AirWave: Hands-Free Cursor Navigation with Face and Voice

Thippeswamy G[1]; Akash R[2]; Ankit Suresh Savalagi[3]; Dayanidhi M K[4]; Dileep R[5]

[1]Professor & HoD,
[1,2,3,4,5]Department of Computer Science and Engineering, BMS Institute of Technology and Management, Bengaluru-560064, India

Publication Date: 2025/06/05

Abstract:

➢ ***Context:***

Hands-free cursor navigation is considered indispensable for improving accessibility for physically disabled individuals. Currently, the control of mouse actions is based on facial gestures and voice commands in most existing systems. However, these solutions often face constraints such as a high sensitivity to environments, user fatigue, and reliance on advance hardware. Therefore, the following discussion will describe a lightweight, yet scalable system optimized for real-world conditions.

➢ ***Objectives:***

The project AirWave develops a hands-free cursor navigation system through facial gestures and voice commands offline. This is focused on real-time performance while minimizing the requirements for hardware on the system, thus becoming accessible and user-friendly for people with disabilities. This application aims at relating facial gestures like head tilts and blinks to cursor actions and integrating voice commands for advanced controls.

➢ ***Method:***

The system uses Dlib for facial landmark detection, OpenCV for video processing, and Vosk for offline speech recognition. The system reads in real-time video input from a webcam, detects facial gestures, and maps these to cursor movements using PyAutoGUI. Voice commands predefined trigger mouse actions, such as opening applications or scrolling.

➢ ***Result:***

The study conducted on the research papers provided critical insights into the current advancements and limitations of hands-free cursor navigation systems. The findings highlight the need for larger datasets and more sophisticated models to improve the accuracy.

➢ ***Conclusion:***

AirWave demonstrates the potential for accessible hands-free computing by addressing environmental sensitivity and hardware constraints. It provides a scalable, efficient solution, with future scope in multilingual commands, adaptive recognition, and IoT integration

Keywords: *Facial Gesture Recognition, Voice Command Integration, Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), Contrast Limited Adaptive Histogram Equalization (CLAHE).*

How to Cite: Thippeswamy G; Akash R; Ankit Suresh Savalagi; Dayanidhi M K; Dileep R. (2025). AirWave: Hands-Free Cursor Navigation with Face and Voice. *International Journal of Innovative Science and Research Technology*, 10(5), 3334-3344. https://doi.org/10.38124/ijisrt/25may1183.

I. INTRODUCTION

The past two decades have undergone a paradigm shift in Human-computer interaction regarding accessibility and accessibility. Even a keyboard and a mouse remain largely inaccessible to an individual who faces physical disabilities in his/her movement. Such movements led to developing hands - free systems that included facial gestures and eye - tracking through voice commands with the needful support to user groups of users with motor disorders.

Facial gesture-based systems have become a focal point in this domain. Techniques like Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) effectively detect eye blinks and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

mouth movements, which are then mapped to mouse actions. For instance, studies such as "Hands-Free Mouse Control Using Facial Feature" (2024) [1] and "Mouse Cursor Control Using Facial Movements" (2022) [3] report accuracies of up to 91% in cursor navigation. These systems rely on tools like Dlib, OpenCV, and MediaPipe for real-time facial landmark detection. However, their effectiveness diminishes in complex environments due to lighting dependencies and limited hardware adaptability

Hybrid systems, combining gestures with voice commands, exhibit promise in delivering multimodal interaction. "Gesture-Driven Virtual Mouse with a Voice Assistant" (2023) [6] exemplify this approach by integrating facial gesture detection with voice-controlled commands. These systems, however, face challenges in synchronizing multimodal inputs effectively and maintaining user comfort during prolonged use.

Voice-based systems enhance accessibility by leveraging natural language processing to interpret commands. Offline tools like Vosk ensure functionality without internet reliance, as demonstrated in "Computer Navigation Using Audio and Video Aid" (2022) [7]. Despite their advantages, these systems are prone to errors in noisy environments and may require additional hardware for improved accuracy

This system, based on these improvements and addressing all the critical gaps found in prior work. The adaptive algorithms for facial gesture recognition and offline voice recognition enable it to be more robust in real-time performance. Optimized for use with standard webcams and microphones, the design reduces hardware dependency. AirWave supports gestures like head tilts for cursor movement, blinks for clicks, and mouth gestures for scrolling, among others, besides customizable voice commands for mode switching, application control, and so on. This paper integrates insights from research studies, providing a comprehensive review of existing systems and methodologies. It points out the challenges and limitations that persist and introduces the AirWave system as a scalable, lightweight, and inclusive solution for hands-free cursor navigation.

A. Purpose of the Survey

The main purpose of this survey is to explore advancements in hands-free cursor navigation systems by analysing methodologies from 51 research papers. It assesses the effectiveness of facial gestures and voice commands in HCI systems, identifies existing challenges, and underlines gaps like environmental sensitivity, hardware dependency, and user fatigue. This review is to inform the design of the proposed AirWave system to deliver a lightweight, scalable, and inclusive solution for accessible computing.

B. Limitations and Open Challenge

Despite the great advancements in hands-free cursor control systems, there are still many limitations. Most of the systems require controlled lighting conditions, which makes them not adaptable to real-world environments. High hardware dependency, such as advanced cameras and sensors, increases the cost and limits accessibility. Gesture-based systems often cause user fatigue due to prolonged facial or head movements. Voice command systems also suffer from noise sensitivity and may require internet connectivity for real-time processing. There exist issues of synchronizing multimodal inputs within existing hybrid systems, and performance in various environments is not yet consistent.

