This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3554366

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Low-Cost Multifunctional Assistive Device for Visually Impaired Individuals

Richard Kral[1], Patrik Jacko[1] and Tibor Vince[1]

[1] Department of Theoretical and Industrial Electrical Engineering, Technical University of Košice, Košice, 040 01, Slovakia

Corresponding author: Tibor Vince (e-mail: tibor.vince@tuke.sk).

This research was partially supported by the Slovak grant Nadácia Tatra Banky.

ABSTRACT Visually impaired individuals face challenges in mobility, object recognition, and text reading. Existing assistive tools provide partial solutions but lack full integration. This paper presents a low-cost, multifunctional assistive device in the form of wearable glasses, integrating real-time text recognition, obstacle detection, and remote assistance. Built around a Raspberry Pi Zero 2W, the system features an Arducam IMX519 autofocus camera and a tactile button interface for user-friendly control. Real-time audio feedback enhances situational awareness, promoting greater independence. The device was tested at the Rehabilitation Center for the Visually Impaired in Levoča. Performance evaluations confirm high text recognition accuracy and improved navigation support. A comparative analysis highlights its affordability and functionality over commercial alternatives. While effective, the device requires processing speed optimization and hardware miniaturization. Future improvements will explore haptic feedback and AI-driven scene analysis to enhance usability. This research contributes to affordable and inclusive assistive technology, improving accessibility for visually impaired individuals.

INDEX TERMS Assistive technology, wearable computers, visually impaired, artificial intelligence, text recognition.

I. INTRODUCTION

Assistive technology plays a crucial role in improving the mobility and independence of visually impaired individuals. With over 285 million people worldwide experiencing some form of visual impairment [1],[2], there is a growing demand for cost-effective, multifunctional solutions that enhance navigation, independence, and access to information.

Existing tools such as white canes, guide dogs, and smartphone applications offer some level of assistance but exhibit limitations in real-time object detection, environmental awareness, and user adaptability. Research in mobility technologies for visually impaired individuals has advanced significantly [3] – [8] . However, most commercially available solutions remain prohibitively expensive, often costing thousands of euros, making them inaccessible to a large segment of users.

To address these challenges, this paper presents the development of a low-cost multifunctional wearable assistive device in form of assistive glasses designed to address key mobility and accessibility challenges faced by visually impaired individuals. Unlike traditional assistive tools such as white canes and smartphone applications, which provide only partial solutions, this system integrates real-time text recognition, obstacle detection, and remote volunteer assistance into a single compact and wearable device. By leveraging a Raspberry Pi Zero 2W for processing, an Arducam IMX519 autofocus camera for visual input, and a tactile button interface for simplified control, the device ensures accessibility and ease of use while offering real-time situational awareness.

A comparative evaluation against commercially available assistive technology highlights the advantages of this system in providing affordable, multifunctional support for visually impaired users. The study also explores the usability of the device through real-world testing, revealing valuable insights into areas for improvement, including processing efficiency, alternative feedback mechanisms, and hardware miniaturization. By addressing these aspects, the research contributes to the development of more accessible, cost-effective, and inclusive assistive technologies that promote greater independence and mobility for visually impaired individuals

II. RELATED WORKS

Several assistive technologies have been developed to support visually impaired individuals by leveraging artificial intelligence (AI), wearable sensors, and real-time processing to enhance accessibility. These can be categorized into the following main approaches:

A. WEARABLE AND SENSOR-BASED TECHNOLOGIES

Wearable technologies and sensor-based solutions have been designed to assist visually impaired users in navigation and environmental awareness. Smart clothing integrates sensory systems to detect obstacles and improve mobility [9]. Another notable development is the IoT-Smart Stick, which combines sensors with internet connectivity to enhance real-time hazard detection and safety [10].

Additionally, smart glasses have been explored to provide autonomous navigation and real-time access to textual information [11]. While these innovations offer enhanced mobility, they often lack affordability and comprehensive integration of multiple functionalities.

