2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) | 979-8-3503-0570-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/RMKMATE59243.2023.10369794

AI Based Human Computer Interactions for Specially-Abled

Pandi C
Assistant Professor
Computer Science and Engineering
Vel Tech Multi Tech Dr.Rangarajan
Dr.Sakunthala Engineering College
Avadi,India
pandic24@gmail.com

Siva Rama Lingham N
Assistant Professor
Computer Science and Engineering
Vel Tech Rangarajan Dr.Sagunthala
R&D Institute of Science and
Technology,Avadi,India
sivaramalinghamn@veltech.edu.in

Nithya S
Assistant Professor
Information Technology
Vel Tech High Tech Dr.Rangarajan
Dr.Sakunthala Engineering College
Avadi,India
vaniselvin1212@gmail.com

Aruna .R
Associate Professor
Computer Science and Engineering
Vel Tech Rangarajan Dr.Sagunthala
R&D Institute of Science and
Technology,Avadi,India
arunasnandhini@gmail.com

***Abstract*— Artificial intelligence (AI) plays a crucial role in supporting people with special needs. People who are blind, deaf, or possibly disabled require ongoing extra care and attention when it comes to education because they have difficulties with everything from keeping track of attendance to completing tests. They cannot use computers for their needs; hence, they are dependent on someone. The use of a computer requires physical contact or interaction, such as using a mouse, keyboard, or touch screen. We have created a touch less sketching interface utilizing the pattern recognition method to overcome these problems. In pattern recognition, machine learning algorithms are used to automatically identify patterns and regularities in data that is given as words, pictures, sounds, or other recognizable components. Others include a virtual mouse that may be operated with the eyeball for those who are unable to move their hands. To quantify eye movement, common eye tracking algorithms analyze the separation between infrared LEDs projected onto the eye and the pupil of the eye. The mediapipe framework, the tensorflow model, and the openCV library are required for these two algorithms, which enable blind individuals to hear news, music, and other information without having to browse the internet.**

Keywords*— *Artificial Intelligence(AI), Touch less user interface, Eyeball-controlled virtual mouse, virtual desktopvoice assistant, Pattern Recognition, Eye Tracking.

I. INTRODUCTION

Human-computer interaction (HCI) is a field that studies how people interact with computers. The goal of HCI is to create user interfaces that are easy to use, efficient, and enjoyable. HCI is used in the design of all types of digital devices, from traditional desktop computers to mobile devices and wearable devices. HCI incorporates ideas and methods from several academic fields, including sociology, psychology, computer science, engineering, and human factors. Social computing, cognitive psychology, and user interface design are just a few of the many areas covered by the field.

II. LITERATURE REVIEW

1. Doshi-Velez and Kim's "A Systematic Review of Explainable Artificial Intelligence (XAI): Toward the Responsible Use of AI" *Appl. Sci.* **2022**, *12*(3), 1353. This review offers a thorough overview of the rapidly developing topic of explainable artificial intelligence (XAI), which attempts to increase user understanding and comprehension of AI systems. The authors examine more than 600 publications that were published between 2010 and 2019 in order to identify major XAI techniques and to indicate potential and problems for future research.

2. Oinas-Kukkonen and Harjumaa's "A Systematic Review of Persuasive Technology in Health and Wellness" (2018). This review investigates how persuasive technology is used to encourage healthy lifestyle choices. The authors review more than 300 research that were published between 2000 and 2016 to determine the most effective persuasion techniques.

3. Vasalou et al.'s "A Systematic Review of Social Virtual Reality: Designing for Social Interaction" (2020). The use of social virtual reality (VR) to promote social interaction is examined in this article. The authors review 51 studies that were published between 2014 and 2019 in order to identify critical design elements that support social presence in virtual reality (VR) and to explore obstacles and potential for further study.

4. By Vaidyam et al., "A Systematic Review of Virtual Agents in Mental Health Interventions." (2019). the use of virtual agents (VAs) in therapies for mental health is examined in this paper. In order to determine the many functions that VAs can play in mental health treatment and to highlight important design elements that enhance their efficacy, the authors examined 43 studies that were published between 2008 and 2018.