C. Motivation

The motivation for this research is the call to bridge some of the access gaps in human-computer interaction of people with impairments. Despite the promise the existing systems possess, they continue to fall back in real usability challenges. Integrating adaptive algorithms and using hardware standards, AirWave aims at overcoming these barriers and providing accessibility, cost, and efficiency into seamless interaction among diverse environments.

D. Contributions

➢ *This Project has made Several Key Contributions:*

- Integrating adaptive facial gesture recognition and offline voice commands.
- Utilization of lightweight tools like OpenCV, Dlib, and Vosk for real-time performance on minimal hardware.
- Introduction of customizable gestures and voice commands for enhanced functionality.
- Optimization techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE) for diverse environments.
- Evaluation metrics for gesture accuracy, command latency, and user satisfaction, ensuring a robust and scalable design.

II. RELATED WORKS

This section compares methodologies and findings from relevant studies, analysing their contributions and relevance to the base paper, "A Comprehensive Review of Face Recognition Techniques, Trends, and Challenges" by H. L. Gururaj.

Mouse Cursor Control Using Facial Movements focuses on pose estimation through facial landmarks and gaze detection to enable cursor movement and clicks. Using a custom dataset, it achieved 91% accuracy for cursor control. The methodology's reliance on facial landmarks aligns closely with AirWave's use of Dlib for similar gesture mappings, reinforcing the robustness of gesture-based navigation.

Mouse Cursor Movement and Control Using Eye Gaze [5] with gaze direction estimation and pupil detection, this system achieved 90% accuracy in cursor navigation. Its emphasis on dynamic calibration for user preferences aligns with AirWave's adaptability in gesture threshold tuning for diverse users.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

Gesture and Voice Controlled Virtual Mouse for Elderly People [20], combining CNNs for gesture recognition, MediaPipe for hand tracking, and NLP for voice commands, this system provided an inclusive interface for elderly users. Although tailored for hand gestures, its focus on multimodal integration is relevant to AirWave's hybrid approach, which merges facial gestures and voice commands to ensure accessibility.

BLINK-CON [26]: A Hands-Free Mouse Pointer employed HOG and SVM for blink detection to simulate mouse clicks. AirWave adopts a similar mechanism, using EAR to detect blinks for left and right clicks, ensuring precision and efficiency in gesture recognition.

Touchless Head-Control (THC) [27]: Head Gesture Recognition utilizes CNNs for head pose estimation and YOLOv4 for deep neural network that detects face in images and videos. Cursor actions are triggered by head movements with Kalman Filters ensuring smooth trajectories. The methodology informs AirWave's head movement mapping for cursor navigation, emphasizing accurate and stable control.

Mouse Cursor Control Using Facial Movements [36]: This paper presents a system for controlling mouse cursors using facial gestures, leveraging Support Vector Machines (SVM), Convolution Neural Networks (CNN), and Histogram of Oriented Gradients (HOG). It achieves high accuracy by detecting facial landmarks and using Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) for gesture recognition.

Table 1: Comparison of Studies, Methodology & Results

Study	**Algorithms used**	**Proposed Methodology & Results**
Title: Hands-Free Mouse Control Using Facial Feature [1] Publication: IEEE Conference Publication (2024)	Facial key point detection using machine learning. Touchless Head-Control (THC) for cursor control.	Accuracy: The paper mentions an 87% accuracy in detecting facial movements for mouse control. Proposed Methodology: The system captures video input via a standard camera. Facial key points are detected to map head and facial movements to cursor actions. Touchless head control is utilized to simulate mouse functions like clicks and drags.
Title: Gesture Based Mouse Control [2] Publication: IEEE Conference Publication (2018)	Hand gesture recognition using edge detection and motion analysis. Mouse action simulation through body movement tracking.	Accuracy: Gesture recognition achieved an accuracy of 85%. Proposed Methodology: Webcam captures the user's gestures in real-time. Edge detection identifies hand or body movements, mapped to mouse functions. Continuous gesture monitoring allows dynamic cursor movement and actions like clicks.
Title: Mouse Cursor Control Using Facial Movements [3] Publication: IEEE Conference Publication (2022)	Pose estimation via facial landmarks and head movement tracking. Gaze detection for click simulation.	Accuracy: The system achieved an accuracy of 91% for cursor movement and 89% for click recognition. Proposed Methodology: Video input captures facial movements using a webcam. Pose estimation maps head movements to cursor actions. Gaze-based triggers are used for mouse clicks and context-sensitive actions.
Title: Identifying Facial Gestures to Emulate a Mouse [4] Publication: IEEE Conference Publication (2017)	Haar cascades for face detection. Aspect ratio thresholds for eye blinks and facial gesture recognition.	Accuracy: 94% accuracy for right-click detection and 92% for left-click actions. Proposed Methodology: Haar cascades detect faces and isolate eye regions. Eye aspect ratio (EAR) is calculated to identify blinks for clicks. Cursor movement is based on detected facial gestures and mapped to screen coordinates.
Title: Mouse Cursor Movement and Control Using Eye Gaze [5] Publication: IEEE Conference Publication (2023)	Eye gaze tracking using Hough Transform for pupil detection. Gaze direction estimation for cursor navigation.	Accuracy: The system reported a 90% accuracy for gaze detection and cursor control. Proposed Methodology: The webcam captures eye movement, processed using gaze detection algorithms.

