AROGYA SAMVAD: EMPOWERING COMMUNICATION FOR INDIVIDUALS WITH MOTOR IMPAIRMENT

Project Reference No.: *48S_BE_2676*

College : Sahyadri College Of Engineering & Management, Mangaluru
Branch : Department Of Computer Science And Engineering (Artificial Intelligence & Data Science)
Guide(s) : Dr. Pushpalatha K
Student(s) : Ms. Advika Nilkanth Yengde
Ms. K Sakshi R Kini
Ms. Prithvi U Pai
Ms. Varsha Rao

Keywords:

Assistive Technology, Motor Impairment, Artificial Intelligence, Human-Computer Interaction, Healthcare Technology, Alternative and Augmentative Communication

Introduction:

Communication is a fundamental human need, yet individuals with severe motor impairments, such as those with locked-in syndrome or paralysis, often struggle to express their basic needs. Existing assistive technologies, including Brain-Computer Interfaces (BCI) and eye-tracking systems, can be expensive, invasive, or require extensive training, limiting their accessibility. Advances in AI and deep learning have paved the way for more intuitive, real-time, and non-invasive solutions.

Arogya Samvad builds upon these advancements by integrating AI-driven eye-tracking and hand gesture recognition to create a seamless communication system. Unlike traditional devices that may be bulky or complex, this system is adaptive, affordable, and user-friendly. Using computer vision and machine learning, it accurately interprets eye movements and gestures in real time, providing a reliable means of interaction for individuals with severe motor impairments.

Prioritizing accessibility and affordability, Arogya Samvad bridges the gap between patients and caregivers, improving independence and enhancing caregiver efficiency. Through continuous research and iterative development, the system aims to refine its accuracy and expand its capabilities, ultimately contributing to the evolution of AI-driven assistive communication technologies.

Objectives:

1. To extract and preprocess video frames to optimize the accuracy and efficiency of eyeball movement and hand gesture tracking for real-time communication.
2. To develop a robust and precise method for pupil detection, enabling reliable eye-tracking as a communication tool for individuals with motor impairments.
3. To design an accurate and responsive hand gesture recognition system, allowing users with partial mobility to interact effectively and express their needs.
4. To integrate eye-tracking and hand gesture recognition into an AI-powered communication system that enhances independence and improves the quality of life for individuals with severe motor impairments.

Methodology:

Arogya Samvad enables individuals with severe motor impairments to communicate through real-time eye-tracking and hand gesture recognition, bridging the gap between users and caregivers. Using dlib and Mediapipe, the system classifies gestures and eye movements, mapping them to predefined commands like calling a caregiver or indicating "yes" or "no." To ensure accuracy, thresholds are fine-tuned, and models are validated under real-world conditions. Recognized inputs are converted into notifications, SMS alerts, or voice messages, allowing caregivers to respond promptly.

The figure 1 outlines the architecture flow of the solution.

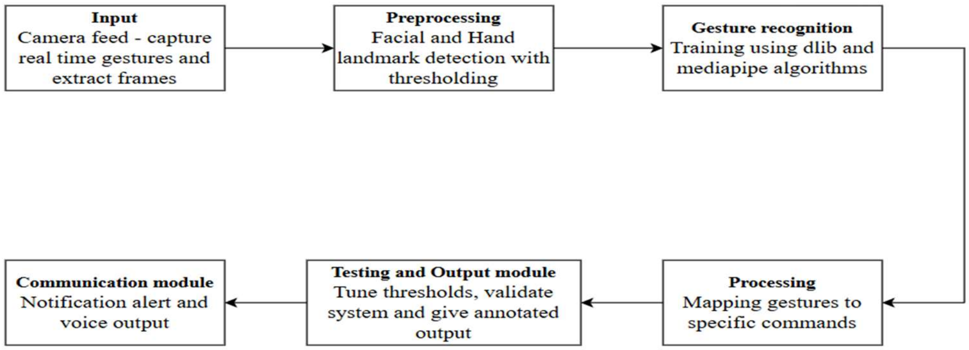

Figure 1: Architecture Diagram of Arogya Samvad

Hand Gesture Recognition:

The hand gesture recognition module is implemented using Google's MediaPipe framework. The process follows these steps:

1. Capturing video frames and preprocessing them by converting images from BGR to RGB for compatibility.
2. Detecting hand landmarks, specifically 21 key points such as fingertips, joints, and palm positions, using MediaPipe's Hand Tracking Module.
3. Classifying gestures by comparing detected landmarks against predefined templates. Common gestures like "Food," "Water," or "Call Caregiver" were assigned specific outputs.
4. Generating output, where recognized gestures are displayed on the screen as text and converted into speech using a Text-to-Speech (TTS) module.
5. Notifications are also sent via Telegram messaging to alert caregivers.

