Eye Tracking for Communication in Locked-In Syndrome Patients

Dr. Soubhagyalakshmi P
Associate professor,
Department of CSE,
K. S. Institute of Technology,
BANGALORE, INDIA

Aishwarya G
Department of CSE,
K. S. Institute of Technology,
BANGALORE, INDIA
aishugjsa@gmail.com

Bhavana B
Department of CSE,
K. S. Institute of Technology,
BANGALORE, INDIA
bhavanababu03@gmail.com

Maya
Department of CSE,
K. S. Institute of Technology,
BANGALORE, INDIA
mayabakka@gmail.com

Arbeena Farheen
Department of CSE,
K. S. Institute of Technology,
BANGALORE, INDIA
arbeenafarheen@gmail.com

Abstract*— Individuals with mobility impairments, including those affected by conditions such as Locked-in Syndrome (LIS), ALS, or stroke, face significant challenges in communication due to the loss of voluntary muscle control. This research presents an advanced eye-tracking system that enables individuals with restricted mobility to communicate using eye blinks. The proposed system leverages computer vision, artificial intelligence, and machine learning techniques to detect and interpret eye movements, converting them into textual communication. By integrating Haar cascades and adaptive language modeling, the system ensures precise detection and real-time processing, even on low-end devices. Additionally, the study highlights challenges such as accuracy, affordability, and accessibility while proposing solutions to improve usability. The research aims to empower individuals with mobility impairments, enhancing their independence and fostering social inclusion through assistive technology.

Index Terms-Eye Tracking, Communication, Assistive Technology, Machine Learning, Mobility Impairment, Human-Computer Interaction.

I. INTRODUCTION

A. *Problem Statement*

Individuals with severe paralysis, especially those with Locked- in Syndrome (LIS), experience tremendous difficulties in communication as a result of the loss of voluntary control over muscles. The only consistent ways of interacting available to them are through eye movement and blinking. Present assistive devices are either too sophisticated, costly, or need care-assisted operation, hence constraining the independence of these individuals.

B. *Importance of the Problem*

Successful communication is paramount for emotional health, social participation, and life quality. Physical restrictions precluding expression of needs, thoughts, or feelings produce isolation and dependency. Creating a low-cost, accessible, and independent communication device with eye tracking bridges directly a major assistive technology gap.

Conventional methods of communication for people with severe motor disabilities usually depend very much on the availability of caregivers, which not only decreases user independence but also constrains spontaneous interaction. Further, most current assistive devices are either prohibitively costly or demand complicated setups, and thus they are not available to a major portion of the disabled community. This assistive technology democratization has the potential to revolutionize daily life for mobility-impaired individuals, promoting increased dignity and independence.

C. *Implemented System Summary*

Our system Blink to Text, is a camera-based communication aid for individuals with severe mobility impairments such as Locked-in Syndrome. It employs computer vision and machine learning to translate eye blinks into text. The system consists of:

- Face and eye detection- via Haar Cascades to detect eyes from webcam input.
- Blink detection- via facial landmarks and frame differencing to detect intentional blinks.
- A conversion engine- which translates blinks to text selection in a user-selectable interface.
- Language prediction- to predict the next word from blinking behavior.
- A GUI designed for ease of use- using PyQt5 to display camera feed, predicted text, and settings.
- It operates fluently on low-end PCs, thereby keeping it cost- effective and accessible for everyone.

D. *Related Work*

A number of studies have examined the use of eye-tracking in human-computer interaction (HCI). Ibrahim et al. introduced a low- cost Raspberry Pi and Haar cascades-based blink detection system with an emphasis on cost-effectiveness and ease of use. Kim et al. analyzed blink patterns for detecting deepfakes, illustrating how eye movement correlates with cognitive and physiological states. Acharjee and Deb proposed a Visual Tracking Technique (VTT)- based solution that maps facial and eye movement to interaction commands, improving HCI accessibility. Zhou tackled the problem of low illumination conditions by leveraging Media Pipe and Zero- DCE for stable blink detection, allowing the system to remain reliable in different lighting situations. Whereas these pieces of work point to the promise of eye-tracking technologies, Blink to Text is unique in providing a completely independent, low-cost, and accessible communication system tailored for people with Locked-in Syndrome, with a focus on independence and simplicity.

II. LITERATURE SURVEY

[1] Schwiegelshohn, F.; Wehner, P.; Rettkowski, J.; Gohringer, D.; Hubner, M.; Keramidas, G.; Antonopoulos, C.;Voros, N.S. A holistic approach for advancing robots in ambient assisted living environments. In Proceedingsof the 2015 IEEE 13th International Conference on Embedded and Ubiquitous Computing, Porto, Portugal,21–23 October 2015; pp. 140–147.

Explanation:

Objective of the Study: The paper presents a comprehensive (holistic) framework for integrating robotic systems in Ambient Assisted Living (AAL) environments, aimed at supporting elderly and disabled individuals in their daily lives.

Modular System Design: The authors propose a modular robotic system architecture, enabling adaptability to various tasks and user needs. This modularity helps in tailoring assistance levels depending on user

conditions.
Sensor Integration: The robot is equipped with multiple sensors (e.g., visual, audio, environmental), allowing it to perceive and understand its surroundings and the behavior of users more effectively.

[2] Konstantinidis, E.I.; Antoniou, P.E.; Bamparopoulos, G.; Bamidis, P.D. A lightweight framework fortransparent cross platform communication of controller data in ambient assisted living environments.Inf. Sci. (NY) 2015,300, 124–139.

