IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Vocal Gaze Mouse Controller

Dr. C. Nandini, Professor & Head, CSE Department, Dayananda Sagar Academy Of Technology And Management,
Bengaluru, Karnataka, India.
Yamini G, Assistant Professor, CSE Department, Dayananda Sagar Academy Of Technology And Management,
Bengaluru, Karnataka, India.
Samarth Srivastava, **Shivam Anand**, **Siddharth Kumar**
Student, CSE Department, Dayananda Sagar Academy Of Technology And Management,
Bengaluru, Karnataka, India.

Abstract: Facial gesture-based cursor control presents a novel, hands-free interaction system that leverages computer vision and machine learning to translate facial expressions into cursor movements. This system utilizes algorithms such as Convolutional Neural Networks (CNNs) to identify facial landmarks, while classifiers like Support Vector Machines (SVMs) or k-Nearest Neighbors (k-NN) interpret gestures, including blinking, smiling, and raising eyebrows, to execute actions like cursor navigation, clicks, and scrolling. Additionally, integrated voice commands provide a seamless way to articulate mouse actions, enhancing the system's functionality. By analyzing real-time video input, this technology delivers an intuitive and accessible interface, enabling gesture-driven control for individuals with mobility challenges and offering an innovative alternative for hands-free interaction.

***Keywords*—** Computer Vision, Machine Learning, Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), Support Vector Machines (SVMs), Voice commands, Hands-free interaction

I. INTRODUCTION

The rapid advancements in computer vision and machine learning have paved the way for innovative human-computer interaction systems. Among these, facial gesture- based cursor control has emerged as a promising hands-free solution, enabling users to operate a computer using facial expressions and gestures. Such systems are particularly beneficial for individuals with physical disabilities, providing an accessible alternative to traditional input devices like a mouse or keyboard.

Facial gesture recognition employs computer vision techniques to detect and analyze facial landmarks. Algorithms like Convolutional Neural Networks (CNNs) identify these features with high precision, while machine learning classifiers such as Support Vector Machines (SVMs) or k-Nearest Neighbors (k-NN) interpret gestures, including blinking, smiling, or raising eyebrows. These gestures are then translated into specific cursor actions like movement, clicking, or scrolling.

This paper explores the development and implementation of a facial gesture-based cursor control system with voice command integration. It delves into the underlying algorithms, system architecture, and potential applications, highlighting its significance in promoting accessibility and offering an innovative interaction paradigm.

Human-Computer Interaction (HCI) has evolved significantly, aiming to create more intuitive, efficient, and accessible systems. One of the most promising advancements in this field is the development of hands-free technologies that allow users to interact with digital devices without traditional input tools like keyboards or mice. Eye-tracking systems have gained considerable attention due to their potential to enable seamless interaction, particularly for individuals with physical disabilities who face challenges using conventional input methods. By leveraging eye movements and blinks, these systems provide a natural and non-invasive way to control computers, opening new possibilities for accessibility usability.

II. LITERATURE REVIEW

A. Mouse Cursor Control with Eye Movement Using OpenCV and Machine Learning

Authors: Murali Manoharan, Jaya Krishna Alagappan Pachamuthu, Surya Ganapathy (2024)

For people with physical challenges, using a mouse can be very difficult. Our proposed solution uses eye movements to control the mouse. Eye gaze is a method where users control their computer with their eye movements, especially helpful for those who can't use touchscreens or traditional mice. This system uses a webcam and machine learning to improve the reliability, mobility, and usability of eye tracking for user-computer interaction, allowing users to control the cursor with just their eyes.

B. System Cursor Control Using Human Eyeball Movement

Authors: Dhanasekar J, Guru Aravindh K B, Kiren A S, Faizal Ahamath A (2024)

This paper develops a hands-free system for controlling a computer cursor using eyeball movements. Using image processing and machine learning algorithms in Python, the system translates eyeball movements and blinks into cursor movements and clicks. This system is accurate and user- friendly, improving computer accessibility for individuals with motor impairments. It has potential applications in gaming.