Eye Gaze Tracking:

Eye-tracking is implemented using Dlib's pre-trained facial landmark detector, focusing on landmarks 36-41 (left eye) and 42-47 (right eye) for gaze estimation. The process includes:

1. Detecting the user’s face and eyes in video frames using Dlib’s shape predictor model.
2. Extracting eye landmarks and computing relative positions to determine gaze direction.
3. Classifying gaze directions into categories such as "Looking Left," "Looking Right," and "Looking Forward."
4. Displaying detected gaze directions as onscreen text and converting them into speech for auditory assistance.

Results and Conclusions:

Arogya Samvad successfully enabled real-time communication for individuals with motor impairments by integrating eye-tracking and hand gesture recognition. The system accurately detected gaze directions and hand movements, converting them into commands with minimal delay. Instant notifications to caregivers ensured quick responses, making communication more seamless and accessible. Testing across different lighting conditions and user variations confirmed the system’s reliability and adaptability.

During development, challenges like glare affecting eye tracking and variations in hand positions were addressed through fine-tuned detection models and preprocessing techniques. Gesture recognition performed well across different backgrounds, and real-time processing ensured smooth interaction without lag. The multimodal approach proved effective, allowing users to switch between eye-tracking and hand gestures based on their comfort and ability. The results of eye gaze tracking and hand gesture recognition are shown in figure 2 and figure 3.

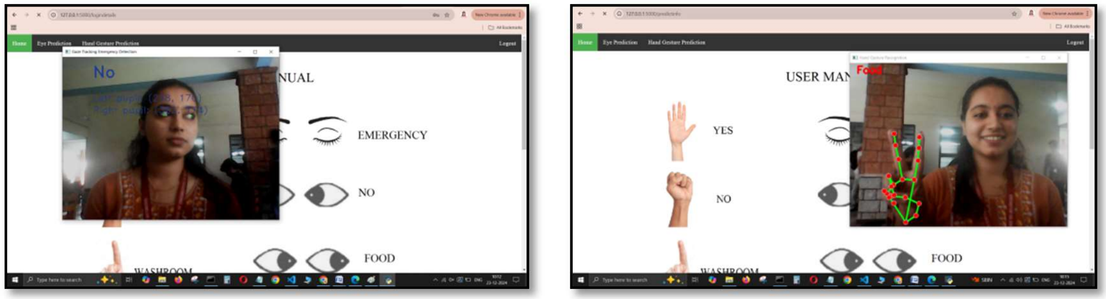

Figure 2&3: Outputs of eye gaze tracking and hand gesture recognition respectively

Overall, Arogya Samvad demonstrated the potential of AI-driven assistive technology in improving accessibility for people with motor impairments. By making communication intuitive and efficient, it empowers users with greater independence while easing the burden on caregivers. The results highlight how AI and real-time processing can bridge communication gaps, enhancing quality of life for those in need.

Project Outcome & Industry Relevance:

Arogya Samvad emerged as a transformative solution that restores a voice to individuals who otherwise struggle to communicate due to severe motor impairments. By leveraging AI for real-time gesture and eye gaze recognition, the system bridges the gap between intention and expression. It empowers users to convey their needs with minimal effort and provides caregivers with timely, clear alerts—reducing guesswork and emotional strain. The model successfully delivered a functional, user-friendly prototype that demonstrated how accessible technology can dramatically improve quality of life, independence, and dignity for both patients and their support systems.

The system is highly relevant to the healthcare and assistive tech sectors. It supports patients with conditions like paralysis, locked-in syndrome, and age-related impairments. Being low-cost, real-time, and non-invasive, it fits well in hospitals, rehab centers, and home care. Its scalable, customizable design makes it suitable for diverse socio-economic groups, promoting inclusive and efficient care.

Working Model vs. Simulation/Study:

Arogya Samvad is a working model, not just a simulation or theoretical study. It uses live video input through a webcam to detect and interpret gestures and gaze in real time. Technologies like MediaPipe and Dlib enable accurate tracking and recognition, which are then mapped to predefined commands. The system sends voice and text alerts to caregivers using Telegram, demonstrating real-world usability. It was implemented, tested, and validated on physical hardware, showing reliable performance under various conditions. The model proves its potential for practical deployment in care settings.

Project Outcomes and Learning's:

Outcomes:

- A fully functional communication assistance system using eye-tracking and hand gestures.
- Successful real-time conversion of non-verbal cues into actionable alerts.
- UI design and integration with platforms like Telegram to deliver messages to caregivers.

Learning's:

- Application of computer vision and machine learning in a real-world assistive technology context.
- Hands-on experience with frameworks like OpenCV, MediaPipe, Flask, and Dlib.
- Deeper understanding of user-centered design, especially considering patients with severe impairments.
- Experience with end-to-end system development: from requirement gathering to design, coding, testing, and deployment.
- Importance of usability, reliability, and responsiveness in critical healthcare applications.

Future Scope:

Planned improvements include integrating deep learning models to boost accuracy in gaze and gesture detection across varied conditions. The system will explore additional input modes like voice or head movements and support personalized calibration for individual users. Hardware optimization will focus on creating compact, affordable, and wearable solutions. Enhancements for low-light performance and noise reduction are also prioritized. Lastly, expanding the command set will enable more intuitive and diverse interactions, making communication even more seamless and effective.