Artificial Vision For Visually Impaired

A. Augustin Ronald[1], Prof. N. Sakthivel[2]
[1]Dept of Computer Applications
[2]Assistant Professor, Dept of Computer Applications
[1, 2] Adhiyamaan College of Engineering (Autonomous), Hosur, Tamil Nadu, India.

Abstract- *This paper presents the investigate the creation and application of artificial vision technologies with the intention of improving blind people's quality of life. The goal of visual implants is to create an artificial vision that can partially restore function. It can enhance the quality of life for visually challenged individuals by allowing them to perceive light, even if they have been blind for a long time. The artificial vision that is made possible by current visual system stimulators has very low resolution because of their small number of microelectrodes. Numerous researchers have sought to enhance artificial vision produced by low resolution implants through the application of machine vision and image processing techniques.*

Keywords- Upload training dataset, Admin authentication

I. INTRODUCTION

The project's goal is to improve the artificial vision experience for those with visual impairments by incorporating cutting-edge technology including image processing algorithms and vision transformers. The objective is to offer visually impaired people a more complete and transparent artificial vision system that enhances their mobility, freedom, and quality of life.

- To Use Vision Transformer Technology: For better artificial vision clarity, use Vision Transformer (ViT) technology.
- To Extract Detailed Data: Create real-time algorithms to extract data on the number of people, their familiarity, gender, age, emotions, items, and distances.
- To Facilitate Real-time Image Processing: Put into practice effective real-time image processing techniques for the analysis of live videos.
- To Improve Accessibility with Audio Output: Combine text- to-speech conversion with audio output to effectively communicate with visually challenged consumers.
- To Verify with Simulated Prosthetic Vision: To ensure compatibility, validate the system using a simulated prosthetic vision that can accommodate 150 microelectrodes.
- To Manage Versatile Tasks: Create algorithms that can manage tasks including object detection, face identification, gender prediction, age estimation, and emotion prediction.
- To Determine User Satisfaction: Determine user satisfaction via experimentation, user input, and contrast with current visual implant options.
- To Investigate Integration with Vision Implants: Investigate and show compatibility with both current and upcoming vision implant technologies.
- To provide the Foundation for Next-Gen Vision Systems: Use cutting-edge technologies to provide the framework for upcoming visual implant systems.
- To Determine Whether the System Is Feasible for everyday Use: Utilize a portable platform to attain a frame rate of 4.5 frames per second in order to determine whether the system is feasible for everyday use.

By giving visually impaired persons a dependable and effective way to recognize obstacles and people in their surroundings and safely navigate around them, the initiative seeks to enhance their quality of life and independence.

II. SYSTEM TESTING

A crucial stage of the software development life cycle is software testing, which entails a methodical assessment of a program to guarantee its operation, quality, and performance. Testing's main objectives are to find errors or faults and confirm that the program complies with requirements. Key components of software testing include the following:

Goal:

Examine and correct any flaws or faults in the software. Verify that the program satisfies all criteria and performs as planned.

- Improve the program's functionality, dependability, and quality.

Testing Types:

- Manual Testing: Without the use of automation tools, testers carry out test cases by hand. It entails the tester watching and evaluating the behavior of the application.
- Automated Testing: Testers run test cases using automation techniques. Regression testing, performance testing, and repetitive testing jobs all benefit from automation.

Methods of Testing:

- Black box testing: testing the program without being aware of its logic or its code. Inputs and outputs are the main focus.
- White Box Testing: Analyzing the software's internal logic, code, and structure. necessitates familiarity with internal mechanisms.

Life Cycle Testing:

- Test Planning: Outlining the general strategy and methodology to testing.
- Test Design: Developing test cases in accordance with specifications and requirements.
- Test execution entails executing the test cases and recording the outcomes.
- Defect Reporting: Finding and recording flaws.
- Test Closure: Recapitulating the steps taken and the conclusions reached.

Types of Software Testing:

1. Functional Testing:

- Unit Testing:

Unit testing involves testing individual units or components in isolation to ensure they function as intended. It is typically performed by developers during the coding phase to catch and fix bugs early.

- Integration Testing:

Integration testing checks the interaction between integrated components or systems. It verifies that the combined units work together seamlessly and identifies any issues that may arise from their integration.

- System Testing:

System testing evaluates the entire software system to ensure it meets specified requirements. It involves testing the software in its entirety to validate its functionality, performance, and reliability.

- Acceptance Testing:

Acceptance testing is performed to ensure that the software meets acceptance criteria and is ready for deployment. It involves validating whether the software satisfies user requirements and business expectations.

2. Non-Functional Testing:

- Performance Testing:

Performance testing evaluates the software's responsiveness, speed, and stability under various conditions. It includes load testing, stress testing, and scalability testing to identify performance bottlenecks and optimize the system.

- Security Testing:

Security testing identifies vulnerabilities in the software to ensure it is secure against unauthorized access, data breaches, and other security threats. It involves testing authentication, authorization, and data protection mechanisms.