B. AI-DRIVEN SCENE AND TEXT RECOGNITION

AI-based solutions have significantly improved accessibility for visually impaired individuals, particularly in text recognition and scene analysis. Attention-Based Convolutional Recurrent Neural Networks have been developed to enable real-time scene text recognition [12]. Similarly, enhanced VGG-based algorithms have been applied in visual prosthetics for object recognition [13].

Lightweight frameworks such as SqueezedText, a binary convolutional encoder-decoder network, have optimized real-time text recognition, making AI-assisted reading more efficient [14]. Despite these advancements, many AI-based systems focus solely on text recognition without integrating navigation or real-time assistance functionalities [15].

C. SCENE ANALYSIS AND SPATIAL AWARENESS

Beyond text recognition, scene analysis and spatial awareness play a crucial role in assistive technologies. Deep learning-based Robotics Framework for Augmented Simulated Prosthetic Vision (RASPV) has been adapted to provide visually impaired users with enhanced spatial perception [16].

Similarly, Simultaneous Localization and Mapping (SLAM) algorithms have been employed to improve real-time navigation capabilities [17], [3]. These approaches enable users to develop an understanding of their surroundings, but they often require complex hardware setups that may not be cost-effective.

D. SENSORY SUBSTITUTION APPROACHES

Sensory substitution techniques offer alternative ways for visually impaired users to perceive their surroundings. The vOICe algorithm, for example, converts visual input into auditory signals, allowing users to interpret spatial information through sound [18]. In addition, wavelet-based real-time hazard warning models have been introduced to enhance road safety for visually impaired pedestrians [19].

While these methods offer innovative solutions, they often require extensive user training, which may hinder widespread adoption.

E. CHALLENGES IN EXISTING TECHNOLOGIES

Despite the rapid advancements in assistive technologies, several limitations persist. One of the most significant challenges is the high cost of commercially available solutions, which often makes them inaccessible to a large segment of visually impaired individuals. Many of these devices are priced beyond the financial reach of users who would benefit most from them. Another major limitation is the lack of integration, as most existing assistive tools are designed to address only a single functionality, such as text recognition or object detection, rather than providing a comprehensive, all-in-one solution. Additionally, usability concerns remain a barrier to adoption, as some assistive systems are difficult to learn and operate, making them challenging for visually impaired individuals to use effectively. These challenges highlight the need for an affordable, multifunctional, and user-friendly assistive device that integrates multiple essential features while maintaining accessibility and ease of use.

This paper presents a cost-effective, multifunctional assistive device that integrates AI-based scene description, text recognition, and real-time remote assistance to address the above limitations. By combining affordable hardware components with optimized software functionalities, the proposed system offers a scalable and practical solution for visually impaired individuals, bridging the gap between high-end commercial devices and affordability.

III. USER NEEDS ANALYSIS

To ensure the proposed device effectively addresses real-world challenges, a qualitative study was conducted with tens of visually impaired individuals from the Rehabilitation Center for the Visually Impaired in Levoča, Slovakia. The study aimed to identify key mobility challenges, limitations of existing assistive technologies, and user preferences for an optimal solution.

A. PARTICIPANT SELECTION AND STUDY METHODOLOGY

Participants were selected to represent a diverse range of backgrounds, including age, gender, and varying degrees of visual impairment. The study included individuals with both congenital blindness and acquired vision loss, ensuring a broad spectrum of experiences and perspectives.

Data collection was conducted through structured interviews, designed to explore participants' daily challenges, experiences with existing assistive technologies, and their expectations for an ideal device. Each participant took part in an individual interview lasting approximately one hour. The questions were categorized into three main areas:

1. **Daily Challenges**: Difficulties encountered in both outdoor and indoor navigation, as well as access to environmental information (e.g., reading bus numbers, identifying street names).
2. **Technology Usage**: Evaluation of currently used tools, such as white canes, guide dogs, or smartphone applications like NVDA and Lazarillo
3. **Device Expectations**: Desired features in an assistive device, including real-time obstacle detection, text recognition, and scene description.

The structured interviews provided valuable insights into the challenges, needs, and expectations of visually impaired individuals, forming the basis for the development of a more effective assistive device.

B. KEY MOBILITY CHALLENGES

Participants consistently highlighted significant obstacles in mobility and navigation, particularly in unfamiliar or crowded environments. A recurring issue was the difficulty in maintaining direction and detecting obstacles, especially those located above ground level.

