   Department & Faculty of Electrical and Electronics Engineering, Nusa Putra University, Sukabumi 43152, West Java, Indonesia; nutral@nusaputra.ac.id
*   Correspondence: ajaysharmadharmani@gmail.com
†   Presented at the 7th International Global Conference Series on ICT Integration in Technical Education & Smart Society, Aizuwakamatsu City, Japan, 20–26 January 2025.

Abstract

Our paper introduces a novel idea of a virtual mouse character driven by gesture detection, eye-tracking, and voice monitoring. This system uses cutting-edge computer vision and machine learning technology to let users command and control the mouse pointer using eye motions, voice commands, or hand gestures. This system's main goal is to provide users who want a more natural, hands-free approach to interacting with their computers as well as those with impairments that limit their bodily motions, such as those with paralysis—with an easy and engaging interface. The system improves accessibility and usability by combining many input modalities, therefore providing a flexible answer for numerous users. While the speech recognition function permits hands-free operation via voice instructions, the eye-tracking component detects and responds to the user's gaze, therefore providing exact cursor control. Gesture recognition enhances these features even further by letting users use their hands simply to execute mouse operations. This technology not only enhances personal user experience for people with impairments but also marks a major development in human–computer interaction. It shows how computer vision and machine learning may be used to provide more inclusive and flexible user interfaces, therefore improving the accessibility and efficiency of computer usage for everyone.

Keywords: virtual mouse; eye ball control; speech control; gesture recognition; computer vision; machine learning; mobility impairments; hands-free interaction; intuitive interface

1. Introduction

In the search for inclusive technology, a lot of work has gone into creating substitute computer input systems catered for those with mobility problems. Because they need exact motions and synchronization, the traditional means of interacting with a computer mouse and keyboard often provide difficult problems. In light of this obstacle, recent developments have concentrated on innovative ideas to democratize computer access for everyone. One such development is eye-tracking systems, which use complex camera and software combinations to monitor and quantify eye motions. These systems convert visual motions into instructions, therefore enabling users to negotiate and manage computer interfaces without depending on conventional manual input methods [1]. Complementing this type of technology, voice recognition capabilities have become more important for improving access. Speech instructions may be understood and carried out by using

specialist software algorithms, therefore allowing a hands-free computer interface. This combination of eye tracking and voice recognition not only improves navigation but also serves those with different degrees of mobility handicap by providing a more inclusive computer environment [2].

Including gesture-tracking technology helps to further change the input modalities. Systems may now read hand gestures as simple instructions by using computer vision and machine learning techniques, therefore extending the repertoire of available interact Faculty of Electrical and Electronics Engineering beyond conventional approaches. This multifarious approach not only meets the many demands of individuals with mobility problems but also promotes a more natural and easy means of computer engagement. The main objective is to design interfaces that are not just accessible but also understandable and user-friendly [3]. The Table 1 provides a comprehensive survey of available algorithm for the gesture recognition, eye tracking movements and voice monitoring algorithm, for the human computer interaction, voice monitoring speech recognition and integration and fusion based models. The Table 2 explains the information about the mathematical expression for the virtual mouse movement.

Table 1. Virtual mouse interface using gesture recognition, eye tracking, and voice monitoring algorithm and methods.

Gesture Recognition (Hand Recognition)	
Algorithm	**Techniques**
Media Pipe Hands (by Google)	Real-time hand tracking and gesture classification.
CNN + LSTM	Convolutional Neural Networks for spatial feature extraction and Long Short-Term Memory for temporal pattern recognition in dynamic gestures.
YOLOv7 Pose or Open Pose	The algorithms detect the body/hand movement.
Support Vector Machine (SVM)	The algorithms classify the hand poses and extract the features for gesture recognition and human–computer interaction.
K-Nearest Neighbour (KNN)	Lightweight method to classify gestures based on landmark distances.
Eye Tracking Algorithm for Human–Computer Interaction	
To control cursor movement with eye movement (Gaze Tracking)	
Pupil Labs SDK/OpenGaze/GazeML	For real-time gaze implementation in eyes pupils.
CNN for Eye Feature Extraction	The map eye landmarks to screen the coordinates. The model is trained on the CNN algorithms.
Kalman Filter	The filter is used to predict the gaze and to reduce the noise from the current images.
FACE MESH Detection	The media pipe located the eyes and pupils to infer directions.
Polynomial Regression Model	Maps the eye features (IRIS center) to screen the positions.
Voice Monitoring (Speech Recognition)	
To trigger the mouse actions with the mouse commands "Clicks"," Scroll"," Right and lift side clicks"	
Google Speech Recognition	Google API for the recognition of voice samples.
Mozilla Deep Speech or Whisper tools by the Open AI	The offline-based model for mobile devices is powered by deep learning-based voice recognition.
NLP (Command Detection)	Tokenization ------> Command-based classification
	Intent recognition with the help of RNN-/transformer-based models.
Wake- Word Detection	When a small CNN/RNN detects a trigger word like "Hey Mouse" or "Hello Mouse", etc.