		Cursor navigation is based on gaze direction and speed calibrated to user preferences. A dwell time mechanism is used for simulating mouse clicks.
Title: Gesture-Driven Virtual Mouse with a Voice Assistant [6] Published in: 2023 6th International Conference on Recent Trends in Advance Computing (ICRTAC).	Gesture recognition using OpenCV and MediaPipe. Speech recognition integrated with system controls. Specific gesture-based mappings for cursor control (e.g., V-shape for cursor movement).	Accuracy: Plain background: 80%-95%. Complex background: 60%-67% depending on the task (e.g., cursor movement, clicks, scrolling). Proposed Methodology: Captures video frames and converts them to HSV color space. Uses MediaPipe for gesture recognition and landmarks detection. Integrates speech recognition to execute additional commands (e.g., Google search). Evaluates performance in different environmental conditions (plain and complex backgrounds).
Title: Computer Navigation Using Audio and Video Aid for Amputees and Parkinson's Patients [7] Published in: 2022 5th International Conference on Advances in Science and Technology (ICAST).	Dlib library with HOG (Histogram of Oriented Gradients) for facial tracking. Vosk speech recognition toolkit for offline voice commands.	Proposed Methodology: Combines facial tracking with voice commands to control cursor movement and mouse actions. Facial movements (e.g., head turns) control the cursor, while voice commands trigger clicks. Avoids false positives by using speech commands instead of blinks for clicks. Ensures low-cost implementation without additional hardware requirements.
Title: Comparative Analysis of Hands-Free Mouse Controlling Based on Face Tracking [8] Published in: 2021 13th International Conference on Information & Communication	HOG (Histogram of Gradients) and Haar Cascade for face tracking. Use of EAR (Eye Aspect Ratio) and MAR (Mouth Aspect Ratio) to detect eye blinks and mouth movements for cursor control.	Proposed Methodology: Face detection using both HOG and Haar Cascade to analyze cursor control performance. Maps facial movements (e.g., head turns) to cursor directions and eye blinks for clicks. Compares the two algorithms in terms of time efficiency and participant performance using ANOVA analysis.
Title: Virtual Mouse using Hand and Eye Gestures [9] Publication Information: Published in the 2023 International Conferences on Data Science, Agents, and Artificial Intelligence (ICDSAAI). IEEE.	Hand Gestures: MediaPipe's hand-tracking module combined with OpenCV for gesture recognition, and PyAutoGUI for controlling the mouse pointer. Eye Gestures: OpenCV's face and eye detection functions for tracking eye movements, MediaPipe FaceMesh for face landmarks, and PyAutoGUI for mouse functions	Proposed Methodology: The proposed system involves: Using OpenCV and MediaPipe to track hand and eye gestures. For hand gestures: Detecting the index finger's position, calculating distances between landmarks, and mapping these to mouse actions. For eye gestures: Detecting eye blinks and calculating gaze positions to map cursor movements and clicks. A simple GUI built with Python's Tkinter to toggle between hand and eye gesture modes. This method enables a cost-effective, efficient virtual mouse that eliminates the need for additional hardware like colored markers or IR pens.
Title: An Analysis on Virtual Mouse Control using Human Eye [10] Publication Information: Published in the 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN). IEEE.	Haar Cascade Classifier for face and eye detection. Support Vector Machine (SVM) for classification of eye movements. Eye-Aspect-Ratio (EAR) for blink detection. Dlib model for real-time face landmark detection.	Haar Classifier: 85% precision, 76% accuracy SVM: 89% precision, 78% accuracy. Proposed Method: 94% precision, 89% accuracy. Proposed Methodology: The system preprocesses live video to detect facial landmarks and calculate eye and mouth aspect ratios for cursor movement and clicks.
Title: Mouse Cursor Control with Eye Gestures [11] Publication Information:	Haar Cascade for eye and face detection. Eye-Aspect-Ratio (EAR) and Mouth-Aspect-Ratio (MAR) for	The system uses a webcam to capture real-time video. Eye and mouth movements are tracked using computer vision techniques to control the cursor.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

