Design for Inclusion, Vol. 169, 2025, 68–73
https://doi.org/10.54941/ahfe1006155

An AI Powered Glasses Attachment for the Visually Impaired

Khac Trong Nguyen and Jan-Torsten Milde

Fulda University of Applied Sciences, Fulda, Germany

ABSTRACT

This research explores a prototype AI-powered glasses attachment designed to aid visually impaired individuals with daily tasks. The device integrates object and text recognition, alongside speech interaction, aiming to improve navigation and information access. Hardware comprises a microcontroller, camera, and touch sensor, encased in a 3D-printed enclosure. User testing, involving participants with varying visual impairments, assessed the system's functionality and usability across different modes. The results highlighted strengths in object/text recognition but exposed limitations related to camera performance, particularly in low light, and touch sensor usability for older individuals. Further refinement is needed, but the glasses show promise as an assistive technology.

Keywords: Visual impairment, Assistive technologies, Human systems integration, User-friendly design, Accessibility, Speech interaction

INTRODUCTION

This research presents the development and evaluation of a prototype AI-powered glasses attachment designed to enhance the daily mobility and independence of individuals with visual impairments. The work addresses the significant challenges faced by millions who experience limitations in navigating, recognising objects, and accessing information. It aims to contribute to the field of assistive technologies by creating a cost-effective and versatile solution. The project emphasises the importance of designing inclusive technology that is not only functional but also user-friendly and accessible to all. The system has several strengths:

- Compact and portable design making it a discrete companion in daily life.
- Versatile functionality combining object/text recognition with speech interaction.
- Cost-effective approach making it accessible to a broader range of users.

The project involved a thorough analysis of user requirements, focusing on the needs of individuals with varying degrees of visual impairment. This included understanding the challenges faced by people with different levels of sight loss, from moderate to complete blindness. The study also highlighted the importance of intuitive design and accessibility.

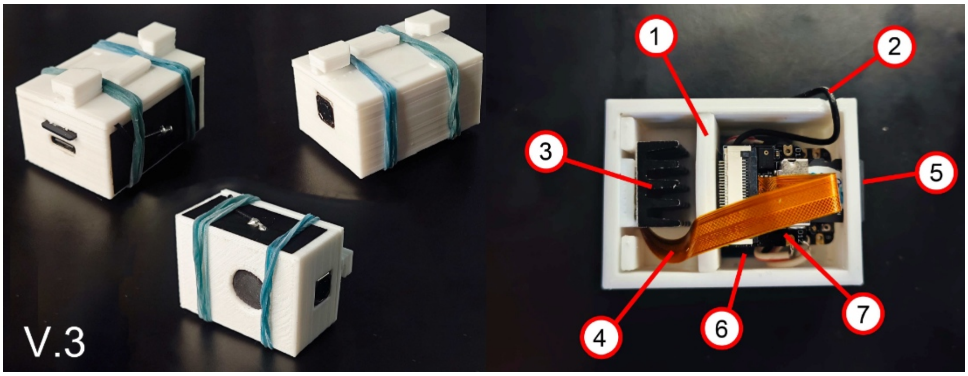

Figure 1: 1. Slide-in module, 2. WiFi antenna, 3. OV5640 camera module + heatsink, 4. Camera ribbon cable, 5. SD card slot + USB C port, 6. TTP223B touch sensor, 7. Pluggable camera sensor board.

Technology of the Prototype

The core of the prototype integrates real-time object and text recognition, along with speech interaction capabilities. The system uses image processing algorithms, particularly convolutional neural networks (CNNs) to identify objects. Optical character recognition (OCR) is implemented to transform printed text into digital formats, enabling text-to-speech functionality. The prototype also incorporates a natural language processing (NLP) powered chatbot, facilitating spoken interaction and information retrieval. A video demonstration can be found under https://youtu.be/f1j1bR50r1c?si=CHybS5tz1M1qg_IB.

Hardware Design

The hardware design incorporates a compact ESP32-S3 microcontroller coupled with an OV5640 camera module, chosen for their balance of performance and power efficiency. The software is developed using Python for data processing and C++ for the microcontroller. Cloud-based AI services, including Google Cloud Text-to-Speech API and OpenAI's GPT API and Whisper API, are used for text-to-speech, object and text recognition, and speech recognition, respectively.