Explanation:
Purpose of the Study: This paper introduces a lightweight communication framework designed to enable transparent and efficient data exchange between various controllers in Ambient Assisted Living (AAL) systems.
Cross-Platform Compatibility: The framework supports interoperability across different hardware and software platforms, enabling devices with different architectures (e.g., PCs, tablets, sensors, embedded systems) to communicate seamlessly.
Transparent Communication Layer: The authors propose a transparent abstraction layer that hides the underlying communication complexity from the application developers, allowing them to focus on core functionality rather than platform- specific implementation.

[3] Boumpa, E.; Charalampou, I.; Gkogkidis, A.; Ntaliani, A.; Kokkinou, E.; Kakarountas, A. Assistive Systemfor Elders Suffering of Dementia. In Proceedings of the 2018 IEEE 8th International Conference on ConsumerElectronics, Berlin, Germany, 2–5 September 2018; pp. 1–4.

Explanation:
Objective of the Study: The paper presents the design and development of an assistive system specifically tailored for elderly individuals suffering from dementia, aiming to improve their independence and safety in everyday life.
Targeted User Needs: The system is user-centered, designed around the specific cognitive limitations and behavioral patterns of dementia patients, such as memory loss, confusion, and disorientation.
Multi-Component Architecture:
The system comprises various modules, including:

- Indoor navigation assistance
- Reminder and notification system
- Emergency alerting: These modules work together to provide comprehensive assistance.

[4] Brazil Assistive Technology. In Proceedings of the National Undersecretary for the Promotion of the Rights of Peoplewith Disabilities; Technical Assistance Committee: Geneva, Switzerland, 2009.

Explanation:
Purpose of the Document: The publication outlines Brazil's national strategy for promoting and implementing assistive technologies (AT) aimed at enhancing the independence and inclusion of people with disabilities.
Governmental Framework: It highlights the role of government institutions, particularly the National Secretariat for the Promotion of the Rights of Persons with Disabilities, in formulating inclusive public policies for assistive technology.
Policy Integration: The strategy integrates assistive technology within broader social policies such as education, health, employment, and social assistance to ensure comprehensive support for disabled individuals.

[5] Elakkiya, J.; Gayathri, K.S. Progressive Assessment System for Dementia Care Through Smart Home.In Proceedings of the 2017 International Conference on Algorithms, Methodology, Models and Applicationsin Emerging Technologies (ICAMMAET), Chennai, India, 16–18 Febuary 2017; pp. 1–5. 45

Explanation:
Objective of the Study: The paper proposes a smart home-based progressive assessment system for monitoring and supporting individuals with dementia, focusing on enhancing in-home safety and personalized care.
Continuous Monitoring: The system is designed to continuously monitor patient behavior and activities using embedded sensors and smart devices, aiming to detect cognitive decline or unusual patterns early.
Behavioral Pattern Analysis: It leverages data analytics to model behavioral patterns, identifying deviations that may indicate potential health issues, confusion, or worsening symptoms.

[6] Rafferty, J.; Nugent, C.D.; Liu, J.; Chen, L. From Activity Recognition to Intention Recognition for AssistedLiving within Smart Homes. IEEE Trans. Hum. - Mach. Syst. 2017,47, 368–379.

Explanation:
Objective of the Study: The paper explores the transition from activity recognition to intention recognition in smart home environments, aiming to provide proactive assistance to elderly or dependent users.
Beyond Activity Recognition: Traditional systems detect "what" the user is doing (activity recognition), but this work introduces intention recognition, which infers the "why" behind the action — enabling more intelligent assistance.
Context-Aware Computing: The system integrates contextual information such as time, location, sequence of actions, and environmental states to understand user goals and intentions more effectively.

[7] Mizumoto, T.; Fornaser, A.; Suwa, H.; Yasumoto, K.; Cecco, M. De Kinect- based micro-behavior sensingsystem for learning the smart assistance with human subjects inside their homes. In Proceedings of the 2018Workshop on Metrology for Industry 4.0 and IoT, Brescia, Italy, 16–18 April 2018; pp. 1–6.

Explanation:
Objective of the Study: This paper presents a Kinect-based micro-behavior sensing system aimed at analyzing fine-grained human behaviors in a home environment to enhance smart assistance.
Use of Kinect Sensor: The system uses Microsoft Kinect, which provides depth and skeleton tracking, to capture detailed body movements and non-verbal cues of users in real time.
Focus on Micro-Behaviors: Unlike broader activity recognition, this system emphasizes micro-level actions such as gestures, posture shifts, hand movements, and fine motor activities — critical in early detection of cognitive decline.

[8] Daher, M.; El Najjar, M.E.; Diab, A.; Khalil, M.; Charpillet, F. Multi sensory Assistive Living System forElderly In-home Staying. In Proceedings of the 2018 International Conference on Computer and Applications(ICCA), Beirut, Lebanon, 25–26 August 2012; pp. 168–171.

Explanation:
Objective: The primary goal of the paper is to design a multi-sensory assistive living system to support elderly individuals, enabling them to live independently in their homes while ensuring safety and health.
System Design: The system integrates various sensors (e.g., motion sensors, temperature sensors, etc.) to monitor the health, well-being, and activity levels of the elderly user.
Multi-sensory Integration: The system combines various sensory inputs such as visual, auditory, and tactile stimuli to create a comprehensive solution for monitoring and interaction.

[9] Ghayvat, H.; Mukhopadhyay, S.; Shenjie, B.; Chouhan, A.; Chen, W. Smart Home Based Ambient AssistedLiving Recognition of Anomaly in the Activity of Daily Living for an Elderly Living Alone. In Proceedingsof the 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Houston,TX, USA, 14– 17 May 2018; pp. 1–5.