Published in the 7th International Conference on Inventive Computation Technologies (ICICT 2024), IEEE.	detecting blinks and mouth movements. Euclidean distance calculations for cursor control and clicks.	Cursor Movement: Determined by the position of the nose tip. Clicks: Detected using EAR for left and right eye blinks. Double Click and Scrolling: Triggered by specific mouth and eye gestures. Calibration and testing ensure personalized and robust performance.
Title: Face Gesture Based Virtual Mouse Using MediaPipe [12] Publication Information: Published in the IEEE 8th International Conference for Convergence in Technology (I2CT 2023).	MediaPipe for facial landmark detection (468 landmarks). Mouth-Aspect-Ratio (MAR) and Eye-Aspect-Ratio (EAR) for mouse operations. OpenCV for image processing.	Proposed Methodology: The system uses MediaPipe to identify facial landmarks and control cursor movements: Cursor Control: Based on nose tip movement. Mouse Operations: Left-click, right-click, double-click, and scrolling mapped to mouth and eye movements. The system leverages MediaPipe's precision to overcome the limitations of earlier models.
Title: Human Computer Interaction Based Eye-Controlled Mouse [13] Publication Information: Presented at the 3rd International Conference on Electronics Communication and Aerospace Technology (ICECA 2019), IEEE.	Histogram of Oriented Gradients (HOG) for feature extraction. EAR and MAR for detecting blinks and yawning. OpenCV and Dlib for image processing.	The system detects facial features and interprets movements: Cursor Activation: Initiated by opening the mouth (MAR threshold). Scrolling and Clicking: Determined by squinting and blinking actions. The approach emphasizes accessibility for individuals with physical impairments.
Title: Cursor Control Based on Eyeball Movement Using Deep Learning [14] Published in: 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS).	Deep learning for gaze classification. HOG (Histogram of Oriented Gradients) for face detection. Accuracy: Not explicitly mentioned but claims high accuracy in eyeball movement detection.	Uses a webcam to capture eyeball movements. Converts RGB frames to grayscale and extracts eye area for analysis. Tracks eye movements and maps them to cursor movements on the screen. PyAutoGUI is used for cursor actions.
Title: Computer Mouse Control Using Iris Tracking: An Accessible and Cost-Effective Approach for People with Mobility Disabilities [15] Published in: 2023 42nd IEEE International Conference of the Chilean Computer Science Society (SCCC).	OpenCV and MediaPipe for iris tracking. PyAutoGUI for cursor control.	Proposed Methodology: Detects iris and tracks movement using MediaPipe's facemesh. Proportional calculations map the iris position to cursor movement. Blinking is detected for mouse clicks using distance metrics between eyelids.
Title: Mouse Cursor Controlled by Eye Movement for Individuals with Disabilities [16] Published in: 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS).	OpenCV for facial and eye feature detection. MediaPipe for eye-tracking and gaze direction.	Tracks eye movements to control the cursor. Utilizes PyAutoGUI for translating gaze into actions (clicks, scrolls). Employs blink detection for cursor commands.
Title: Controlling Mouse Motions Using Eye Tracking Using Computer Vision [17] Published: Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)	Face Detection: Pre-trained dataset iBUG 300-W for facial landmarks (68 coordinates) using Dlib. Eye Detection: Eye Aspect Ratio (EAR) for flicker detection.	Detects facial features like eyes and mouth using Dlib Scrolling is activated by squinting both eyes, followed by head movement up or down Developed in Python with OpenCV, focusing on cost-effective implementation for physically disabled users.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

	Mouth Detection: Mouth Aspect Ratio (MAR) to identify if the mouth is open.	
Title: Control the Movement of Mouse Using Computer Vision Technique [18] Published: Proceedings of the Sixth International Conference on Electronics, Communication, and Aerospace Technology (ICECA 2022)	Face Mesh Algorithm: Converts RGB to grayscale to locate pixel intensity. Canny Edge Detection: Used to detect edges and calculate gradient intensity of the image. Deep Neural Network: To detect face and eye movement in real-time.	The system uses a webcam to capture the user's facial and eye movements. Converts the image into grayscale and applies Canny edge detection to identify edges. Eye movements determine the cursor's position: Moving the eye left/right controls the cursor's horizontal movement.
Title: EyeGaze Control: Enhancing Mouse Cursor Precision Through Eyeball Movements [19] Publication: 2024 IEEE Students Conference on Engineering and Systems (SCES)	Keypoints Algorithm Thresholding Contouring PyAuto GUI integration	Utilized OpenCV and Dlib libraries for real-time eye movement tracking. Processed eye region data to determine pupil position and translated it into cursor movements.
Title: Gesture and Voice Controlled Virtual Mouse for Elderly People Publication: 2024 IEEE International Conference on Networking and Communications (ICNWC) [20]	Convolutional Neural Networks (CNN) for gesture recognition. MediaPipe framework for hand and fingertip detection. Natural Language Processing (NLP) for voice command interpretation.	Used webcams for hand gesture detection and movement. CNN models trained to classify gestures. Integrated MediaPipe and OpenCV for gesture tracking. Implemented Proton for voice commands using pyttsx3.

Table 2: Key Metrics Comparison

Study	Methodology	Dataset	Accuracy %
Hands-Free Mouse Control Using Facial Feature (2024)	Touchless Head-Control (THC)	Custom video dataset	87%
Mouse Cursor Control Using Facial Movements (2022)	Facial landmarks and gaze detection	Custom dataset of head & eye movements.	91% (cursor), 89% (click recognition)
Eye Gaze Controlled Virtual Keyboard (2019)	HoG and Shape Predictor	Real-time video input from standard webcams.	90.13%
Gesture-Driven Virtual Mouse (2023)	MediaPipe and OpenCV with offline speech recognition	Real-time video capture	80–95% (plain backgrounds); 60–67% (complex backgrounds)

➢ *Relevance to Project*

The analysis of 51 research studies provides critical insights into the design and development of the AirWave system. These studies explore facial gesture and voice command technologies, highlighting their potential for accessible computing. Techniques like Eye Aspect Ratio (EAR) [12] and Mouth Aspect Ratio (MAR) [17], frequently employed in facial gesture recognition, inform the selection of lightweight and efficient algorithms. Similarly, advancements in voice recognition frameworks, including offline solutions like Vosk, validate the feasibility of incorporating robust, real-time command processing.

Identifying challenges such as environmental sensitivity, hardware dependency, and user fatigue from prior systems shapes AirWave's focus on adaptive algorithms and minimal hardware requirements. The incorporation of tools like OpenCV and Dlib is inspired by their proven efficacy in existing research for accurate and real-time facial tracking. Moreover, the use of CLAHE for contrast enhancement directly addresses issues of lighting variability, a limitation frequently noted in the literature.

Additionally, these studies provide benchmarks for evaluating system performance, particularly in accuracy, latency, and user satisfaction. This foundational knowledge enables AirWave to optimize gesture and command execution while ensuring inclusivity and cost-efficiency. By addressing these gaps, AirWave integrates proven methodologies with novel solutions to advance hands-free cursor navigation systems.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

III. PROPOSED METHODOLOGY

The proposed AirWave system is designed to provide a robust and lightweight hands-free cursor navigation solution using facial gestures and voice commands. The following steps outline the methodology in detail:

The AirWave system integrates computer vision and offline voice recognition to perform cursor navigation and control. It is optimized for real-time performance on standard hardware configurations without requiring expensive peripherals or continuous internet connectivity.