Integration and Multimodel Fusion	
Combine inputs from gestures, voice, and eye movements.	
Decision-Level Fusion	The rule-based system/voting strategy to decide dominant inputs (eyes are idle; use gestures to control the decision level).
Context-Aware Fusion	Use of Machine Learning or probabilistic models, e.g., HMM (Hidden Markov Model).
Weighted Sensor Fusion	Apply the weighted input sources on the signal confidence and the noise level.
User Experience Enhancers for the HCI	
Adaptive Learning	To train user-specific models for few-shot learning.
Latency Reduction	Use a lightweight technique.
Accessibility Feedback	Use the audio/visual techniques for confirming the actions (clicks, selections, etc.).

Table 2. Table showing the mathematical expression, logic-based condition.

Mathematical Expression of Virtual Mouse System	
Input Function	Let: $G(t) \rightarrow (a\_g, c\_g)$: Gesture input at time t, where $\quad - a\_g$: Gesture action (e.g., "click", "scroll"). $\quad - c\_g$: Confidence score of gesture. $-E(t) \rightarrow (x\_e, c\_e)$: Eye input at time t, where $\quad - x\_e$: Cursor coordinates from gaze. $\quad - c\_e$: Confidence score of gaze estimation. $-V(t) \rightarrow (a\_v, c\_v)$: Voice input at time t, where $\quad - a\_v$: Voice command (e.g., "right click"). $\quad - c\_v$: Confidence score of voice recognition.
Fusion Weight	Let the weights for each modality be $w\_g = 0.4, \ w\_e = 0.4, \ w\_v = 0.2$
Decision Thresholds	$\theta\_g = 0.6, \ \theta\_e = 0.6, \ \theta\_v = 0.5$
Control Logic	At each time step t, the system evaluates If $w\_g * c\_g(t) > \theta\_g$, then $M(a\_g)$ Else if $w\_e * c\_e(t) > \theta\_e$, then $M(x\_e)$ Else if $w\_v * c\_v(t) > \theta\_v$, then $M(a\_v)$ where $\quad - M(a)$: Executes mouse action a. $\quad - M(x)$: Moves mouse cursor to position x.
Optimal Visual Feedback Function	Let $F(G(t), E(t), V(t))$ represent a feedback function that shows the current inputs visually.
Loops	The system runs the decision model continuously, $\forall \ t \in T$, and evaluates inputs and controls mouse according to the logic above.
Mathematical Control	Mathematical Control Rule (per frame): $M\_t =$ $\quad M(a\_g), \ if \ w\_g * c\_g(t) > \theta\_g$ $\quad M(x\_e), \ else \ if \ w\_e * c\_e(t) > \theta\_e$ $\quad M(a\_v), \ else \ if \ w\_v * c\_v(t) > \theta\_v$ $\quad$ No action, otherwise.

The design philosophy of these systems is based mostly on the idea of welcoming. These systems constantly adapt and grow by using innovative technologies like machine learning, thereby customizing the user experience to fit particular tastes and capabilities. Furthermore, they are designed to fit not just those with limited mobility but also those who may have difficulty with conventional input devices because of character problems or other circumstances. This inclusive strategy goes beyond simple utility to include a basic change toward a more human-centered computer environment. In the search for universal access to computers, the confluence of eye-tracking, voice recognition, and gesture-tracking technologies essentially marks a major breakthrough. These technologies enable people with mobility problems to completely engage in the digital era by offering alternate input techniques that go beyond the restrictions of conventional devices [3,4]. They also open the path for a more inclusive and human-centric technology, where designs give top priority to the requirements and preferences of the user.