While white canes remain a widely used mobility tool, they were frequently described as insufficient for detecting overhead hazards, such as: tree branches, hanging signs, protruding shelves, among others. These limitations often led to collisions and injuries. Additionally, participants expressed frustration with their limited ability to access environmental information, such as reading street names, identifying bus numbers, and locating building entrances. These challenges were further compounded in complex indoor spaces including but not limited to shopping malls, offices or transport hubs.

A lack of tactile or audio guidance systems in such environments posed additional barriers. Furthermore, reliance on passersby for assistance was described as socially uncomfortable and inconvenient, emphasizing the need for a device that promotes greater autonomy in navigation.

The main challenges faced by visually impaired individuals, expressed as percentages, are shown in Figure 1.

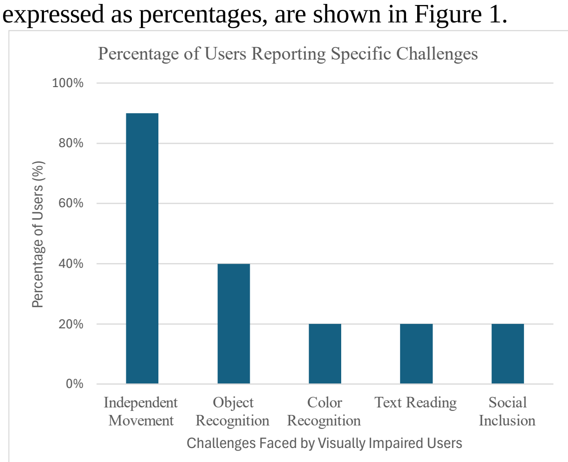

FIGURE 1. Main challenges faced by visually impaired individuals.

C. EVALUATION OF EXISTING ASSISTIVE TECHNOLOGIES

The study also examined the effectiveness and limitations of existing assistive tools used by visually impaired individuals. White canes and guide dogs remain the most common solutions for mobility, providing essential tactile or guided navigation support. In addition to these traditional methods, smartphone applications equipped with accessibility software, such as NVDA, TalkBack, and JAWS, help users read text and navigate their surroundings more effectively. Furthermore, GPS-based applications, including Lazarillo and Corvus, offer turn-by-turn navigation specifically designed for visually impaired individuals. While these tools provide valuable assistance, they often lack full integration, requiring users to rely on multiple devices to address different challenges.

While these tools provide valuable support, they still have significant limitations in terms of functionality and integration. GPS applications help users navigate but do not offer real-time obstacle detection or text recognition, making independent mobility more challenging. Many assistive applications also struggle with response times and become less effective in visually complex environments, such as crowded streets or poorly lit areas, where quick and accurate feedback is essential. Additionally, the high cost of these technologies and the difficulty of learning how to use them create barriers for many users, especially those with financial limitations or limited experience with digital tools. These challenges highlight the need for a more affordable, fully integrated assistive solution that is easy to use and provides real-time support across different environments.

The functional limitations of assistive technologies based on user feedback, expressed as percentages, are shown in Figure 2.

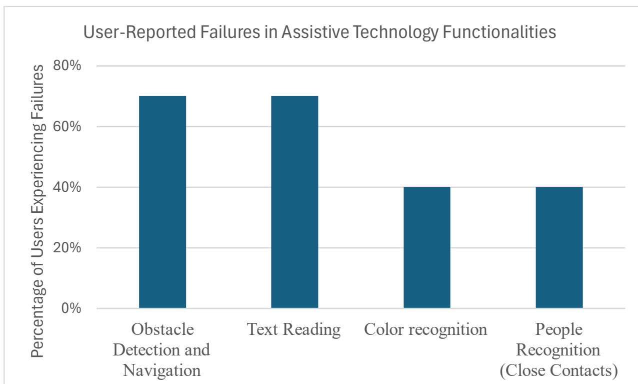

FIGURE 2. Functional limitations of assistive technologies based on user feedback.