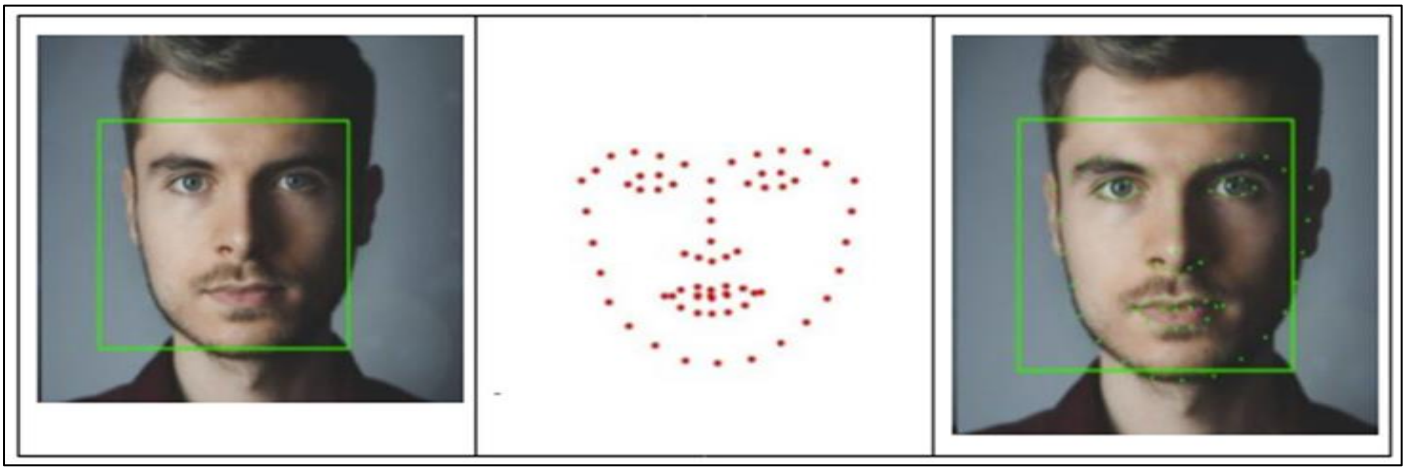

Fig 1: Landmark Detection

A. *Facial Gesture Recognition*

- Facial Landmark Detection: Using the Dlib library, 68 facial landmarks are extracted to track key regions such as the eyes, mouth, and head position.
- Gesture Parameters: Gestures are defined using the following metrics:

✓ Eye Aspect Ratio (EAR): Detects blinking for click actions.
✓ Mouth Aspect Ratio (MAR): Identifies mouth gestures for scrolling or zooming.
✓ Head Orientation: Head tilts are tracked for directional cursor movements using angles derived from the landmarks.

- Preprocessing: Images captured via the webcam undergo preprocessing with OpenCV techniques like grayscale conversion and Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance contrast under varying lighting conditions.
- Gesture Classification: Adaptive thresholds are employed to ensure gesture recognition remains accurate across diverse user profiles and environments.

B. *Voice Command Recognition*

- Offline Voice Processing: Vosk, an offline voice recognition toolkit, is used to ensure low-latency command execution without internet dependency.
- Command Vocabulary: The system recognizes predefined commands for actions like:

✓ Mode switching (e.g., toggling between gesture-only and voice-only modes).
✓ Application control (e.g., opening/closing software, volume adjustments).
✓ Text typing for basic input tasks.

- Noise Robustness: Noise filtering techniques are applied to improve recognition accuracy in different acoustic environments.

C. *Gesture-Command Mapping*

➢ *Each recognized facial gesture and voice command is mapped to specific cursor actions. Examples include:*

- Blinking of left and right eyelids for left and right clicks respectively.
- Opening the mouth for scroll up or down.
- Tilting the head for directional cursor movement.
- Voice commands like "Scroll up" or "Open Browser" for advanced functionality.
- Blink and hold for click, drag and drop function of the mouse pointer.

D. *Real-Time Execution*

- Cursor Movement and Control: PyAutoGUI handles real-time cursor navigation based on gesture inputs and executes mouse clicks and scrolls.
- Synchronization: A multimodal integration layer ensures smooth synchronization between gesture and voice inputs to avoid conflicts.

E. *Hardware Requirements*

The system is designed to operate efficiently on minimal hardware configurations, including standard webcams and microphones. No additional sensors or advanced processing units are required, making it accessible for a wide audience.

F. System Optimization

Lightweight algorithms are employed to minimize latency. Computational bottlenecks are addressed by optimizing image processing tasks and reducing the complexity of voice command recognition models.

G. Evaluation Metrics

➢ *The System's Performance will be Assessed using:*

- Gesture Recognition Accuracy: Precision and recall metrics for each gesture.
- Voice Command Latency: Time taken to recognize and execute commands.
- User Satisfaction: Surveys to gauge the ease of use and comfort.