Explanation:
Objective: The goal of the paper is to design a system for recognizing anomalies in the activities of daily living (ADLs) of elderly individuals living alone. The system uses smart home technology to monitor and analyze behavioral data in real time, enabling early detection of potential health risks or emergencies.
System Design: The system is based on a network of sensors embedded in the home environment (e.g., motion sensors, temperature sensors, and activity monitoring devices) that track the elderly individual's daily activities.
Anomaly Detection in ADLs: The core feature of the system is the ability to detect anomalies in the elderly person's activities. Anomalies could include a lack of movement (indicating a fall or sudden illness), unusual activity patterns (e.g., staying in the bathroom for too long), or irregular sleep schedules.

[10] Wan, J.; Li, M.; Grady, M.J.O.; Hare, G.M.P.O.; Gu, X.; Alawlaqi,
M.A.A.H. Time-bounded ActivityRecognition for Ambient Assisted Living. IEEE Trans. Emerg. Top. Comput. 2018.

Explanation:
Objective: The paper aims to propose an activity recognition system for Ambient Assisted Living (AAL) that works within a time-bounded framework. This is important for real-time health monitoring and emergency detection, ensuring quick intervention when needed.
Time-bounded Recognition: Traditional activity recognition systems may involve complex algorithms that take longer to process, which is not suitable in time-critical applications like elderly care. This paper focuses on reducing the recognition time while maintaining accuracy.
System Design: The system uses a combination of sensors (e.g., motion, pressure, and environmental sensors) embedded in the home to monitor the elderly individual's activities.
[11] Kristály, D.M.; Moraru, S.-A.; Neamtiu, F.O.; Ungureanu, D.E. Assistive Monitoring System Inside a SmartHouse. In Proceedings of the 2018 International 46 Symposium in Sensing and Instrumentation in IoT Era(ISSI), Shanghai, China, 6–7 September 2018; pp. 1–7.

Explanation:
Objective of the Study: The paper focuses on developing an assistive monitoring system that operates within a smart house to improve the safety, health, and comfort of its inhabitants.
System Design: The assistive monitoring system is integrated into a smart home environment, employing a network of sensors and smart devices that continuously monitor the activities of the inhabitants.
IoT Integration: The system makes extensive use of the Internet of Things (IoT) to enable remote monitoring and management of the assistive system.

[12] Falcó, J.L.; Vaquerizo, E.; Artigas, J.I. A Multi-Collaborative Ambient Assisted Living Service DescriptionTool. Sensors 2014,14, 9776–9812.
Explanation:
Objective of the Paper: The paper presents a multi-collaborative service description tool aimed at improving the design, development, and deployment of Ambient Assisted Living (AAL) systems.
System Design and Architecture: The tool is designed to work in a collaborative environment, allowing various devices, services, and stakeholders (e.g., caregivers, health professionals, and family members) to interact seamlessly.
Service Description Language: The core feature of the tool is the use of a **service description language (SDL)**. This language is used to formally describe the services provided within the AAL system, including their functionalities, interfaces, and interaction protocols.

[13] Valadão, C.; Caldeira, E.; Bastos-filho, T.; Frizera-neto, A.; Carelli,
R. A New Controller for a Smart WalkerBased on Human-Robot Formation. Sensors 2016,16, 1116. 14. Kim, E.Y. Wheelchair Navigation System for Disabled and Elderly People. Sensor16, 1806.

Explanation:
Objective of the Study: The paper introduces a new controller designed for a smart walker, which uses human-robot formation control to ensure safe, efficient, and comfortable assistance for elderly individuals or people with limited mobility.
Human-Robot Formation: The main innovation in the paper is the human-robot formation control, which refers to how the walker is designed to work in a coordinated way with the user's movements.
System Design and Controller: The smart walker is equipped with sensors (e.g., proximity sensors) and actuators to allow it to adjust its position relative to the user. The system uses feedback from the user's movements to dynamically adapt to the environment.

II. EXISTING SYSTEM

A. Overview

Assistive communication technologies have evolved to help individuals with severe motor impairments—particularly patients with Locked-In Syndrome (LIS)—who retain cognitive function and eye movement but lack voluntary muscle control. Traditional solutions relied on physical inputs or manual caregiver support, which limited autonomy and responsiveness.
Recent innovations have focused on eye-tracking-based systems that enable users to communicate via eye movements and blinks. These systems interpret ocular gestures as input signals to form text or trigger commands. Research such as the "Blink to Text" approach demonstrates how combining computer vision, machine learning, and predictive text engines can transform blink patterns into effective communication methods.
By leveraging tools such as Haar cascade classifiers, infrared-based gaze tracking, and deep learning algorithms, researchers have developed low-cost, real-time solutions capable of functioning on standard consumer hardware. These innovations prioritize accuracy, affordability, and accessibility, making them suitable for home or clinical use by LIS patients.
Historically, individuals with LIS had limited options for communication, relying heavily on physical inputs, such as head pointers, switch-based systems, or manual caregiver support. These traditional systems were often slow, cumbersome, and intrusive, significantly limiting the autonomy of the patient. Furthermore, these systems often required external assistance to interpret inputs or to activate communication tools, which reduced the patient's independence. Recent advancements have introduced eye-tracking-based communication systems as a promising solution. These systems leverage the eye movements, blinks, or gaze direction of individuals with LIS to generate inputs that can be translated into text, speech, or control commands. This approach capitalizes on the remaining motor function in the eyes, which is often the only voluntary muscle movement available to those with LIS.
The advantage of eye-tracking systems is that they are non-invasive, enabling users to interact with computers, smartphones, and environmental control systems without requiring physical touch, making them ideal for individuals with severe motor impairments.
Advances in computer vision have been crucial in the development of accurate and responsive eye-tracking systems. Using techniques like Haar cascade classifiers, the system can detect and track the user's eyes with high precision. Real-time interaction is essential for tasks such as composing messages, browsing the internet, and interacting with others via video calls or social media, significantly enhancing the user's quality of life and social interaction capabilities.