C. Human-Eye Controlled Virtual Mouse

Authors: Mr. Dhanaraju, Dr. Sreenivas Mekala, A. Harsha Vardhan Rao, CH. Pavan Kumar, R. Lokesh (2024)

In the digital era, hands-free computing is essential for individuals with disabilities, like quadriplegics. This paper introduces a Human-Computer Interaction system that uses eye movements (blinking, staring, squinting) to control the mouse cursor. The system is implemented using Python, OpenCV, NumPy, and basic camera technology, making it completely hands-free without requiring additional hardware or sensors.

D. Eye-Controlled Mouse Cursor for Physically Disabled Individuals

Authors: Mohamed Nasor, Mujeeb Rahman K K, Maryam Mohamed, Haya Ansari, Farida Mohamed (2018)

This paper presents a new algorithm for controlling a computer screen cursor using eye movements. By accurately detecting the position of the iris, the algorithm maps this to a specific position on the screen, enabling physically disabled individuals to control the cursor's movements (left, right, up, down) and perform actions like opening and closing files or applications through a clicking mechanism.

E. Eyeball Movement-Based Cursor Using Deep Learning

Authors: Shubham Rane, Nihal Salian, Mohit Shetty, Rohan Shinde, Swati Rane (2021)

This paper proposes a hands-free human-machine interaction system for better communication between humans and computers. Long hours of using a mouse can lead to wrist and hand issues. This system, designed for differently-able people, allows them to control the mouse with their eyes. Using deep learning and machine learning, the system accurately tracks eye movements to control the cursor. The system includes noise reduction filters to ensure smooth and accurate operation.

III. PROPOSED METHODOLOGY

The methodology for this research is designed to evaluate and enhance machine learning models for predicting cricket match outcomes by combination of systematic approaches.

A. Data Acquisition and Facial Landmark Detection

Utilize a standard webcam to capture live video feed. Ensure the video stream provides sufficient resolution and frame rate for accurate eye tracking. Initial process the input frames to convert them into grayscale for reducing computational complexity and enhancing feature detection. Employ a pre-trained model to identify key facial landmarks. Extract regions of interest corresponding to the eyes using specific landmark indices. Analyze the extracted Regions Of Interest to track the position of the eyes. Compute the Eye Aspect Ratio (EAR) to assess the opening and closing states of the eyes. Use threshold techniques to determine if the eyes are open or closed based on EAR values.

B. Gaze and Blink Detection

Monitor the EAR values over consecutive frames. Detect a blink when the EAR drops below a predefined threshold and returns to normal within a short duration. Implement temporal filtering to avoid false-positive blink detection due to noise. Divide the eye ROI into regions and analyze the distribution of pixel intensities or positions of the pupil. Calculate gaze direction (left, right, up, or down) by measuring relative shifts in the pupil's position. Map gaze directions to control the cursor or trigger specific actions.

C. Voice Command Integration

Speech Recognition Module:

- Integrate voice recognition using libraries like Speech- recognition or Google Speech-to-Text API.
- Recognize specific voice commands such as "left click" or "right click" to trigger corresponding mouse actions.

Error Handling:

- Implement noise filtering and keyword detection to minimize misinterpretation of commands.

D. System Integration and Real-Time Processing

Integrate all modules to ensure synchronous operation in real-time. Optimize computational performance using efficient algorithms and hardware acceleration. Implement error-handling mechanisms to maintain robustness in varying lighting and environmental conditions. Combine gaze detection and blink interpretation to control cursor movement and perform mouse actions. Optimize system latency to ensure real-time responsiveness by employing multi threading and efficient algorithms.