III. STRUCTURE OF HCI STYLING

The typical structure of human-computer interaction (HCI) consists of the user, the computer system, and the interface that connects them all. Each of these components interacts with the others in order to facilitate communication, interaction, and information sharing.
The user: This component also includes the individual using the computer system. Depending on their individual features, abilities, and interests, users interact with the system in different ways. User qualities include things like age, gender,

technological background, physical and mental prowess, and cultural background, to name just a few. Understanding these user characteristics is necessary for designing interfaces that are user-friendly and accessible for all users.

IV. SYSTEM REQUIREMENTS

The hardware and software of the computer system are both part of the computing equipment. The system might include input and output devices like keyboards, mice, and screens in addition to processing and storage components. The system's capabilities and limitations can have an impact on the interface's design and how the user and system communicate.

This component includes the design and implementation of the user interface, which is the means by which a user interacts with a computer system. The interface is made up of both digital and analogue components, including screens, buttons, menus, and dialogue boxes. The user interface should be designed to meet the needs of the user. This means considering the user's characteristics, such as their age, ability, and experience. It also means considering the user's preferences, such as their preferred style of interaction and their level of comfort with technology. Finally, it means considering the user's objectives, such as what they want to achieve by using the interface.

A. VScode

The free and open-source source code editor Visual Studio Code, also referred to as VS Code, was built by Microsoft. Windows, macOS, and Linux are just a few of the operating systems it is designed to work with. VS Code is one of the most popular code editors among developers all around the world and is well known for its adaptability, performance, and myriad customization options. Visual Studio Code is a powerful, flexible, and scalable code editor that provides a wide range of capabilities for modern software development. Developers regularly utilise it for a range of programming tasks due to its user-friendly interface, wide range of customization options, and supportive community.

B. Python

Python is a strong, adaptable, and multifaceted programming language that is employed in a wide range of industries. It is renowned for its expressiveness, readability, and simplicity. In 1991, Guido van Rossum debuted his Python programming language. It has a sizable and vibrant development community, and there are numerous libraries and frameworks available for it. The building of websites, scientific computing, data analysis, and artificial intelligence are all common uses for Python.

C. Tensorflow

Google's open-source TensorFlow machine learning library offers a base for developing and improving machine learning models. Like speech recognition, computer vision, and natural language processing, it is one of the best-known machine learning libraries and is widely used in both business and academia.

The concept of a computational graph, TensorFlow uses a directed graph to represent computations. The edges between nodes in the graph reflect the flow of data, and every single node in the network represents a mathematical operation .This makes TensorFlow a powerful tool for building machine learning models, as it allows for the efficient execution of complex computations. This graph-based approach enables TensorFlow to efficiently conduct computations on large datasets using distributed computing over several CPUs or GPUs.

D. Mediapipe

Building cross-platform applications on a variety of devices is possible with the help of the robust open-source machine learning framework known as MediaPipe. It was developed by Google and is still under active development. It provides a platform for building pipelines for processing and analysing multimedia data, such as images, videos, and audio. MediaPipe includes a set of pre-built models and components for common applications, including object detection, face detection, and hand tracking, in addition to tools for building and fine-tuning bespoke models.

Real-time processing capability is one of MediaPipe's key features, which makes it perfect for applications like gesture recognition, augmented reality, and virtual try-on. MediaPipe is designed to be portable across several platforms, including desktop, mobile, and embedded devices, and may be combined with a broad range of programming languages and frameworks, including C++, Python, and TensorFlow.

E. OpenCV:

The software library OpenCV (Open Source Computer Vision Library) is used for computer vision and machine learning. Intel first created it in 1999, it is now maintained by the OpenCV community. OpenCV provides a set of tools and techniques for the processing and analysis of visual data, such as images and movies. OpenCV is a cross-platform computer vision library.It is capable of supporting an extensive variety of computing platforms and programming languages. It can be used on Windows, Linux, macOS, Android, and iOS. The APIs for C++, Python, and Java are the most popular. It provides a wide range of modules for numerous computer vision applications, including processing pictures and videos, matching features, recognising and tracking objects, and machine learning.