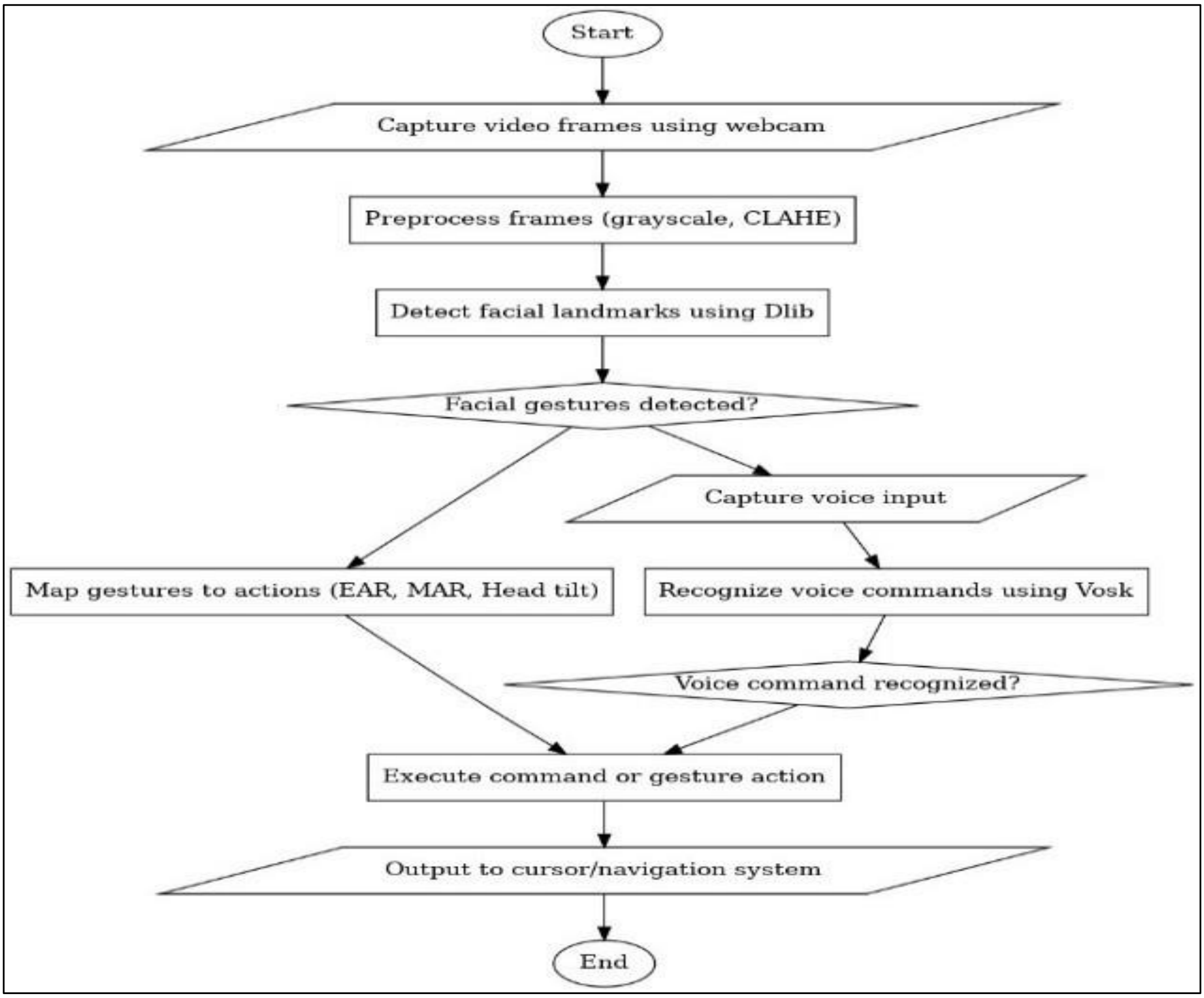

Fig 2: Flowchart of the Proposed Methodology

IV. PERFORMANCE AND EVALUATION PARAMETERS

Performance and evaluation are critical components in assessing the effectiveness and reliability of any research or project. In the context of Hands-Free cursor navigation with Face Gestures and Voice Commands, performance evaluation ensures that the proposed system meets its objectives in terms of accuracy, efficiency, and reliability.

A. Performance Parameters:

➢ *Gesture Accuracy: Measures the percentage of correct facial gesture detections for cursor navigation.*

- Metric: Precision, recall, and F1-score for gesture recognition.
- Evaluation Tools: Python libraries (scikit-learn) for classification report generation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

- *Voice Command Recognition Rate: Evaluates the accuracy of offline voice recognition for command execution.*

- Metric: Word error rate (WER) and command recognition accuracy.
- Evaluation Tools: Vosk performance analyser and Python-based speech evaluation modules.

- *Latency: Assesses the time delay between gesture or voice command input and corresponding system response.*

- Metric: Average response time for gesture and command execution (milliseconds).
- Evaluation Tools: Python time library and system profiling tools like Py-Performance.

- *System Resource Usage: Monitors CPU, memory, and GPU usage to ensure compatibility with minimal hardware.*

- Metric: CPU, GPU, and memory consumption during system operation.
- Evaluation Tools: Resource monitoring tools like psutil (Python) and Windows/Linux task managers.

- *User Fatigue: Evaluates the physical strain caused by prolonged use of facial gestures.*

- Metric: User-reported fatigue levels on a Likert scale and observation of sustained usage effects.
- Evaluation Tools: User feedback surveys and time-usage analysis.

B. Evaluation Parameters

- Accuracy Metrics: Precision and recall for gesture and voice recognition.
- Latency Benchmarking: Average delay time for various gestures and commands.
- Usability Testing: Surveys and interviews to assess user satisfaction and comfort.
- Environment Robustness: Performance across varying lighting conditions and noise levels.
- Scalability: System adaptability to diverse hardware configurations and user demographics.