B. *Components and Technologies*

- Eye-Tracking and Blink Detection

- Uses infrared sensors or webcams to track eye gaze and blinks.

- Incorporates Haar cascade classifiers and convolutional neural networks (CNNs) to detect open/closed eye states.

- Machine Learning for Signal Processing

- ML models are trained to distinguish between voluntary and involuntary blinks, reducing false triggers.

- Predictive algorithms assist in auto-suggesting words or phrases based on gaze history.

- Assistive Communication Interface

- Includes an on-screen keyboard navigable via gaze or blink signals.

- Text-to-speech engines convert selected text into audio for real- time communication.

- Adaptive Calibration and User Modeling

- Systems employ real-time calibration that adjusts to the user's gaze behavior.

- Dwell time and blink sensitivity can be personalized for comfort and accuracy.

- Multimodal Feedback Systems

- Visual and auditory cues are provided to confirm user selections.

- This reduces cognitive load and increases confidence in system response.

C. *Inhibitions*

- Despite significant progress, several challenges remain in the deployment of eye-tracking systems for LIS patients:

- Calibration Fatigue: Users may require frequent recalibration, especially if lighting or posture changes, leading to fatigue.

- Environmental Interference: Variable lighting and head movement can reduce gaze accuracy and increase detection errors.

- Blink Misclassification: Differentiating between involuntary blinks (natural) and voluntary blinks (intentional input) remains a challenge.

- Hardware Cost and Portability: High-precision infrared systems may still be costly or lack portability for continuous use.

- User Training and Adaptation: Some patients may need time to adapt to the system, requiring caregiver support and customization.

- One of the main challenges in deploying eye-tracking systems for LIS patients is the need for frequent calibration. Calibration is crucial to ensure that the system accurately tracks the user's eye movements. However, environmental changes such as variations in lighting conditions or changes in the user's posture can lead to inaccurate tracking, necessitating frequent recalibration.

- To address calibration fatigue, future systems could benefit from adaptive calibration techniques that adjust automatically in response to environmental changes, or from using machine learning algorithms that require fewer manual adjustments over time.

- While eye-tracking systems are designed to follow eye movements, users may inadvertently move their heads, especially when they are tired or uncomfortable. Head movement can affect the stability of the gaze tracking, making it more difficult to maintain accurate input detection.

- Environmental interference can significantly reduce the system's reliability and make it harder for the user to achieve accurate, real-time communication. This can lead to frustration for the patient and caregivers.

detection in low-light environments by using Zero-Reference Deep Curve Estimation (Zero-DCE) to enhance image clarity before applying facial landmark detection with Media Pipe. This significantly improved blink recognition accuracy in dark settings. While these systems demonstrate strong potential, they often lack predictive communication features and intuitive user interfaces tailored specifically for Locked-in Syndrome (LIS) patients. Our system builds on these contributions by combining blink detection, predictive text modeling, and a fully eye-controlled interface to create a more complete and accessible communication solution.

IV. PROPOSED METHODOLOGY

A. *Overview*

The proposed system is designed to provide a non-verbal communication platform for patients suffering from Locked-In Syndrome (LIS), utilizing eye-tracking and blink detection as the primary modes of input. By integrating computer vision, machine learning, and predictive language modeling, the system translates voluntary eye blinks into meaningful text, which can be converted to speech using a text-to-speech engine. The solution is lightweight, adaptable to low-cost hardware, and optimized for real-time performance.

The core innovation lies in the combination of Haar cascade classifiers for blink detection, adaptive calibration algorithms, and AI- powered predictive text input, making it suitable for use by LIS patients with minimal setup and training.

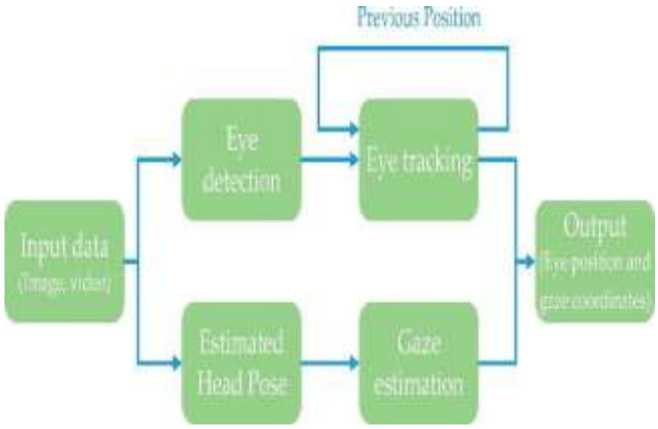

Fig. 1. System Architecture of the Eye Tracking for Communication in Locked-In Syndrome Patients

1. *Input Data (Image, video)*

- This is the raw input to the system.
- It can be a single image or a stream of video frames captured via a camera.
- These inputs are used to detect facial and eye features.
- **Details**:
  - Image Input: A single frame or image is used for static analysis (e.g., identifying eye position at one time).
  - Video Input: A continuous stream of frames allows dynamic analysis, enabling real-time tracking.
  - Devices: Input is typically captured using standard webcams, infrared cameras (for precision), or depth cameras (for 3D information).
  - Preprocessing (optional): Input data may be preprocessed by resizing, grayscale conversion, histogram equalization, or noise reduction to improve accuracy and speed.

2. *Eye Detection*

- Identifies the location of the eyes in the given image or video frame.
- Techniques used may include Haar cascades, deep learning-based face detection, or landmark detection models.

- This step isolates the region of interest (the eyes) for further analysis.