V. PROPOSED SYSTEM METHODOLOGY

A. AI Virtual Sketch Board for people affected by Dyslexia:

Reading is challenging for people with dyslexia because they have trouble understanding how speech sounds relate to letters and words. It is not caused by IQ, hearing, or vision problems. Early detection and intervention are the best ways to help dyslexic children succeed in school. Emotional support is also important. Although dyslexia has a currently unidentified treatment, there are therapies that are capable of helping patients manage their symptoms. We recommend an AI-based touch less drawing board for people who are impacted by dyslexia.

B. Hand Landmark Model:

The Hand Landmark Model is a machine learning algorithm for hand tracking and gesture recognition. It is a part of Google's MediaPipe architecture, which provides

several pre-built models and pieces for managing multimedia data. The Hand Landmark Model is based on a convolutional neural network (CNN) architecture that has been trained on massive datasets of hand pictures in order to detect and identify key hand features, including the fingertips, knuckles, and wrist. The model can distinguish and follow the hand in real time, despite complex environments and poor lighting. The Hand Landmark Model has several uses, including hand gesture control, sign language recognition, and virtual try-on.

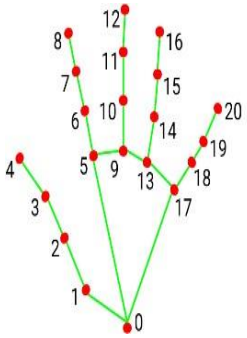

Fig. 1 Hand Landmark Model

C. Eye Ball Cursor Movement Using OpenCV:

Today, people who are paralyzed require assistance to complete any work. Only one person is required to care for the person. Using the eyeball tracking technique, we can track the eye movements of a paralyzed person by fixing the centroid on the eye. The centroid is the center of mass of the eye, and it can be used to track the eye's movement. This information can then be used to control a cursor or other device, allowing the paralyzed person to interact with their environment. This eyeball tracking technique uses a variety of programmers, such as home automation with Python GUI robotic control and a virtual keyboard programme.

Our proposed technique employs OpenCV to track eye movement and manage computer cursor movement. The camera picks up the eyeball's movement, which OpenCV then analyses. This makes cursor control possible.

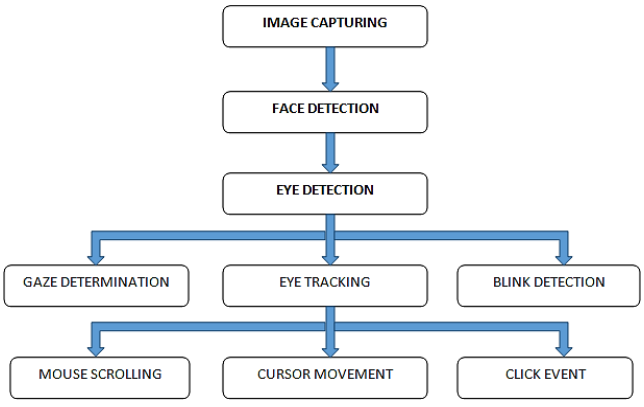

Fig. 2. Eye Controlled Mouse Block Diagram

The user is required to sit in front of a computer or laptop screen while an eye-testing camera is mounted above the screen. The laptop calculates where the user is looking on the screen while continuously analysing the attention video. Nothing is attached to the purchaser's head or body. While the user just needs to blink their eye to "press" a key, they must "pick out" a key by focusing on it for a specific amount of time. With this gadget, there is no need for a calibration process. The simplest choice in this system is enter. There is no connection to or need for additional gear.

The eye supplies input to the camera. As soon as the cameras' streaming movies arrive, they will degrade into frames. In order to prevent error warnings from displaying on the screen, cameras need appropriate lighting from outside sources; hence, after receiving frames, they will check the lighting conditions. The already-in-RGB-mode collected frames are transformed to black and white. Five. Iris detection is performed on images (frames) from the entrance supply that concentrate the eye. (centre of the eye). After that, a midway is determined by subtracting the indicated sites of the left and right eyes. Eventually, the mouse will move from one spot on the screen to another, and the user will click by blinking their eyes for five seconds.