V. RESULTS

The study conducted on the research papers provided critical insights into the current advancements and limitations of hands-free cursor navigation systems. Most systems demonstrated high accuracy in gesture recognition, with many achieving above 90% using algorithms like Haar Cascade [10], Eye Aspect Ratio (EAR) [12], Mouth Aspect Ratio (MAR) [13], and convolutional neural networks (CNNs) [20]. Like those systems, the voice recognition system utilized strong tools, including Vosk and Microsoft Speech SDK for robust command recognition rates. On the other hand, similar challenges of environment sensitivity, hardware dependency, and user fatigue were dominant among all analysed systems.

It also came out that the hybrid systems were trying to merge gesture and voice control but encountered problems in synchronizing these with real-world inconsistencies in performance, which makes their practicality somewhat limited. Such findings influenced the design of the AirWave system, where it solves these issues with adaptive algorithms and optimized resource usage.

The proposed methodology of AirWave will be tested against these insights using lightweight tools such as OpenCV, Dlib, and PyAutoGUI. The system will able to perform robustly under various conditions by incorporating features such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and offline voice recognition through Vosk. This approach not only addresses the major limitations identified in the reviewed studies but also provides real-time responsiveness and low hardware requirements, thus making it accessible to a larger population.

VI. CONCLUSION AND FUTURE SCOPE

This survey paper identifies the progress and limitations in hands-free cursor navigation systems by discussing the research studies. The findings from the review were used to design the AirWave system, which addresses challenges such as environmental sensitivity, hardware dependency, and user fatigue. Adaptive computer vision and offline voice recognition are leveraged to offer a lightweight, scalable, and inclusive solution for accessible computing that bridges critical gaps in current systems.

A. Future Scope

- *Extension to Wearable and IoT Devices*

The system can be extended to work on wearable devices such as AR/VR headsets or IoT-enabled environments. This would enable hands-free interaction in various domains like immersive gaming, smart home control, or industrial automation, broadening the application of gesture and voice-controlled systems.

- *Real-Time Emotion Detection for Context-Aware Interaction*

Researchers could integrate real-time emotion detection to modify the behaviour of the system based on the user's emotional state. For instance, frustration might prompt assistance features, and fatigue could suggest switching to voice commands to minimize physical strain. This would make the system more intuitive and user-centric.

- *Robustness Against Adverse Environmental Conditions*

Future work can improve the robustness of the system under extreme conditions, such as heavy background noise or poor lighting. Real-time algorithms for environmental adaptation, such as noise suppression on voice commands and lighting compensation by using advanced computer vision techniques.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

➢ *Accessibility to Multilingual Voice Commands*

Future research could then be on how to integrate support for multilingual voice commands that would cater for users from all linguistic backgrounds. Training the system with datasets of various languages and accents can therefore expand the scope of inclusiveness of hands-free navigation systems, particularly in regions that have limited access to English-centric technology.

REFERENCES

[1]. "Hands-Free Mouse Control Using Facial Feature," IEEE Conference Publication, 2024. [Online]. Available: IEEE XPLORE.

[2]. "Gesture Based Mouse Control," IEEE Conference Publication, 2018. [Online]. Available: IEEE XPLORE.

[3]. "Mouse Cursor Control Using Facial Movements," IEEE Conference Publication, 2022. [Online]. Available: IEEE XPLORE.

[4]. "Identifying Facial Gestures to Emulate a Mouse," IEEE Conference Publication, 2017. [Online]. Available: IEEE XPLORE.

[5]. "Mouse Cursor Movement and Control Using Eye Gaze," IEEE Conference Publication, 2023. [Online]. Available: IEEE XPLORE.

[6]. "Gesture-Driven Virtual Mouse with a Voice Assistant," 2023 6th International Conference on Recent Trends in Advance Computing (ICRTAC).

[7]. "Computer Navigation Using Audio and Video Aid for Amputees and Parkinson's Patients," 2022 5th International Conference on Advances in Science and Technology (ICAST).

[8]. "Comparative Analysis of Hands-Free Mouse Controlling Based on Face Tracking," 2021 13th International Conference on Information & Communication Technology and System (ICTS).

[9]. "Virtual Mouse Using Hand and Eye Gestures," 2023 International Conferences on Data Science, Agents, and Artificial Intelligence (ICDSAAI), IEEE. DOI: 10.1109/ICDSAAI59313.2023.10452550.

[10]. "An Analysis on Virtual Mouse Control using Human Eye," 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN), IEEE. DOI: 10.1109/ICIPCN63822.2024.00045.

[11]. "Mouse Cursor Control with Eye Gestures," 7th International Conference on Inventive Computation Technologies (ICICT 2024), IEEE.

[12]. "Face Gesture Based Virtual Mouse Using Mediapipe," IEEE 8th International Conference for Convergence in Technology (I2CT 2023).

[13]. "Human Computer Interaction Based Eye-Controlled Mouse," 3rd International Conference on Electronics Communication and Aerospace Technology (ICECA 2019), IEEE.

[14]. "Cursor Control Based on Eyeball Movement Using Deep Learning," 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS).

[15]. "Computer Mouse Control Using Iris Tracking: An Accessible and Cost-Effective Approach for People with Mobility Disabilities," 2023 42nd IEEE International Conference of the Chilean Computer Science Society (SCCC).

[16]. "Mouse Cursor Controlled by Eye Movement for Individuals with Disabilities," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS).

[17]. "Controlling Mouse Motions Using Eye Tracking Using Computer Vision," Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020), IEEE.