2. Literature Review

Aiming to find out whether gaze-based data, along with hand–eye coordination information, can precisely forecast user interaction intents inside digital environments, Chen and Hou's work focuses on researching gaze-based interaction intention detection in virtual reality (VR) settings. Using Gradient Boosting Decision Trees, the research shows the effective prediction of item selection and teleporting events with an amazing F1-Score of 0.943 by means of analyzing eye-tracking data from 10 individuals engaged in VR activities, hence surpassing previous classifiers. Furthermore, this research underlines the possibility of using hand–eye coordination characteristics to improve interaction intention identification in virtual reality, thus stressing the need for multimodal data fusion in improving forecasts. This study offers an insightful analysis of the evolution of predictive interfaces in virtual reality, therefore improving user experience and immersion in many different applications [1].

In the current paper, Pawłowski, Wroblewska, and Sysko-Romanchuk take the pressing challenge faced by robotics in effectively processing multimodal data, emphasizing the critical need for building comprehensive and common representations. They highlight the abundance of raw data available and underscore the importance of smart management, positioning multimodal learning as a pivotal paradigm for data fusion. Despite the success of various techniques in building multimodal representations, the authors note a gap in the literature regarding their comparative analysis within production settings. To address this, the paper explores three prominent techniques like late fusion, early fusion, and the sketch method assessing their efficacy in classification tasks. Notably, the study encompasses diverse data modalities obtainable from a wide array of sensors serving various applications. Through experiments conducted on Amazon Reviews, MovieLens25M, and MovieLens1M datasets, the authors elucidate the significance of fusion technique selection in maximizing model performance through the optimal modality combination [2].

The study discusses the multifaceted application of eye movements in human–computer interaction (HCI) (Figure 1) and usability research, both as a tool for analyzing interfaces and measuring usability and as a direct control medium within a user–computer dialogue. While these areas have traditionally been approached separately, the text emphasizes the importance of integrating them. In usability analysis, eye movements are recorded and analyzed retrospectively, providing insights into user behavior but not directly influencing the interface in real time. Conversely, as a control medium, eye movements are utilized in real time as input for the user–computer dialogue, either as the sole input or in conjunction with other devices. However, both retrospective analysis and real-time use present significant challenges, including interpreting eye movement data effectively and

responding appropriately without over-responding to eye movement input. Despite the promises of eye-tracking technology in HCI, progress in its practical implementation has been relatively slow [3,4].
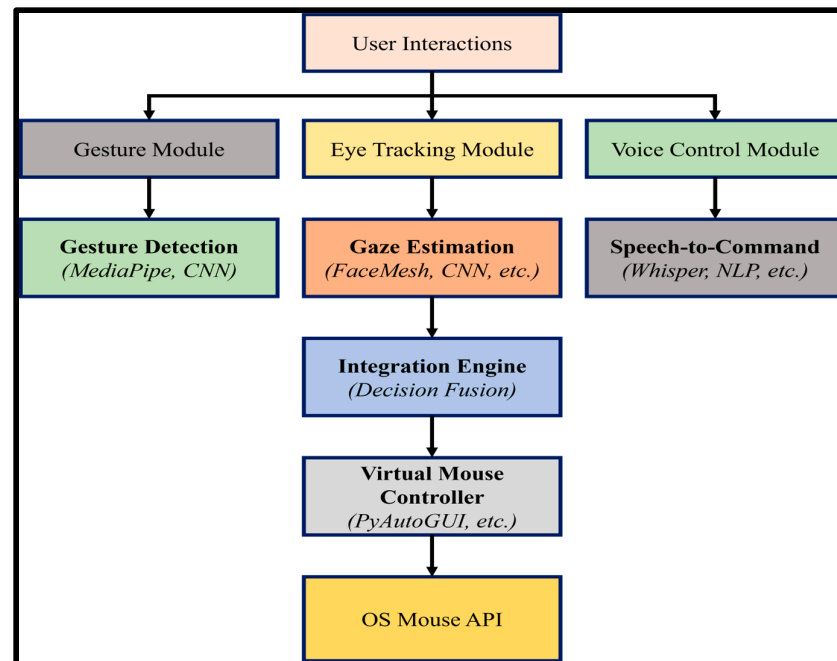
Figure 1. High-level architecture diagram showing how gesture, eye, and voice input come together to drive the virtual mouse.