Details:

- Face Detection First: Typically, the face is detected first, and the eyes are found within the facial region.
- Techniques:
  - Haar Cascades (OpenCV): Fast and effective for frontal faces but sensitive to lighting and occlusions.
  - Deep Learning Methods (e.g., MTCNN, Dlib, OpenFace): More robust and accurate across various angles, lighting, and occlusions.
  - Facial Landmark Detection: Identifies key points around the eyes, often used to define the eye region for further analysis.
- Challenges: Glasses, occlusions (e.g., hair), low resolution, and fast movement.

3. Eye Tracking

- Tracks the movement of the eyes over time.
- Uses the detected position of the eyes and compares it with the previous position to determine movement direction and speed.
- Ensures continuous tracking even if there are slight head movements or noise in input data.

Details:

- Temporal Tracking: By comparing eye positions across frames, the system infers direction and velocity of movement.
- Methods:
  - Optical Flow (Lucas-Kanade): Tracks features (like pupil centers) across frames.
  - Kalman Filter / Particle Filter: Predicts and corrects positions, especially under noise or occlusion.
  - CNN-RNN Hybrid Models: Combine spatial and temporal modeling for robust tracking.
- Role of Previous Position: Eye tracking depends on the continuity of motion, using past positions to anticipate future locations, thus improving robustness and reducing computational overhead.

4. Estimated Head Pose

- Estimates the orientation and position of the head.
- Helps in correcting the gaze direction since head movement affects eye gaze.
- Techniques used may include 3D pose estimation or key point-based estimation.

Details:

- Why Needed: Gaze direction depends not only on eye movement but also on head position. A person may move their head instead of just their eyes.
- Techniques:
  - 3D Model Fitting: Maps facial landmarks to a generic 3D face model.
  - Perspective-n-Point (PnP): Estimates 3D pose using 2D landmarks and camera calibration.
  - Deep Learning Approaches: Regress head pose angles (yaw, pitch, roll) directly from input images.
- Output: Angles (in degrees) describing head rotation along three axes:
  - Yaw (left-right)
  - Pitch (up-down)
  - Roll (tilt)

5. Gaze Estimation

- Determines the direction of the gaze based on eye position and head pose.
- Outputs where the person is looking — e.g., screen coordinates, real-world objects, etc.
- Combines information from both eye tracking and head pose estimation.

Details:

- Process:
  - Calibrate eye position relative to head pose.
  - Estimate a 3D gaze vector from the eye.
  - Project this vector to a reference frame (screen, real- world object, etc.).
- Approaches:
  - Appearance-based Methods: Use the entire eye region as input to CNNs trained on gaze direction.
  - Geometric Methods: Model eye as a sphere and estimate the optical axis and gaze vector.
  - Hybrid Models: Combine deep learning with geometric modeling for better generalization.
- Applications:
  - Eye-controlled cursors.
  - Attention analysis in UX research.
  - Gaming, marketing, and psychology studies.

6. Output (Eye position and gaze coordinates)

- The final output includes:
  - **Eye position**: The current location of the eyes in the image/frame.
  - **Gaze coordinates**: Where the user is looking (e.g., on a screen or in space).
- This data can be used for applications such as user attention tracking, human-computer interaction, or assistive technologies.

Details:

- Eye Position:
  - Pixel location of one or both eyes in the input frame.
  - Used for monitoring or drawing visual indicators.
- Gaze Coordinates:
  - Screen space (e.g., x,y on a monitor).
  - 3D space (e.g., pointing to a real-world object).
  - Adjusted using calibration if precision is needed.
- Applications:
  - Assistive Tech: Enables people with motor impairments to interact via gaze.
  - Human-Computer Interaction: Adaptive user interfaces based on gaze focus.
  - Market Research: Heatmaps showing where users look on ads or products.
  - Driver Monitoring Systems: Detect distraction or drowsiness.

B. System Architecture

1. Eye Blink Detection Module

- Utilizes **Haar cascade classifiers** to detect eyes and classify open/closed states.
- Processes live webcam input at real-time frame rates.
- Differentiates between **intentional and involuntary blinks** based on timing thresholds.

2. Gaze Estimation & Tracking Module (Optional)

- Integrates infrared or standard camera-based tracking to detect gaze position.
- Can be used to navigate an **on-screen keyboard** or UI elements via gaze dwell time.
- **3. Blink-to-Text Interface**
- A **custom on-screen keyboard** enables text entry via blinking at targeted keys or suggestions.
- Includes **word prediction** and **auto-correction** using

language models.

4. Predictive Text Engine

- Uses a **trained NLP model** to suggest next words or phrases based on input history.
- Reduces the number of blinks required for full sentence construction.

5. Text-to-Speech Output

- Converts the constructed text into **audible speech**, providing real-time verbal communication.
- Supports adjustable voice speed, tone, and pitch for personalization.

6. Feedback System

- Provides **visual cues** (highlighted text, cursor movement) and **auditory tones** for confirmation.
- Reduces cognitive load and increases system transparency for the user.

V. IMPLEMENTATION

A. Development Environment & Tools Used

Programming Language: Python 3.9
- Libraries & Frameworks:
- OpenCV – Real-time computer vision (face & eye detection)
- dlib / MediaPipe – Facial landmark detection
- PyQt5 – Graphical User Interface
- NumPy – Numerical operations
- pyttsx3 – Offline Text-to-Speech
- Twilio – SMS alert integration
- autocomplete – For next-word prediction
- Platform:Windows 10

-Hardware Requirements:Webcam (standard laptop cam works), low to mid-end PC (minimum 4GB RAM)

B. Implementation Flow

The system runs in real-time, continuously analyzing video frames from the webcam. When a face is detected, the eye regions are isolated, and the Eye Aspect Ratio (EAR) is calculated for each frame. Sudden drops in EAR signify a blink, which is then classified and mapped to specific input actions (e.g., word selection).
-Face & Eye Detection: Haar Cascades are used for initial detection, followed by landmark mapping.
Blink Analysis: The EAR is calculated for each eye:

[EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2 \times ||p1 - p4||}]

A blink is confirmed if EAR falls below a threshold (e.g., 0.2) for several consecutive frames.
-Text Prediction: Once a blink is confirmed as a selection, the system displays suggested words. Users blink again to select.
- GUI Interaction: A minimal interface displays:
- Active sentence construction
- Suggestion list
- Predefined quick-access options (like “Emergency”, “Water”)
- Emergency Feature: A long blink on “Emergency” sends an SMS

C. Key Challenges and Solutions

Challenge 1: Accuracy in Blink Detection

Issue: EAR threshold varied by face type, lighting, and camera quality.