The 3D-printed enclosure for the attachment was designed with several key considerations, including compactness, component integration, and user comfort. Key aspects of the hardware design are:

- Component Layout: The enclosure was designed to house the ESP32 microcontroller, the OV5640 camera module, and a TTP223B touch sensor. The internal structure includes a separate compartment for the camera with a heatsink to manage its operating temperature.
- Touch Sensor Integration: The touch sensor is located on the underside of the enclosure, designed with a concave depression for easy access and covered with a rubber membrane. This placement allows for user interaction and control of the system's functions.

- Size and Weight: The prototype has a weight of 23 grams with dimensions less then 5cm by 5 cm by 3 cm. Its design aimed to be as small as possible whilst being able to accommodate the necessary hardware.

The conducted functional tests focused on the accuracy of text and object recognition, and the efficacy of the chatbot. The text and object recognition performed well, however, issues with the camera quality, particularly in low-light conditions, limited the system's performance with smaller details and text. The speech interaction, while functional, encountered some difficulties with complex questions. The prototype is promising, future development should focus on addressing its limitations.

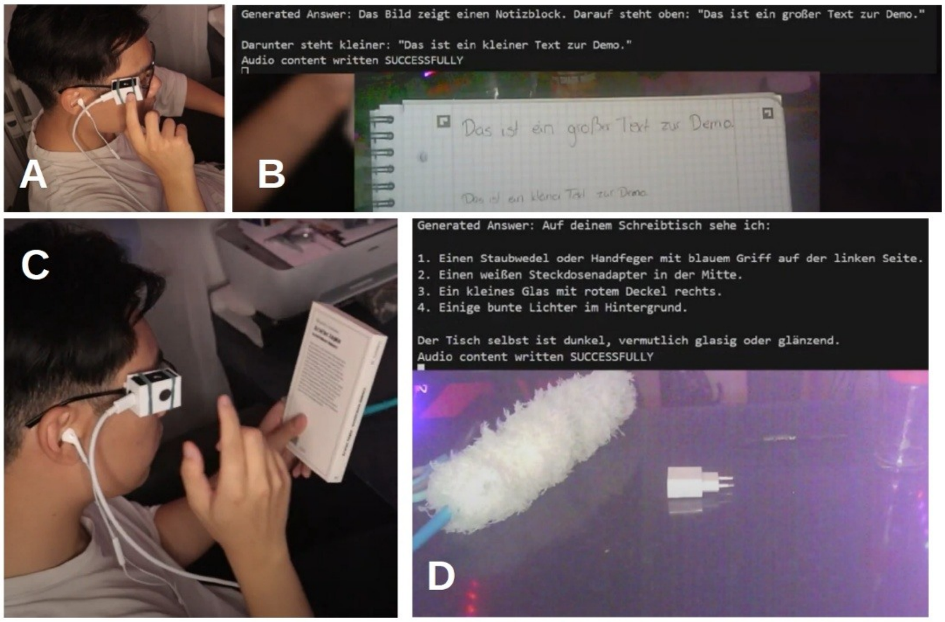

Figure 2: A. Select operation mode with touch sensor, B. OCR of hand written text, C. OCR of printed text, D. Scene description of surrounding environment.

User Tests

A set of user tests were conducted to evaluate the system. The study involved 20 participants with varying degrees of visual impairment.

The aim of the user test was to comprehensively assess the accessibility and user-friendliness of the KI system. The tests were carried out in a controlled, realistic environment to investigate the impact of different variables on system performance.

METHODOLOGY

The methodology employed a mixed-methods approach, combining qualitative and quantitative data collection, to evaluate the system's performance. Testing was conducted in controlled, enclosed environments, such as living rooms and kitchens, to minimize external variables like fluctuating lighting and ambient noise. A standardized testbed was established to ensure consistent scenario presentation across all participants. The test procedure involved a rigorous assessment of each operational mode—text recognition, object recognition, and chatbot functionality—followed by structured interviews. These interviews consisted of three targeted questions designed to elicit detailed feedback specific to the tasks performed within each mode, allowing for a comprehensive understanding of the system's usability and effectiveness.

While the test subject performed the tasks, their behaviors, specific approaches, and reaction times were recorded. In the subsequent short interview, they were asked whether the recognized text could be understood and whether any difficulties arose with certain fonts or sizes. They were also asked whether a specific positioning could achieve a better result. Following this, the test subject was asked to evaluate this mode. This allows the test subject to rate the mode on a scale of 1–5 (1 being very poor, 5 being very good). The process is the same for the remaining two modes.