According to Noussaiba Jaafar, security systems have concentrated on unattached modalities like video or audio, but their ambiguity and complexity make the identification of complex behaviors like aggressiveness difficult. The article presents an approach that combines audio, video, and text modalities, along with meta-information, using deep learning techniques to detect aggression. Four multimodal fusion methods are explored, including an intermediate-level fusion approach utilizing deep neural networks (DNNs), a concatenation of features with meta-features, and element-wise operations such as product and addition. With an unweighted average accuracy of 85.66% and a weighted average accuracy of 86.35%, the second fusion technique beats others, exceeding current methods according to the research. Meta-features greatly enhance prediction performance using all fusion techniques. The results imply that every fusion technique has benefits and constraints for estimating various degrees of aggressiveness, thereby stressing the need for multimodal techniques and meta-information in surveillance networks for improved violence detection [4].

Emphasizing the need for usability and connection between humans and computers, the article explores the development of HCI within the framework of growing computer technology. It describes many methods of HCI research, including examining present trends and pointing out areas that need further growth. Examining how consumers' knowledge and expectations affect their interactions with computer systems, one focuses on the idea of mental models in HCI. Further exploring the function of emotional intelligence in HCI, the research seeks to produce more interesting and user-friendly interfaces. It also underlines the importance of fidelity prototypes in the design process and the need for automated solutions to help with these chores. Emphasizing the interaction between users and computers, the introduction defines HCI as the study of how people use computers to successfully and enjoyably complete activities [5,6].

Usually using tensors, multimodal data fusion is the process of converting data from many single-modal representations into a compact multimodal representation. This method, nevertheless, rapidly increases computing complexity. Hu Zhu's work suggests using a low-rank tensor multimodal fusion approach with an attention mechanism to handle this problem, therefore lowering computing cost and improving efficiency. Three multimodal fusion challenges using public datasets, CMU Multimodal Opinion Sentiment and Emotion Intensity (CMU-MOESI), Interactive emotional dyadic motion capture database (IEMOCAP), and Project Object Model (POM), evaluate the proposed model. The effective capture of global and local linkages helps the model to show an excellent performance. Furthermore, the proposed model shows better and stable outcomes across many attention processes than prior tensor-based multimodal fusion approaches [7].

3. Implementation of the Proposed Work

The design of a virtual mouse driven by gestures, voice, and eye motion requires many complicated elements and approaches. Key components employed in face movement, eye tracking, and vocal command detection include image processing techniques, decision trees, and machine learning. The system uses decision tree or regression tree techniques to find the iris, therefore converting data into edges and borders to provide detailed pictures for exact detection in tracking eye movements. Using the eye location from a normal camera as input, facial motions are controlled to scroll the mouse cursor in many directions [1]. The proposed implementation along with the working methodology has been explained in the Figure 1, the architecture for the gesture input and output controls. The Figure 2 shows the developmental process for the human computer interaction.
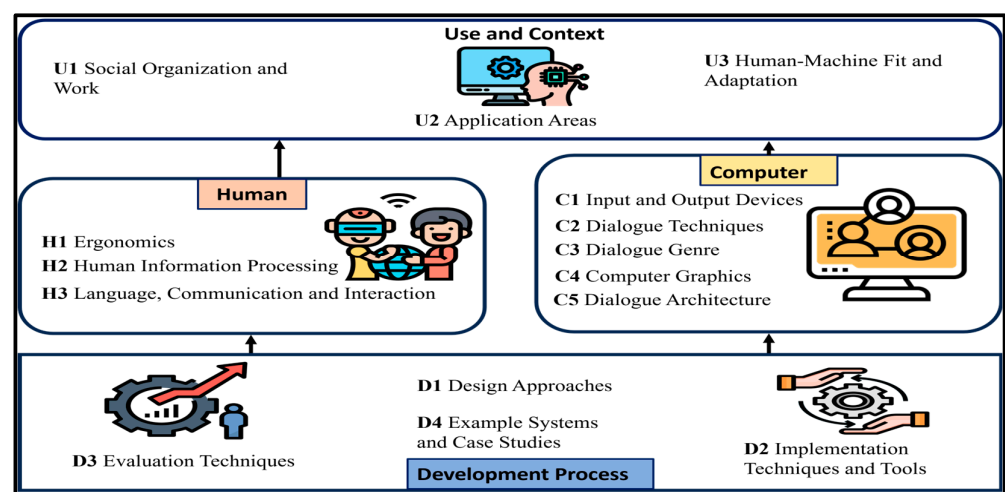
Figure 2. A diagram of a human–computer interaction (HCI) development process.