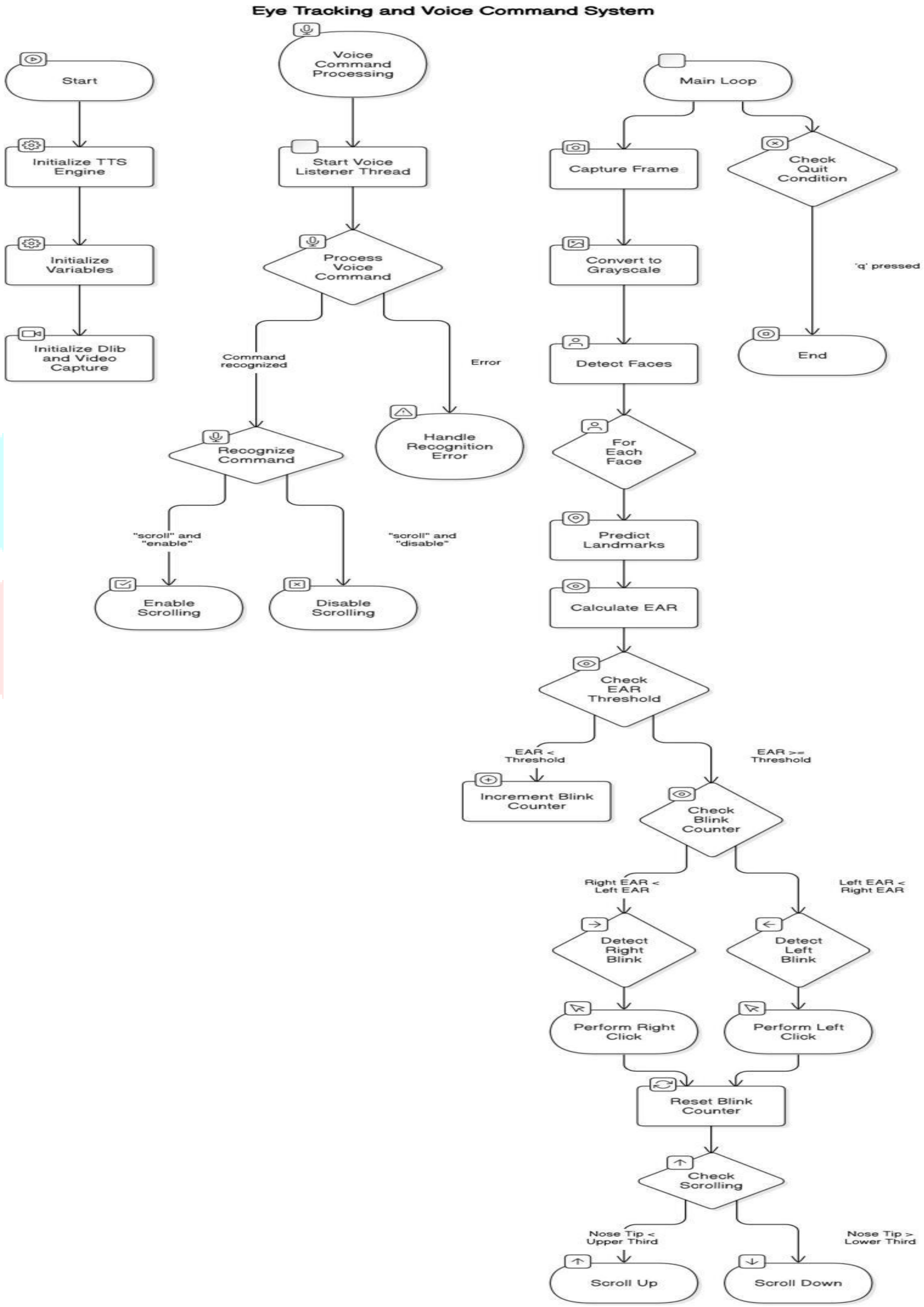
E. System Validation and Testing

Test the system under diverse conditions, including variations in lighting, user position, and background noise. Evaluate performance metrics such as accuracy, response time, and user satisfaction. Conduct usability tests with individuals, particularly those with physical disabilities, to assess the system's effectiveness and identify areas for improvement.

F. Implementation

The implementation of the real-time eye tracking and control system integrates advanced computer vision and voice recognition techniques to provide a hands-free interaction mechanism. Using Dlib's pre-trained facial landmark detection model, the system identifies key facial features, particularly around the eyes, by leveraging Histogram of Oriented Gradients (HOG) features. This enables precise tracking of eye movements and the calculation of the Eye Aspect Ratio (EAR), which is crucial for determining whether the eyes are open or closed. The EAR values, computed in real-time, form the basis of blink detection. When the EAR falls below a specified threshold for a predefined number of consecutive frames, the system recognizes it as a blink. Single and double blinks are mapped to distinct actions, such as scrolling up or down, offering intuitive control.