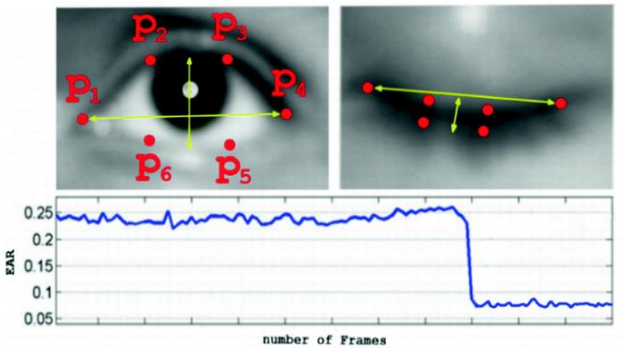

Fig. 3. Eye Mapping

D. Desktop Assistant-Jarvis

AI systems that can seamlessly facilitate natural human-machine communication are becoming increasingly popular. They can understand and respond to a wide range of human input modalities, including voice, dialogue, gestures, and facial expressions. One of the most studied and popular paths of interaction was the one based on the machine's understanding of the machine's natural human language.

Virtual assistants are computer programs that can help with everyday tasks, such as checking the weather, making shopping lists, and setting alarms. They can be controlled by voice or text, and they often require a wake word to activate. Some popular virtual assistants include Amazon Alexa, Apple Siri, and Microsoft Cortana.

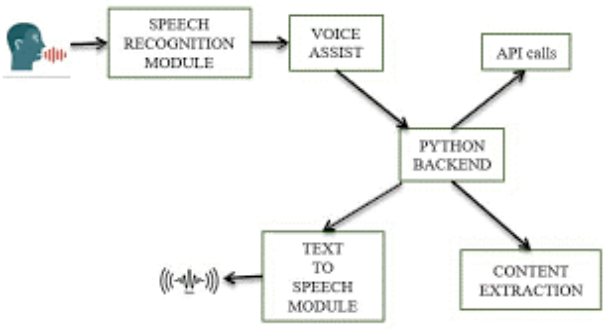

Fig.4. Methodology of Virtual Assistant Using Python

E. Facial Recognition Attendance System:

1) Face Detection:

Face detection is the skill of recognising individual individuals in digital photographs. This technology can recognise a human face in a photograph or video. To detect whether a photo or video has a face (or numerous faces), we need to define the general anatomy of a face. Eyes, noses, foreheads, mouths, and chins are common features on human faces. Finding the location and size of the face in a picture is the goal of face detection. The facial recognition system then makes use of the found face.

2) Feature Extraction:

In this phase, we are taking the features out of the discovered face. The first local binary pattern pictures are generated in LBPH, and a facial recognition histogram is produced. Thus, a template is created. A template is a collection of information that depicts the special and distinguishing characteristics of the detected face.

3) Face Recognition:

Face recognition is the ability to compare and analyse a biometric person's face in order to uniquely identify and authenticate that person. An application called a face recognition system is used to recognise or confirm a person from a digital image.