Another control technique is speech recognition, which interprets mouse event user instructions. Face identification, eye extraction, and blinking sequence interpretation using a webcam input guarantee a non-invasive human–computer interaction (Figure 2). A movement-sensing camera included in headgear with an eye motion sensor allows gesture control. This replaces conventional accuracy and physical effort with image-processing methods and artificial intelligence, therefore allowing cursor adaptability on the screen [8].

Using probabilistic techniques like pupil detection and screen position detection, a real-time pupil-centered machine vision system including two cameras, infrared markers, and electrical components follows the user's eye gaze to move the pointer appropriately [3]. The technology seeks to offer those requiring a natural engagement with computers or those with mobility limitations an easy interface. For those with impairments in particular,

this method improves autonomy and productivity, therefore creating fresh opportunities for effective computer connection [5]. Three main components define the proposed system: nonverbal figure identification, speech understanding, and eye movement detection.

3.1. Eye Movement Detection

The Tobii Eye Tracker 4C is a specialist tool using infrared light to monitor and track user eye movements during concentration activities, therefore enabling eye movement detection. The camera of the gadget records reflections from the user's eyes; specialist software examines these reflections in real time to find the gaze direction.

3.2. Speech Comprehension

A commercial speech recognition engine, like Google Speech-to-Text, helps one to grasp speech. This engine analyzes user voice using machine learning techniques; then, it very precisely generates text. Users of the system may command or input vocally, therefore interacting with the computer without physical touch.

3.3. Nonverbal Figure Detection

Highlighting pupil monitoring, the real-time machine vision system tracks user gaze and cursor movement using two cameras, electrical components, and infrared markers. Together with screen position detection, probabilistic methods for pupil detection, arranged in a cascade, allow the exact monitoring of eye movements in real time [3].

Especially for single-spot examinations, the Tobii Eye Tracker 4C is included in the system for tracking and confirming user eye movements. While the camera underneath records reflections, infrared light technology shines on the user's eyes. This data is processed by specialized software to ascertain the user's gaze direction, therefore improving the accuracy of the eye movement detecting component [6]. By use of voice recognition and eye-tracking technology, users may communicate with the system orally and via eye motions, therefore improving user comfort and efficiency. The exact translation of voice to text guarantees flawless connection with the system, therefore allowing users to provide orders or input without direct physical contact with the computer [7–9].

4. Proposed Model

People with physical mobility constraints may efficiently engage with computers utilizing the accessibility and usability approach of the many models suggested in this study of a PC controlled using eye ball control, voice speech control, and hand gesture control. There are numerous parts to the system [10,11].

4.1. Eye Movement Detection

The system distinguishes the iris from one person to another by employing a decision tree algorithm or regression tree algorithm, which function to see the position of the iris in the eyes of the user. Given that the recorded image is transformed into edges and corners/junctions that can be precisely observed by the system, we can track all the movements of the eyes [12–14].

4.2. Facial Movement Detection

Indeed, the main method of controlling this system is by facial expressions, such as leading the cursor by moving the face to the left or right, upwards and downwards. The system uses a normal camera as the input device to capture an input image and arranges eyes in two positions to control the movement of the cursor [15–17].