The gaze detection feature enhances functionality by determining the user's point of focus based on the relative positions of eye landmarks. By analyzing horizontal and vertical displacements of the iris, the system maps gaze directions to control the mouse cursor. Similarly, blinks are extended to include mouse control actions: left and right blinks simulate left and right mouse clicks, respectively. These actions are implemented using Python's pyautogui library, which interacts seamlessly with the operating system to execute the required events.



IV. FUTURE RESEARCH

The proposed eye tracking system presents numerous opportunities for future advancements and applications. One promising direction is enhancing accessibility for creating systems that adapt to natural human behaviors, offering an accessible and non-invasive solution for diverse applications. By addressing challenges like environmental variability and computational efficiency, this project lays the groundwork for future innovations in accessibility tools, virtual environments, automotive safety, and wearable technology. The findings underscore the transformative potential of eye tracking systems, fostering new possibilities for interaction in an increasingly digital world.

The potential of eye-tracking systems extends far beyond the current implementation, opening doors to a multitude of advanced applications and innovations. Future advancements in this technology could significantly impact various domains, fostering inclusivity, enhancing productivity, and enabling smarter human-computer interactions.

One promising direction is the development of wearable eye- tracking devices, such as smart glasses or augmented reality (AR) headsets, which could make the technology portable and more accessible.

These devices could enable users to interact seamlessly with both physical and virtual environments, revolutionizing industries like gaming, healthcare, and education. For instance, students could navigate interactive learning modules using gaze and voice commands, while healthcare professionals could benefit from hands-free control in surgical or diagnostic settings.

Integrating the system with artificial intelligence (AI) and deep learning models could enhance its accuracy, adaptability, and functionality. AI-powered gaze analytics could identify behavioral patterns, offering valuable insights into cognitive states, such as attention levels or emotional responses. This has applications in mental health monitoring, personalized advertising, and user experience optimization. Additionally, real-time emotion recognition through eye tracking could assist therapists in diagnosing and treating psychological conditions more effectively.

In automotive technology, the system could be expanded to monitor driver behavior, detect fatigue, and recognize distractions, contributing to safer road experiences. Advanced eye-tracking systems could also enable autonomous vehicles to understand and respond to passenger needs, providing a more personalized travel experience.

V. CONCLUSION

The proposed eye tracking system represents a significant advancement in human-computer interaction, offering an intuitive and accessible interface for users to interact with digital systems through natural gaze and blink behaviors. By combining real-time eye tracking, blink detection, gaze direction analysis, and voice commands, the system demonstrates a versatile approach to enabling seamless control over various functionalities, such as scrolling, mouse clicks, and voice-activated tasks. The integration of these features provides a foundation for applications in diverse domains, including accessibility tools for individuals with motor impairments, driver safety systems, and immersive experiences in virtual and augmented reality environments.

The project's emphasis on real-time processing, user- friendly design, and adaptability highlights its potential for practical deployment in everyday life. Furthermore, the scalability of the system allows for future enhancements, such as improved accuracy through advanced machine learning models, expanded functionality with emotion detection, and integration with wearable and IoT devices. This innovative framework not only addresses existing limitations in eye-tracking technologies but also opens avenues for further research and development. Ultimately, the system contributes to advancing assistive technologies and fostering a more inclusive digital ecosystem.