- Usability Testing:

Usability testing assesses the user-friendliness and overall user experience of the software. It evaluates factors such as navigation, interface design, and user satisfaction to ensure the software is intuitive and easy to use.

III. EXISTING SYSTEM

These systems record the user's environment by using cameras that are attached to eyewear or other wearable technology. Computer vision algorithms are used to process the acquired images in real-time and extract pertinent information, such as text, obstructions, and facial expressions. The user is subsequently provided with tactile or audible feedback with the information that has been processed. The wearable Horus gadget and the OrCam MyEye are two examples.

A lot of applications for smartphones make use of the processor and built-in camera to enable artificial vision features. These programs have the ability to read language, recognize objects, and describe scenes. Examples include Microsoft's Seeing AI and Be My Eyes, which pairs visually impaired people with sighted volunteers to provide in-the-moment support.

To assist users in navigating both interior and outdoor areas, navigation systems for the visually handicapped combine GPS, inertial sensors, and computer vision. In order to direct users along predetermined paths, steer clear of obstacles, and locate objects of interest, these systems employ tactile or audible cues. Examples are the Sunu Band, which uses ultrasonic detection of obstructions, and the BlindSquare app.

BCIs are devices that communicate directly with the user's brain in order to interpret or restore visual data. In order to decipher neural activity, these systems usually need invasive procedures, including implanting electrodes in the visual cortex. They also rely on complex signal processing and machine learning techniques. Although BCIs are still in the early stages, they have the potential to give blind people high-resolution artificial vision.

All things considered, current artificial vision systems for the blind include a variety of features, ranging from simple object identification to sophisticated navigation and visual enhancement. Prolonged investigation and advancement in this domain have the potential to augment the self-sufficiency and standard of living of persons with visual impairments.

IV. PROPOSED SYSTEM

The proposed system aims to revolutionize artificial vision through the integration of a Vision Transformer-based approach, specifically designed for next-generation visual implants.

Artificial vision build and train:

In the initial phase of Artificial Vision Build and Train, the project focuses on assembling a diverse dataset, gathering images relevant to face recognition, gender prediction, age estimation, and emotion prediction. Following dataset collection, meticulous pre-processing steps, including grayscale conversion, resizing, noise filtering, and binarization, ensure standardized input conditions for optimal model training.

Advanced segmentation techniques like the Region Proposal Network (RPN) are then applied to identify and extract facial regions, providing a foundation for subsequent analyses. Feature extraction, utilizing methods like Local Binary Patterns (LBP), captures distinctive features crucial for nuanced tasks. The final step involves implementing a Convolutional Neural Network (CNN) for unified classification tasks, consolidating face recognition, gender prediction, age estimation, and emotion prediction into a cohesive framework. This comprehensive approach establishes a robust foundation for the subsequent development of an advanced artificial vision system.

Vision prediction system using vision transformer:

In the Vision Prediction System using Vision Transformer, real-time video is captured through a device-mounted camera, serving as input for subsequent analysis by the Vision Transformer. This technology predicts whether individuals in the user's vicinity are known or unknown, extracting features and enhancing spatial awareness.

In the presence of unknown individuals, the system predicts gender, age, and facial expression. The MiDaS algorithm estimates distances, and object detection identifies and locates objects, contributing to heightened situational awareness. To convey predictions, the system utilizes audio output with text-to-speech conversion, ensuring accessible feedback for visually impaired users.

Advantages of proposed system

- Enhances artificial vision for better clarity.
- Promotes greater independence for visually impaired users.
- Integrates real-time image processing for swift and dynamic analyses of live video feeds.
- Audio output ensures effective information delivery.
- Handles diverse tasks, from face recognition to object detection.
- Provides timely information for better situational awareness.

V. METHODOLOGY

The problems and tasks that the artificial vision system is meant to solve in the problem definition. This could include tasks related to the needs of visually impaired people, such as object recognition, text reading, navigation, and facial recognition.

Compile a variety of representative datasets with pictures or videos that are pertinent to the activities at hand. To guarantee the system's resilience and generalizability, these datasets ought to include a range of settings, luminaries.

Improve the quality and usefulness of the gathered data for further analysis by preprocessing it. This could entail methods like color correcting, denoising, scaling down images, and normalizing them in order to enhance the quality of the input for later processing stages.

To represent pertinent information for the intended tasks, extract significant features from the preprocessed data. Handcrafted feature extraction techniques like edge detection, texture analysis, or corner detection may be used in traditional approaches. From the raw data, deep learning models may automatically extract hierarchical features.

Determine whether machine learning or deep learning models are best for the tasks at hand and the data that is available. Using the provided datasets, train the chosen models, modifying optimization techniques and hyperparameters as needed to get acceptable results.

Assess the accuracy, speed, robustness, and usability of the trained models by utilizing cross-validation techniques or independent validation datasets. To increase performance, make necessary iterations to the model's design and training regimen.

Construct a working system or application that enables visually impaired people to use the trained models. This could entail creating user interfaces, creating interaction modalities (such tactile or aural feedback), and adjusting the system to work in practical situations.