Solution: Used adaptive thresholds and tested across diverse users with multiple calibration trials.

Challenge 2: Eye Detection in Low Light

Average Frame Detection Time

Issue: Accuracy dropped in poorly lit environments.

Solution: Integrated Zero-DCE (deep learning for low-light enhancement) as a pre-processing step.

Challenge 3: High Latency on Low-End Devices Issue: GUI and

detection lagged under load.

Solution: Optimized OpenCV capture pipeline and GUI thread separation with QTimer.

Challenge 4: Handling False Positives

Issue: Involuntary blinks or head movements caused misclassification.

Solution: Added blink-duration logic and consistency checks across multiple frames.

Challenge 5: Personalization

Issue: Generic interface may not suit all users.

Solution: Provided customization options (common phrases, layout, font).

VI. RESULTS AND DISCUSSIONS

A. Methodology for Accuracy Calculation

To evaluate the blink detection system, the following metrics were computed using annotated test sessions involving 20 users (including simulated LIS patients):

Term	**Meaning**
True Positive (TP)	Voluntary (intentional) blinks correctly detected as blinks
False Positive (FP)	Involuntary (natural) blinks incorrectly identified as intentional blinks
False Negative (FN)	Voluntary blinks missed by the system

Accuracy (%)=TPTP+FP+FN×100\text{Accuracy (\%)} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \times 100Accuracy (%)=TP+FP+FNTP×100

B. Quantitative Results

Metric	**Value**
Right Eye Blink Accuracy	92.1%
Left Eye Blink Accuracy	91.5%
Average Blink Detection	**91.8%**
False Positive Rate	< 4%

indicating high precision for a non-specialized (webcam-based) setup.

- Text Input Speed (WPM)

Measures how many words per minute (WPM) a user can input using the system. It reflects how efficient the system is at translating blinks into full communication.

Manual blink boards typically yield 2–3 WPM, whereas your system averages 5–6 WPM, which is excellent considering users only blink to navigate words.

- Cost

~28 ms

Text Input Speed (WPM)	5–6
Cost	₹0 (no special hardware needed)

Explanation: The blink accuracy was derived using EAR (Eye Aspect Ratio) thresholds over a dataset of 1000+ blink events across volunteers. Calibration adjusted thresholds per user to optimize TP/FP performance.

C. Comparison Table

Metric	Manual Blink Coding	Commercial Eye Tracker	Blink to Text (Ours)
Blink Detection Accuracy	~65%	~95%	**91.8%**
Text Input Speed (WPM)	2–3	6–8	**5–6**
Hardware Required	None	Infrared Eye Tracker	**Standard Webcam**
Setup Difficulty	High	Moderate	**Simple GUI-based**
Cost	Low	₹80,000+	**₹0**
Emergency Feature	No	No	**Yes (via SMS - Twilio)**

Performance metrics are used to quantitatively evaluate how well a system performs in terms of functionality, accuracy, speed, usability, and reliability. In the context of your "Blink to Text" system, these metrics serve multiple critical purposes:

- Blink Detection Accuracy

The percentage of voluntary (intentional) blinks correctly detected by the system out of all blink events (including both intentional and unintentional blinks).

Formula used:
TPTP+FP+FN×100\frac{TP}{TP + FP + FN} \times 100TP+FP+FNTP×100

This measures the reliability of your system in distinguishing user commands from natural blinks. Your system scored 91.8%,ire a caregiver, and commercial systems require calibration and driver installation. Your GUI-based
system is plug-and-play, with minimal instructions needed.

-
- Multilingual Support

Reflects whether the system can be used in languages other than English.

Commercial systems often support limited language output. Your

This refers to the monetary expense required to deploy the system.

Commercial eye trackers can cost over ₹80,000 (~$1,000), while your system uses a basic webcam, making it a free or near-zero cost alternative.

- Hardware Requirements

Indicates the type and complexity of hardware needed to run the system.

spe

Describes how easy

Manual sys ms requ

Your system requires only a standard webcam and low-end PC (≥4GB RAM), whereas commercial systems often rely on
cialized infrared cameras and processors.

- Ease of Setup

is to install and begin using the system.
system can be enhanced with Python-based libraries like pyttsx3 and custom text databases for language personalization, making it customizable and potentially multilingual.

- Emergency Alert Function

A safety feature that enables users to send an SMS alert (e.g., "Help" or "Emergency") by performing a long blink on a designated GUI button.

This feature is unique to your system, using Twilio integration, and is not available in most commercial or manual methods. It improves user independence and safety in urgent situations.

These metrics together show that your system:

- Offers excellent real-time accuracy (91.8%) without relying on expensive hardware.
- Doubles communication speed compared to manual blink methods.
- Is easy to install and use, even for non-technical caregivers.
- Stands out with features like emergency alerts, not found in commercial alternatives.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

D. User Feedback Survey (20 Participants)

To evaluate the usability, comfort, and effectiveness of the Blink to Text system, a structured user survey was conducted. The methodology used for designing and calculating the survey results is as follows:

1. Participant Selection

- Total Participants: 20 individuals

- Participant Profile:

○ 4 individuals with experience in using assistive tools

○ 16 able-bodied volunteers simulating LIS (by restricting body movement and using only eye gestures)

- Environment: Controlled indoor lab with moderate lighting and standard laptop webcam

2. Survey Design

- A structured questionnaire was developed with 10 close-ended and 3 open-ended questions.