VI. REFERENCES

Priya Nandihal, Vijay Shetty, Tapas, Piyush Pareek " Giloma Detection Using Improved Artificial Neural Network in MRI Images", 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), 2022, pp 1-9. [1]

N. Kumar, P. Nandihal, M. R. B, P. K. Pareek, N. T and S. S. R, "A Novel Machine Learning-Based Artificial Voice Box," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), Bangalore, India, 2022, pp. 1-7, doi: 10.1109/ICATIECE56365.2022.10046967. [2]

Kumar, P. R., Meenakshi, S., Shalini, S., Devi, S. R., & Boopathi, S. (2023). Soil Quality Prediction in Context Learning Approaches Using Deep Learning and Blockchain for Smart Agriculture. In R. Kumar, A. Abdul Hamid, & D. Binti Ya'akub (Eds.), Effective AI, Blockchain, and E-Governance Applications for Knowledge Discovery and Management (pp. 1-26). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-6684-9151-5.ch001 [3]

P. Nandihal, P. K. Pareek, V. H. C. De Albuquerque, M. R. B, A. Khanna and V. S. Kumar, "Ant Colony Optimization based Medical Image Preservation and Segmentation," 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), Bangalore,India,2022,pp. 1-7,doi:10.1109/ICATIECE56365.2022.10047584. [4]

S A, K. ., Nandihal , P. ., K, S., D R , M. ., &Liyakathunisa. (2022). PRIOR DETECTION OF ALZHEIMER'S DISEASE WITH THE AID OF MRI IMAGES AND DEEP NEURAL NETWORKS. Malaysian Journal of Computer Science, 16–28. https://doi.org/10.22452/mjcs.sp2022no2.2 [5]

A. Padthe, M. Mathapati, P. M S and P. Nandihal, "APOA based Multi-scale Parallel Convolution Blocks with Hybrid Deep Learning for Gastric Cancer Prediction from Endoscopic Images," 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), Ballari, India, 2023, pp. 1-7, doi: 10.1109/AIKIIE60097.2023.10390430. [6]

N. Scott, C. Green, and S. Fairley, "Investigation of the use of eye tracking to examine tourism advertising effectiveness," Current Issues in Tourism, vol. 19, no. 7, pp. 634–642, 2016. [7]

K. Takemura, K. Takahashi, J. Takamatsu, and T. Ogasawara, "Estimating 3-D point-of-regard in a real environment using a head-mounted eye-tracking system," IEEE Transactions on Human-Machine Systems, vol. 44, no. 4, pp. 531–536, 2014.[8]

O. Ferhat and F. Vilarino, "Low cost eye tracking: the current panorama," Computational Intelligence and Neuroscience, vol. 2016, Article ID 8680541, pp. 1–14, 2016.[9]

M. A. Eid, N. Giakoumidis, and A. El Saddik, "A novel eye- gaze-controlled wheelchair system for navigating unknown environments: case study with a person with ALS," IEEE Access, vol. 4, pp. 558–573, 2016. [10]

L. Sun, Z. Liu, and M.-T. Sun, "Real time gaze estimation with a consumer depth camera," Information Sciences, vol. 320, pp. 346–360, 2015. [11]

Nagaraj, G., & Channegowda, N. "Video Forgery Detection using an Improved BAT with Stacked Auto Encoder Model". Journal of Advanced Research in Applied Sciences and Engineering Technology, 42(2), 175–187.(2024). https://doi.org/10.37934/araset.42.2.175187 Scopus indexed, SJR Q2. [12]

Nandihal, Priya; Pareek, Piyush Kumar; Suhas, G K; Deepak, N R; Bhagappa.” The Clinical and Public Health Impact of The Combined Work of hysicians and AI Systems” NeuroQuantology; Bornova Izmir Vol. 20, Iss. 16, (2022):2154-2163. DOI:10.48047/NQ.2022.20.16.NQ88214 [13]

Reddy, Shiva & Channegowda, Nandini. (2021). Edge Boost Curve Transform and Modified ReliefF Algorithm for Communicable and Non Communicable isease Detection Using Pathology Images. International Journal of Intelligent Engineering and Systems. 14. 463-473.10.22266/ijies2021.0430.42. [14]

Manjunath, D. R., Selva Kumar S., Sai Shiva Sumanth Reddy, and Lohith J. J. "Enhancing Personalized Learning Based on AI-Driven Lesson Plan Generator Using Mistral- 7B for Efficient Content Extraction and Summarization." Gradiva Journal, June 2024, https://gradiva.it/june-2024/ [15]