PERFORMANCE TESTS

The performance test was divided into four sections. To evaluate the system's functionality across diverse applications, participants engaged in a series of targeted tasks designed to simulate real-world scenarios. The Text Recognition mode was assessed through reading exercises involving short texts on paper, digital screens (smartphones and computer monitors), and product packaging (food and medication). Object Recognition capabilities were tested by requiring participants to locate common household items (e.g., remote controls, glasses), identify food items (fruits, beverages), and recognize specific objects (flower types, individuals). Finally, the Chatbot functionality was evaluated by prompting participants to pose questions and subsequently provide feedback on the accuracy and relevance of the system's responses, thereby assessing its conversational efficacy and information retrieval capabilities.

DATA COLLECTION

Data collection during the study employed a multi-faceted approach to comprehensively assess system performance and user experience. We recorded participant behavior, specific task execution strategies, and reaction times during each task. Following each task completion, short structured interviews were conducted to gather qualitative feedback. These interviews focused on assessing the participant's comprehension of recognized text, identifying challenges associated with specific font types or sizes, and evaluating the influence of device positioning on recognition accuracy. Furthermore, participants provided quantitative evaluations of each operational mode (text recognition, object recognition, and chatbot) using a 5-point Likert scale, where 1 represented "very poor" and

5 represented “very good,” enabling a systematic quantification of user satisfaction and perceived system effectiveness.

OUTCOMES

The user tests yielded crucial insights into the system’s operational strengths and weaknesses, revealing key factors for successful implementation among visually impaired users. Notably, the importance of high-quality camera resolution, rapid processing speeds, and intuitive, user-friendly design was underscored. The evaluation also demonstrated a discernible age-related difference in system adaptation, with younger users readily embracing the technology, while older participants expressed a preference for tactile physical buttons over touch-based interfaces. The collected user feedback was instrumental in pinpointing specific areas requiring refinement, directly informing the development of targeted improvements for subsequent iterations of the AI-supported glasses attachment.

Table 1: Evaluation of functions by test subject.

Criterion	Evaluation	Percentage of Test Subjects
Recognition of Text on Paper	Good	80%
Recognition of Text on Packaging	Very good	85%
Speed and Precision	Very good	85%
Answer Accuracy	Good	70%
Intuitive Operation	Very good	65%
Wearing Comfort (Weight)	Acceptable	30%
Average Overall Rating	7.5/10	100%

CONCLUSION

The AI-powered glasses attachment demonstrated notable strengths in text and object recognition, particularly with clear visuals and larger objects, while the chatbot’s familiarity proved highly beneficial. Younger users found the interface intuitive, highlighting the system’s potential for user-friendly assistive technology. However, limitations persisted with small font sizes and occasional speech input errors, indicating areas for future refinement.

The evaluation also revealed limitations of the approach: specifically older users encountered difficulties with touch controls, favoring physical buttons, while the device’s weight and wired design presented challenges for comfortable daily use.

These factors highlight key areas requiring improvement for future iterations. Overall, the presented AI glasses demonstrate promising potential. Potential improvements include an automated configuration process, the integration of a navigation system and a more robust camera. Enhancing the user interface with a more tactile feedback system could improve accessibility.

REFERENCES

A Ali, Sanjeev Rao, Swaroop Ranganath, Ashwin T S, and Rammohana Reddy Guddeti (2021). A google glass based real-time scene analysis for the visually impaired. IEEE Access, PP:1±1, 12 2021.

Cremers and Verhaeghe (2020). The impact of visual impairment on quality of life: A review of the literature. International journal of environmental research and public health, 17(11):3896, 2020.

macularsociety.org (2021). New smart glasses help people with macular disease to read, recognise faces and identify objects, 2021.

Subhankar Bhowmick and SM Hazarika (2020). Smart glasses for visually impaired: A systematic review. Journal of King Saud University-Computer and Information Sciences, 32(10):1193±1202, 2020.

Vision Science Academy (2023). Smart vision glasses ± gen 2.0: Even without sight, there is still vision, 2023.

World Health Organization (2021). World report on vision. 2021.

Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu (2020). Object detection with deep learning: A review. IEEE transactions on pattern analysis and machine intelligence, 40(11):2547±2563, 2020.