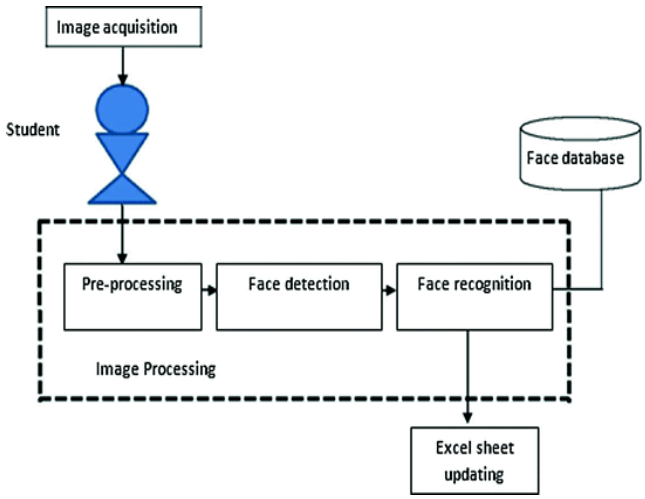

Fig. 5. Block Diagram for attendance and storage process

4) Image Capture And Image Processing

To get good results, we need an HD camera. We can manually capture each and every image from the webcam, or we can manually take the images from the video stream. While performing frame capture from the video stream would produce results faster, we won't be able to correctly record the face in the event that light is lost or if the face is not properly captured. A branch of digital signal processing called "digital image processing" employs algorithms to edit digital photos. Greater accuracy, adaptability, and scalability are just a few of its benefits over analogue image processing.

Digital image processing is a multidimensional system that can be modeled using discrete mathematics. It may minimise noise and distortion and enable the application of a greater variety of algorithms to the input data. The development of computers and the rise in demand for image processing applications in a variety of fields have all contributed to the growth of digital image processing.

5) Convolution Neural Network:

CNNs are a type of deep neural network that are commonly used for image analysis. They are characterized by their shared weight architecture and translation invariance properties, which make them well-suited for tasks such as image classification and object detection. CNNs are also used in natural language processing, recommender systems, and other domains. Multilayer perceptron's MLPs are neural networks that have multiple layers. In a fully connected network, each neuron in one layer is connected to every neuron in the next layer. This can lead to over fitting, As a result, MLPs are usually regularised by adding weights with a magnitude measurement to the loss function. Convolutional neural networks (CNNs) tackle regularisation differently. They combine more complex patterns out of smaller, simpler ones using the data's hierarchical structure. This makes CNNs less complex and less prone to over fitting than MLPs. Biological processes inspired CNNs, as each and every cortical neuron can only process information in a limited region of the visual field known as the receptive field.

VI. CONCLUSIONAND FUTURE ENHANCEMENTS:

It is concluded that using technologies like mediapipe, openCV, and tensorflow, we can interact with the system without physically touching it, so people with disabilities can easily interact with the computer. In the future, it will be combined into a single package and enhanced as an AI-based operating system.

REFERENCES

[1] Dayan Wang; Jue Qu; Wei Wang; Gangsheng Zhang; Shuai Zhu "The verification system for interface intelligent perception of human-computer interaction ," 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), IEEEXplore: 10 May 2021.

[2] Danny Plass-Oude Bos, Boris Reuderink, Bram van de Laar, Human-Computer Interaction for BCI Games, IEEE Xplore: 03 December 2010.

[3] Ran Liu; Lanlan Gao; Yijing Liu; Jian Le , "Research on the Usability Evaluation Model of Human-Computer Interaction for Command and Control System ," 2021 2nd International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) IEEE Xplore: 21 February 2022.

[4] Jingsi Ma; Yi Zhang; Hesiqi Bin; Kang Wang; Jinfang Liu; Hanrui Gao," The Development of Students' Computational Thinking Practices in AI Course Using the Game-Based Learning: A Case Study", 2022 International Symposium on Educational Technology (ISET), IEEE Xplore: 29 August 2022.

[5] Mobeen Nazar; Muhammad Mansoor Alam; Eiad Yafi; Mazliham Mohd Su'ud," A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques", IEEE Access (Volume: 9) Page(s): 153316 – 153348, 12 November 2021.

[6] Yunlong Wu; Jinghua Li; Haoxiang Jin; Jiexin Zhang; Yanzhen Wang," RBT-HCI: A Reliable Behavior Tree Planning Method with Human-Computer Interaction", 2022 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE Xplore: 18 January 2023

[7] Kening Zhu; Artur Lugmayr; Xiaojuan Ma; Florian Floyd Mueller; Ulrich Engelke; Simeon Simoff, " Interface and Experience Design with AI for VR/AR (DAIVAR'18) and AI/ML for Immersive Simulations (AMISIM'18) ", 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), IEEE Xplore: 17 January 2019.

[8] Rabail Tahir, "Analyzing the intelligence in user interfaces", 2015 SAI Intelligent Systems Conference (IntelliSys), IEEE Xplore: 21 December 2015.