To get input on the efficacy, accessibility, and usability of the system, do user testing with visually impaired people. To fix any problems and raise consumer happiness, iterate the design in response to customer input.

Keep an eye on the artificial vision system at all times, fixing any glitches, decreasing performance, or changing user requirements. To maintain the system's applicability and efficacy over time, update it frequently with fresh information and enhanced algorithms.

VI. WORKING

Artificial vision systems for the blind operate by utilizing a number of essential parts and procedures, which are usually carried out by combining hardware and software:

Using cameras installed on wearable technology, cellphones, or other devices, the procedure starts with taking pictures or video frames of the user's environment. Depending on the particular application, these cameras may be positioned to offer a larger field of view or a first-person perspective.

Relevant visual information is extracted from the collected photos in real-time using computer vision algorithms. Tasks including object detection, text recognition, scene segmentation, and depth estimation may be a part of this processing. For this, a variety of approaches, such as deep learning models and conventional image processing techniques, may be used.

Key visual signals including objects, text, barriers, and surrounding context are represented by features that are extracted from the processed images. These qualities aid in the user's comprehension of the scene and the identification of pertinent information.

Following the extraction of features, the system classifies and identifies any relevant objects, text, or other scene elements. In order to classify objects, recognize language, or carry out other recognition tasks, machine learning algorithms trained on annotated datasets are frequently used in this step.

Depending on the system's architecture and the user's choices, the identified visual information is provided to the user via either tactile or audible feedback. This feedback could come in the form of vibrations, synthesized speech, or spoken descriptions.

The system helps the user safely navigate the surroundings by providing guidance and aid based on the processed visual information. This could entail pointing out impediments, directing the user along routes that have been predetermined, alerting them to potential dangers, and helping them with tasks like reading signs or finding interesting objects.

Voice commands, gestures, and tactile interfaces are just a few of the input modalities that users can use to interact with the artificial vision system. Through these interactions, users can ask for specific information, take control of the system, and express their requirements and preferences.

To enhance their performance over time, artificial vision systems may integrate adaptive algorithms that take in feedback and user interactions. This adaption might entail changing the parameters.

VII. CONCLUSION

In summary, the project's goal is to completely transform the artificial vision experience for those with visual impairments by combining cutting-edge technologies and creative solutions. By utilizing Vision Transformer technology, real-time image processing algorithms, and information extraction methodologies, the project aims to improve visually impaired individuals' overall quality of life, independence, and accessibility.

Through tackling major issues such restricted information availability, difficulties with navigation, restricted social interactions, and barriers to education and work, the project aims to enable visually impaired people to live more fulfilled and independent lives.

Accessible feedback is ensured by the integration of audio output with text-to-speech conversion, and the project's potential impact is further reinforced by validation using a simulated prosthetic vision and a feasibility study for daily use.

Contributions to the larger area of artificial vision technology in addition to trying to enhance the quality of life for people who are visually impaired by offering an enhanced artificial vision experience. Through promoting independence, accessibility, and user pleasure, the initiative hopes to establish a standard for next advancements in visual implant systems.

Furthermore, the research establishes the foundation for ongoing developments in artificial vision by investigating integration with current visual implant technologies and building the framework for next-generation visual implant systems. The project intends to continuously refine and improve the produced solutions to better fulfill the needs of visually impaired individuals through feedback collecting and user satisfaction surveys.

REFERENCES

[1] F. Catherine, Shiri Azenkot, Maya Cakmak, "Designing a Robot Guide for Blind People in Indoor Environments," ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, 2015.

[2] H. E. Chen, Y. Y. Lin, C. H. Chen, I. F. Wang, "Blindnavi: a mobile navigation app specially designed for the visually impaired," ACM Conference Extended Abstracts on Human Factors in Computing Systems, 2015.

[3] K. W. Chen, C. H. Wang, X. Wei, Q. Liang, C. S. Chen, M. H. Yang, and Y. P. Hung, "Vision-based positioning for Internet-of-Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no.2, pp. 364–376, 2016.

[4] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[5] J. Ducasse, M. Macé, M. Serrano, and C. Jouffrais, "Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users," ACM CHI Conference on Human Factors in Computing Systems, 2016.

[6] J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large- scale direct monocular SLAM," European Conference on Computer Vision, 2014.

[7] S. Gilson, S. Gohil, F. Khan, V. Nagaonkar, "A Wireless Navigation System for the Visually Impaired," Capstone Spring, 2015.

[8] J. Guerreiro, D. Ahmetovic, K. M. Kitani, and C. Asakawa, "Virtual Navigation for Blind People: Building Sequential Representations of the Real-World," International ACM SIGACCESS Conference on Computers and Accessibility, 2017.

[9] Kendall, M. Grimes, and R. Cipolla, "PoseNet: a convolutional network for real-time 6-DOF camera relocalization," International Conference on Computer Vision, 2015.

[10] Kendall, and R. Cipolla, "Geometric loss function for camera pose regression with deep learning," International Conference on Computer Vision, 2017.