4.3. Speech Recognition

The system is based on a speech recognition background to understand the user's voice commands, which drives the mouse processor to carry out mouse actions. The system usually relies on a regular webcam to input an image. Image processing methods, such as face detection, eye extraction, and interpretation of blinks in real time, are performed for controlling the human–machine interface in a non-intrusive, eye-scanning way [17–19].

4.4. Gesture-Controlled Mouse

The technique uses a headpiece with an eye motion sensor for recording the eye movements, and no hands are required in performing these movements on the screen. Nonetheless, the improvement of image processing techniques and the cognitive capacities of artificial intelligence make this task feasible without any need for specialized equipment, just a simple web camera attached to a computer [16–19].

4.5. Real-Time Daylight-Based Eye Gaze Tracking System

The system is based on two cameras, infrared markers, and electronic components. In detecting pupils and the screen position, two methodologies work at the same time. They ensure the user's eyes are nominatively followed and the cursor moves as per the same track [19]. The Figure 3 highlights the different type of the interfaces used in the various application for the human computer interaction.
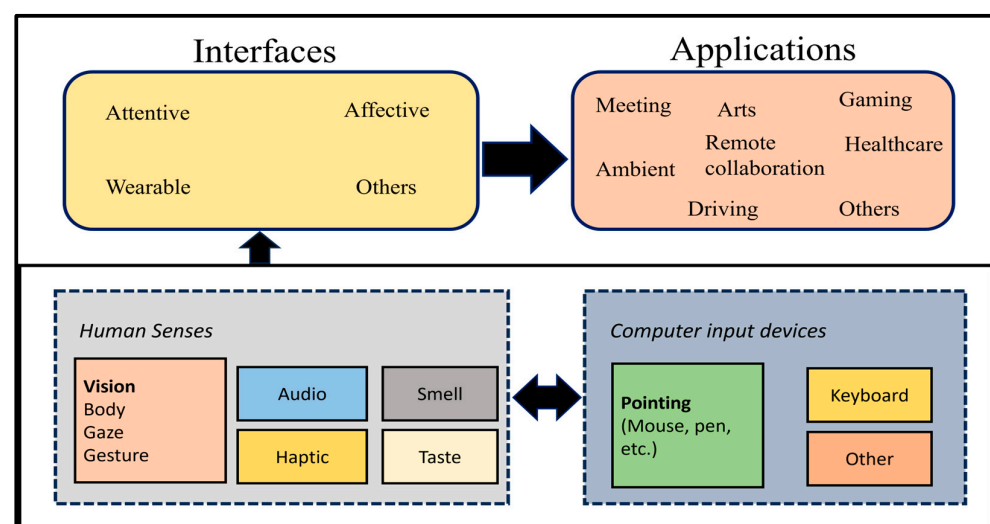
Figure 3. A diagram illustrating different types of interfaces and applications for human–computer interaction.

5. Results and Discussion

Aiming to transform human–computer interaction, this paper addresses a new system that precisely converts eye motions into mouse movements and effectively integrates gesture control techniques with voice recognition. During concentration tasks, the system tracks and measures user eye movements using the Tobii Eye Tracker 4C. Using infrared light and a camera, this eye tracker records user reflections. After that, specialized software examines these reflections to ascertain the user's gaze direction. The system employs a commercial voice recognition engine like Google voice-to-text which, using machine learning algorithms, very precisely converts user speech into text for speech understanding. This lets users communicate vocally, giving orders or input without having to physically engage with the computer.

Based on pupil monitoring, the real-time artificial intelligence system tracks user gaze and cursor movement by including two cameras, electrical components, and infrared markers. To achieve the exact real-time monitoring of eye (pupils) movements/detection the system (model) use the probabilistic algorithm and screen position detection. The recording and validating of user eye movements made possible by the integration of the Tobii Eye Tracker 4C during single spot testing improve the dependability of the eye movement-detecting component. Empirical test participants said that the effective voice recognition and gesture control techniques helped them to engage in a variety of activities without difficulties.