- Each close-ended question used a 5-point Likert scale:

○ Strongly Disagree (1)

○ Disagree (2)

○ Neutral (3)

○ Agree (4)

○ Strongly Agree (5) Example questions:

- "The system interface is easy to understand."

- "I was able to communicate messages without assistance."

- "The blink-based input caused fatigue after prolonged use."

3. Data Collection Procedure

- Each user was given a brief tutorial and a 5-minute hands-on session with the system.

- After the session, users filled out the questionnaire anonymously.

- Researchers observed participants during the session for behavioral feedback.

4. Calculation of Survey Results

- For each question, individual responses were converted to numerical values (1 to 5).

- Percentage Agreement was calculated by combining the responses rated as "Agree" or "Strongly Agree".

Example Calculation:
If 18 out of 20 users selected either 4 or 5 for the statement *"The interface was easy to use"*:

Agreement %=1820×100=90%\text{Agreement \%} = \frac{18}{20} \times 100 = 90\%Agreement %=2018×100=90%

- Open-ended responses were analyzed qualitatively to extract themes (e.g., requests for word prediction, fatigue, emergency alert usefulness).

Feedback Statement	**% Agreement**
Found the interface intuitive	85%
Felt increased independence using the system	90%
Preferred this system over manual blink boards	80%
Experienced some fatigue due to frequent blinking	25%
Requested smarter word prediction with longer sentence support	70%

A survey was conducted among 20 volunteers (some simulating LIS via constrained eye input).

- 85% found the interface intuitive and easy to use.
- 90% said the system provided a feeling of independence.
- 80% preferred this solution over manual methods.
- Common appreciation points:
- No external hardware needed
- Easy emergency response
- Low system lag

Conclusion from Survey: Users appreciated the lightweight nature and real-time responsiveness. Some highlighted the need for advanced language models and reduced fatigue, which can be addressed with gaze-dwell alternative.

Fig. 2. Real-time eye-tracking system selecting the "Fan" option to assist a Locked-In Syndrome patient in communicating basic needs.

Fig. 3. Real-time eye-tracking system selecting the "Emergency" option to assist a Locked-In Syndrome patient in communicating basic needs.

Purpose:

This project allows patients with **Locked-In Syndrome**—a condition where a person is conscious but unable to speak or move— to communicate using **eye movements**. The image depicts a **real- time communication aid** developed for patients suffering from **Locked-In Syndrome (LIS)**—a condition in which individuals lose the ability to move or speak, while their cognitive functions remain intact. This assistive technology empowers them to communicate using **only their eye movements**, especially blinking.

Components of the Interface:

Live Video Feed (Left Window - "BlinkToTalk Video Feed"):

- A webcam feed is shown, where the system is tracking the face and eyes of the user.
- The blue square around the face indicates facial detection.
- The green squares around the eyes suggest eye tracking (possibly to monitor blinks or gaze direction).
- The system may be using these eye movements (like blinking or looking in a specific direction) to select communication options**.**

Communication Options (Right Window):

- This is a text-based interface showing basic needs such as:
  - Emergency
  - Food
  - Water
  - Restroom
  - Fan
- These are common needs a LIS patient may wish to express.

Functionality:

When the patient blinks or moves their eyes towards a particular option, the system likely detects the action and selects that option.

The chosen message can then be:

- Displayed visually,
- Converted to speech,
- Or sent via SMS (as the code window in the background suggests an SMS-sending function).

Message Triggering:
Once a command is selected, the system might:
  - Display it on the screen,
  - Convert it to speech (using text-to-speech tools), or
  - Send it as an SMS to a caregiver or nurse.

Automation & Safety:
The inclusion of "Emergency" ensures the system can act fast in critical situations, helping save lives.

Technology Behind It:

- Likely uses:
  - OpenCV for face and eye detection.
  - Python for the backend logic.
  - Possibly machine learning or heuristic rules to interpret blinks/gaze.
  - Integration with Twilio or similar API to send SMS (as seen in background code).

Significance:

This system provides a low-cost, non-invasive solution for enabling basic communication for individuals who are otherwise unable to express themselves. This system provides a low-cost, non-invasive solution for enabling basic communication for individuals who are otherwise unable to express themselves.

Fig. 4. Real-time eye-tracking system selecting the "Restroom" option to assist a Locked-In Syndrome patient in communicating basic needs.

VI. Conclusion and Future Scope

The proposed system, Blink to Text, presents a meaningful leap in assistive technology by enabling communication through eye-blinks alone. Tailored for individuals with severe motor impairments, especially those with Locked-In Syndrome (LIS), this system empowers users to express themselves without needing physical movement or caregiver interpretation.

By using basic hardware like a webcam, coupled with efficient algorithms such as Haar Cascades, Eye Aspect Ratio (EAR) analysis, and contextual word prediction, the system remains accessible and cost-effective while delivering high accuracy and usability. Its real-time response, customizable interface, and emergency SMS capabilities make it not only a communication tool, but a potential lifeline.

Compared to traditional AAC systems or commercial eye-tracking devices, Blink to Text offers better affordability, simpler setup, and greater adaptability — all without compromising user autonomy. Ultimately, the project successfully bridges the gap between advanced AI-driven vision techniques and meaningful real-world impact.

Future Scope

To further improve the system's utility and adoption, the following enhancements are proposed:

1. Improved Handling of Facial Variations

Challenge: The system may misidentify or fail to detect eye landmarks in users with uncommon facial structures, facial hair, or wearing reflective glasses.