[18]. "Control the Movement of Mouse Using Computer Vision Technique," Proceedings of the Sixth International Conference on Electronics, Communication, and Aerospace Technology (ICECA 2022), IEEE.

[19]. "EyeGaze Control: Enhancing Mouse Cursor Precision Through Eyeball Movements," 2024 IEEE Students Conference on Engineering and Systems (SCES).

[20]. "Gesture and Voice Controlled Virtual Mouse for Elderly People," 2024 IEEE International Conference on Networking and Communications (ICNWC).

[21]. "Facial Movement and Voice Recognition Based Mouse Cursor Control," 2023 IEEE International Conference on Smart Electronics and Communication (ICOSEC).

[22]. "An Efficient Mouse Tracking System Using Facial Gestures," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS).

[23]. "Assisting the Differently Abled Person Using Eye Mouse," 2024 9th International Conference on Science Technology Engineering and Mathematics (ICONSTEM).

[24]. "Computer Cursor Control Using Eye and Face Gestures," 2020 11th ICCCNT.

[25]. "Eyeball-Based Cursor Movement Control," International Conference on Communication and Signal Processing, July 28-30, 2020, India.

[26]. "BLINK-CON: A Hands-Free Mouse Pointer Control with Eye Gaze Tracking," 2021 IEEE Mysore Subsection International Conference (MysuruCon).

[27]. "Touchless Head-Control (THC): Head Gesture Recognition for Cursor and Orientation Control," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 30, 2022.

[28]. "Facial Movements Based Mouse Cursor Control for Physically Disabled Individuals," International Journal of Engineering & Science Research (IJESR), Sep 2023, Vol-13, Issue-3.

[29]. "Hands-Free Gesture and Voice Control for System Interfacing," International Journal on Recent and Innovative Trends in Computing and Communication (IJRITCC), ICEMTE-2017.

[30]. "Hands-Free Gesture and Voice Control for System Interfacing," International Conference on Emanations in Modern Technology and Engineering (ICEMTE-2017).

[31]. "A Prototype System for Controlling a Computer by Head Movements and Voice Commands," Technical Report, 2017.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may1183

[32]. “A Overview on Designing of Hands-Free Mouse Pointer for Motor Impairment People Using Motion Tracking and Speech Recognition,” International Journal of Engineering Research & Technology (IJERT), June 2013, Volume 2, Issue 6.
[33]. “Hands-Free PC Control: Controlling of Mouse Cursor Using Eye Movement,” International Journal of Scientific and Research Publications, Volume 2, Issue 4, April 2012.
[34]. “Cursor Control with Facial Gestures Using CNN,” Technical Report, Tribhuvan University, May 2023.
[35]. “Face Gesture and Speech Based Virtual Mouse and Virtual Assistant,” International Journal of Creative Research Thoughts, Volume 12, Issue 5, May 2024.
[36]. “Mouse Cursor Control Using Facial Movements,” International Journal of Creative Research Thoughts (IJCRT), Volume 12, Issue 4, April 2024.
[37]. “Computer Cursor Tracking Using Eye Movement, Gesture Sign Language, and Voice Commands,” International Journal of Research Publication and Reviews, Volume 5, Issue 1, January 2024.
[38]. “Facial-Expression Based Mouse Cursor Control for Physically Challenged Individuals,” International Research Journal of Engineering and Technology (IRJET), Volume 10, Issue 4, April 2023.
[39]. “Mouse Cursor’s Movements Using Voice Controlled Mouse Pointer,” International Journal of Computer Applications, Volume 71, Issue 7, May 2013.
[40]. “Low-Cost Human–Machine Interface for Computer Control with Facial Landmark Detection and Voice Commands,” Sensors 2022, Volume 22, Article 9279 (MDPI).
[41]. “Cursor Tracking by Sensory Organs for Handicapped People,” International Journal of Advance Research, Ideas, and Innovations in Technology, Volume 5, Issue 6, 2019.
[42]. “Human-Computer Interaction Based Head-Controlled Mouse,” MZUJHSS, Volume X, Issue 1, June 2024.
[43]. A Comprehensive Review of Face Recognition Techniques, Trends, and Challenges H. L. GURURAJ, 2 July 2024
[44]. "Multimodal Biometric Human Recognition for Perceptual Human-Computer Interaction," IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 40, no. 6, pp. 740–751, Nov. 2010.
[45]. "Comparing the Use of Single Versus Multiple Combined Abilities in Conducting Complex Computer Tasks Hands-Free," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 9, pp. 1736–1745, Sep. 2018.
[46]. "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 3, pp. 478–500, Mar. 2010.
[47]. "A Multi-Gesture Interaction System Using a 3-D Iris Disk Model for Gaze Estimation and an Active Appearance Model for 3-D Hand Pointing," IEEE Transactions on Multimedia, vol. 13, no. 3, pp. 513–526, Jun. 2011.
[48]. "Smart Wheelchair Based on Eye Tracking," in Proceedings of the Biomedical Engineering International Conference (BMEiCON), Laung Prabang, Laos, 2016, pp. 1–5.
[49]. "Facial Landmark-Based Cursor Control and Speech-to-Text System for Paralyzed Individuals," in Proceedings of the IEEE International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 1–8.
[50]. "Video Face Detection Based on Improved SSD Model and Target Tracking Algorithm," Journal of Web Engineering, vol. 21, no. 2, pp. 135–152, 2022.
[51]. "HeadTrack: Real-Time Human Computer Interaction via Wireless Earphones," IEEE Journal on Selected Areas in Communications, vol. 42, no. 4, pp. 1014–1025, Apr. 2024.