Ten people participated in field testing, including activities like swiping through folders, tapping targets, and manipulating things. With participants reaching great accuracy and speed in task execution, the findings revealed that the system efficiently controlled the mouse cursor utilizing eye motions, vocal instructions, and hand gestures. Combining natural language processing, eye movement tracking, and hand gesture identification, this creative system presents a more inclusive, immersive, and quick means of human–computer interaction [20,21]. It has great potential to improve virtual reality applications, enhance gaming experiences, and provide comfort tools to handicapped people. All things considered, the research shows an interesting advancement that may make using computers more natural and accessible, therefore benefiting user comfort and efficiency.

6. Conclusions and Future Scope

Aiming to remove manual user control in everyday chores and improve engagement for those with limited mobility, this paper addresses cutting-edge technology for PC control utilizing eye motions, voice, and gestures. Through natural and intuitive media, the system has shown great accuracy in reading eye, facial, and spoken instructions, thereby greatly boosting user experience and independence.

The suggested system uses voice recognition, detects eye and face motions, and provides gesture-activated mouse and daylight-based eye gaze tracking. Testing and validation have proved their effectiveness in improving user experience and, hence, offering more freedom. The possibilities of the developed system will make several changes in the various areas, including entertainment, healthcare, and education. For instance, it might help impaired students make use of instructional programs and support medical personnel in identifying and treating diseases.

This paper emphasizes the promise of technology for touchless human–computer interaction in perhaps substituting conventional input devices for those with mobility problems. Without using a keyboard or mouse, the handless/contactless interfaces the user can engage and do various activities like messaging, web surfing, and TV viewing. Face identification, eye extraction, and real-time interpretation of eye blinks are among the many image processing techniques the system uses to discreet human–computer interface management to perform the task. Mouse movement is detected and controlled using a decision tree method for eye motion.

The development of the system has solved all the identified problems, proving its capacity to enhance independence and engagement. More sophisticated hand motions and improved spoken commands derived from powerful machine learning algorithms might constitute future enhancements. These improvements could let the system interface with other programs like productivity tools, email, and social networking. Trials with bigger user groups might also assist in improving the usability and efficiency of the system.

In summary, all things considered, the suggested technology presents a breakthrough means of human–computer interaction that would greatly help those with limited mobility.

In addition to making technology more accessible and user-friendly, it has the potential to alter the way in which people use computers.

References

1. Chen, X.-L.; Hou, W.-J. Gaze-Based Interaction Intention Recognition in Virtual Reality. *Electronics* **2022**, *11*, 1647. [CrossRef]
2. Pawłowski, M.; Wroblewska, A.; Sysko-Romańczuk, S. Effective Techniques for Multimodal Data Fusion: A Comparative Analysis. *Sensors* **2023**, *23*, 2381. [CrossRef] [PubMed]
3. Jacob, R.J.K.; Karn, K.S. Eye Tracking in Human-Computer Interaction and Usability Research. In *The Mind's Eye*; Hyönä, J., Radach, R., Deubel, H., Eds.; Elsevier: Amsterdam, The Netherlands, 2003; pp. 573–605. [CrossRef]
4. Jaafar, N.; Lachiri, Z. Multimodal Fusion Methods with Deep Neural Networks and Meta-Information for Aggression Detection in Surveillance. *Expert Syst. Appl.* **2023**, *211*, 118523. [CrossRef]
5. Bansal, H.; Khan, R. A Review Paper on Human Computer Interaction. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2018**, *8*, 53. [CrossRef]
6. Zhu, H.; Wang, Z.; Shi, Y.; Hua, Y.; Xu, G.; Deng, L. Multimodal Fusion Method Based on Self-Attention Mechanism. *Wirel. Commun. Mob. Comput.* **2020**, *2020*, 8843186. [CrossRef]
7. Wilms, L.K.; Gerl, K.; Stoll, A.; Ziegele, M. Technology Acceptance and Transparency Demands for Toxic Language Classification—Interviews with Moderators of Public Online Discussion Fora. *Hum.-Comput. Interact.* **2024**, *40*, 285–310. [CrossRef]
8. Céspedes-Hernández, D.; González-Calleros, J.M.; Guerrero-García, J.; Rodríguez-Vizzuett, L. Gesture-Based Interaction for Virtual Reality Environments Through User-Defined Commands. In *Human-Computer Interaction*; Agredo-Delgado, V., Ruiz, P.H., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2019; Volume 847, pp. 143–157. [CrossRef]
9. Razaque, A.; Frej, M.B.H.; Bektemyssova, G.; Amsaad, F.; Almiani, M.; Alotaibi, A.; Jhanjhi, N.Z.; Amanzholova, S.; Alshammari, M. Credit Card-Not-Present Fraud Detection and Prevention Using Big Data Analytics Algorithms. *Appl. Sci.* **2023**, *13*, 57. [CrossRef]
10. Eldhai, A.M.; Hamdan, M.; Abdelaziz, A.; Hashem, I.A.T.; Babiker, S.F.; Marsono, M.N. Improved Feature Selection and Stream Traffic Classification Based on Machine Learning in Software-Defined Networks. *IEEE Access* **2024**, *12*, 34141–34159. [CrossRef]
11. Aldughayfiq, B.; Ashfaq, F.; Jhanjhi, N.Z.; Humayun, M. A Deep Learning Approach for Atrial Fibrillation Classification Using Multi-Feature Time Series Data from ECG and PPG. *Diagnostics* **2023**, *13*, 2442. [CrossRef] [PubMed]
12. Javed, D.; Jhanjhi, N.Z.; Khan, N.A. Explainable Twitter Bot Detection Model for Limited Features. *IET Conf. Proc.* **2023**, *11*, 476–481. [CrossRef]
13. Rana, P.; Batra, I.; Malik, A. An Innovative Approach: Hybrid Firefly Algorithm for Optimal Feature Selection. In Proceedings of the 2024 International Conference on Electrical, Electronics and Computing Technologies (ICEECT 2024), Greater Noida, India, 29–31 August 2024; IEEE: Piscataway, NJ, USA, 2024. [CrossRef]