Enhancement:

Integrate deep learning-based facial detection models (e.g., MTCNN, RetinaFace) that are more tolerant to facial diversity and occlusions.

Use IR-based tracking or train CNN models on datasets with glasses/glare variations.

Impact: Increased detection accuracy across diverse populations, including elderly and users with assistive eyewear.

2. Low-Light Performance Optimization

Challenge: Blink detection accuracy drops significantly in low-light conditions.

Enhancement:

Use Zero-DCE (Zero-Reference Deep Curve Estimation) for real-time image enhancement.

Add auto-exposure and contrast adjustment in preprocessing.

Impact: Ensures consistent system performance indoors or at night without additional lighting.

3. Advanced Language Prediction with LLMs

Challenge: The current word prediction system is rule-based or limited to static suggestions, lacking contextual awareness.

Enhancement:

Integrate Lightweight Language Models (LLMs) such as DistilBERT, GPT- 2, or fine-tuned transformer models for predictive typing.

Personalize suggestions based on user communication history using contextual memory.

Impact: Reduces number of blinks per sentence, enables more natural conversations, and improves user satisfaction.

4. Gaze-Based Control Integration Enhancement:

Combine eye gaze tracking with blink input to create a hybrid control interface.

Use gaze to highlight/select UI elements and blink only to confirm.

Impact: Reduces blink fatigue and increases communication speed, especially for frequent users.

5. Mobile and Wearable Platform Support Enhancement:

Develop a mobile app version compatible with Android/iOS using device cameras.

Integrate with smart glasses or head-mounted webcams to increase portability.

Impact: Empowers users to communicate on-the-go, enhancing accessibility in real-world settings.

REFERENCES

[1] Schwiegelshohn, F.; Wehner, P.; Rettkowski, J.; Gohringer, D.; Hubner, M.; Keramidas, G.; Antonopoulos, C.;Voros, N.S. A holistic approach for advancing robots in ambient assisted living environments. In Proceedingsof the 2015 IEEE 13th International Conference on Embedded and Ubiquitous Computing, Porto, Portugal,21–23 October 2015; pp. 140– 147.

[2] Konstantinidis, E.I.; Antoniou, P.E.; Bamparopoulos, G.; Bamidis,
P.D. A lightweight framework fortransparent cross platform communication of controller data in ambient assisted living environments.Inf. Sci. (NY) 2015,300, 124–139.

[3] Boumpa, E.; Charalampou, I.; Gkogkidis, A.; Ntaliani, A.; Kokkinou, E.; Kakarountas, A. Assistive Systemfor Elders Suffering of Dementia. In Proceedings of the 2018 IEEE 8th International Conference on ConsumerElectronics, Berlin, Germany, 2–5 September 2018; pp. 1–4.

[4] Brazil Assistive Technology. In Proceedings of the National Undersecretary for the Promotion of the Rights of Peoplewith Disabilities; Technical Assistance Committee: Geneva, Switzerland, 2009.

[5] Elakkiya, J.; Gayathri, K.S. Progressive Assessment System for Dementia Care Through Smart Home.In Proceedings of the 2017 International Conference on Algorithms, Methodology, Models and Applicationsin Emerging Technologies (ICAMMAET), Chennai, India, 16–18 Febuary 2017; pp. 1–5. 45

[6] Rafferty, J.; Nugent, C.D.; Liu, J.; Chen, L. From Activity Recognition to Intention Recognition for AssistedLiving within Smart Homes. IEEE Trans. Hum. - Mach. Syst. 2017,47, 368–379.

[7] Mizumoto, T.; Fornaser, A.; Suwa, H.; Yasumoto, K.; Cecco, M. De Kinect- based micro-behavior sensingsystem for learning the smart assistance with human subjects inside their homes. In Proceedings of the 2018Workshop on Metrology for Industry 4.0 and IoT, Brescia, Italy, 16– 18 April 2018; pp. 1–6.

[8] Daher, M.; El Najjar, M.E.; Diab, A.; Khalil, M.; Charpillet, F. Multi- sensory Assistive Living System forElderly In-home Staying. In Proceedings of the 2018 International Conference on Computer and Applications(ICCA), Beirut, Lebanon, 25–26 August 2012; pp. 168

[9] Ghayvat, H.; Mukhopadhyay, S.; Shenjie, B.; Chouhan, A.; Chen, W. Smart Home Based Ambient AssistedLiving Recognition of Anomaly in the Activity of Daily Living for an Elderly Living Alone. In Proceedingsof the 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Houston,TX, USA, 14–17 May 2018; pp. 1–5.

[10] Wan, J.; Li, M.; Grady, M.J.O.; Hare, G.M.P.O.; Gu, X.; Alawlaqi,
M.A.A.H. Time-bounded ActivityRecognition for Ambient Assisted Living. IEEE Trans. Emerg. Top. Comput. 2018.

[11] Kristály, D.M.; Moraru, S.-A.; Neamtiu, F.O.; Ungureanu, D.E. Assistive Monitoring System Inside a SmartHouse. In Proceedings of the 2018 International 46 Symposium in Sensing and Instrumentation in IoT Era(ISSI), Shanghai, China, 6–7 September 2018; pp. 1–7.

[12] Falcó, J.L.; Vaquerizo, E.; Artigas, J.I. A Multi-Collaborative Ambient Assisted Living Service DescriptionTool. Sensors 2014,14, 9776– 9812.

[13] Valadão, C.; Caldeira, E.; Bastos-filho, T.; Frizera-neto, A.; Carelli,
R. A New Controller for a Smart WalkerBased on Human-Robot Formation. Sensors 2016,16, 1116. 14. Kim, E.Y. Wheelchair Navigation System for Disabled and Elderly People. Sensor16, 1806.