14. Sharma, A. Smart Agriculture Services Using Deep Learning, Big Data, and IoT (Internet of Things). In *Handbook of Research on Smart Farming Technologies for Sustainable Development*; IGI Global: Hershey, PA, USA, 2024; pp. 166–202. [CrossRef]

15. Sharma, A.; Kala, S.; Guleria, V.; Jaiswal, V. IoT-Based Data Management and Systems for Public Healthcare. In *Assistive Technology Intervention in Healthcare*; CRC Press: Boca Raton, FL, USA, 2021; pp. 189–224. [CrossRef]

16. Lankadasu, N.V.Y.; Pesarlanka, D.B.; Sharma, A.; Sharma, S. Security Aspects of Blockchain Technology. In *Blockchain Applications for Secure Smart Cities*; IGI Global: Hershey, PA, USA, 2024; pp. 259–277. [CrossRef]

17. Sharma, A.; Kala, S.; Kumar, A.; Sharma, S.; Gupta, G.; Jaiswal, V. Deep Learning in Genomics, Personalized Medicine, and Neurodevelopmental Disorders. In *Intelligent Data Analytics for Bioinformatics and Biomedical Systems*; Wiley: Hoboken, NJ, USA, 2024; pp. 235–264. [CrossRef]

18. Sharma, A.; Kumar, R. Recent Advancement and Challenges in Deep Learning, Big Data in Bioinformatics. In *Studies in Big Data*; Springer: Cham, Switzerland, 2022; Volume 105, pp. 251–284. [CrossRef]

19. Aldughayfiq, B.; Ashfaq, F.; Jhanjhi, N.Z.; Humayun, M. YOLO-Based Deep Learning Model for Pressure Ulcer Detection and Classification. *Healthcare* **2023**, *11*, 1222. [CrossRef] [PubMed]

20. Aherwadi, N.; Mittal, U.; Singla, J.; Jhanjhi, N.Z.; Yassine, A.; Hossain, M.S. Prediction of Fruit Maturity, Quality, and Its Life Using Deep Learning Algorithms. *Electronics* **2022**, *11*, 4100. [CrossRef]

21. Ray, S.K.; Sinha, R.; Ray, S.K. A Smartphone-Based Post-Disaster Management Mechanism Using WiFi Tethering. In Proceedings of the 2015 10th IEEE Conference on Industrial Electronics and Applications (ICIEA 2015), Auckland, New Zealand, 15–17 June 2015; pp. 966–971. [CrossRef]