D. USER EXPECTATIONS FOR AN OPTIAL ASSISTIVE DEVICE

When describing their ideal assistive device, participants expressed a strong preference for a solution that integrates multiple functionalities into a single, user-friendly system. Real-time navigation and obstacle detection were identified as

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3554366

essential features, enabling users to maintain situational awareness and navigate safely without external assistance. Another critical requirement was integrated text recognition, allowing users to read street signs, product labels, bus numbers, and other printed materials across various languages.

Additionally, participants highlighted the value of scene description and object identification capabilities, which would further enhance independence and safety. A wearable, hands-free design was emphasized as a key requirement, ensuring that the device could be used without interfering with other tasks. Simplicity of operation and affordability were also prioritized, as these factors are crucial in ensuring accessibility for a broad audience, including individuals with limited technical proficiency.

Participants' expectations can be summarized as follows:

- Integrated Multi-Functionality: A strong preference for a device that combines real-time navigation assistance, text recognition, and scene description into a single system.
- Hands-Free Operation: A wearable format was deemed highly desirable to enable ease of use without interfering with mobility.
- Affordability & Accessibility: Cost was a major concern, with participants emphasizing the need for a budget-friendly solution that remains technologically robust.

The findings underscored the significant potential of assistive technologies to improve the daily lives of visually impaired individuals, while also revealing critical design and technical challenges. A key priority was the development of compact, lightweight hardware that remains comfortable during extended use. To address functional requirements, the device would need to incorporate advanced artificial intelligence algorithms capable of performing real-time obstacle detection, text recognition, and scene analysis within the computational constraints of a portable system.

Additionally, the interface must be intuitive enough to accommodate users with diverse technical backgrounds while providing the necessary depth for more advanced users. The need for affordability and accessibility further complicates the challenge, requiring innovative approaches to both hardware and software development.

This analysis provided the foundation for defining the functional requirements of a new assistive device, highlighting the importance of user-centered innovation. By addressing the identified challenges and integrating the desired features, the proposed system has the potential to transform the daily experiences of visually impaired individuals, promoting independence and enhancing their quality of life. These insights not only guided the development process but also underscored the broader impact of assistive technology in fostering social inclusion and equal opportunities.

IV. SYSTEM PROTOTYPE DEVELOPMENT

The assistive system prototype was designed as a wearable device in the form of assistive glasses, addressing the needs identified in previous analyses. This form factor was chosen to enable hands-free operation while integrating multiple functionalities, ensuring continuous accessibility and ease of use for visually impaired users. The development process focused on creating a lightweight, portable device capable of providing real-time assistance in various scenarios.

The block diagram of the proposed solution is shown in Figure 3. A single-board computer serves as the core processing unit, interfacing with input and output peripherals, while key servers provide AI-based or volunteer-assisted support.

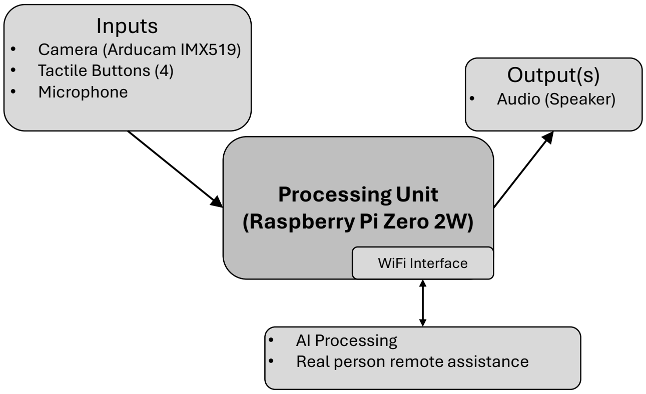

FIGURE 3. **Block diagram of the proposed solution.**

A. HARDWARE DESIGN

The selection of hardware components for the assistive glasses was guided by requirements for portability, functionality, affordability, and cost efficiency. Each component was chosen to ensure compatibility, performance, and user comfort, while maintaining a lightweight and compact design. Additionally, a strong emphasis was placed on selecting cost-effective components to develop a low-cost yet reliable assistive device, making it accessible to a broader audience, including individuals with limited financial resources.

The hardware configuration of the prototype is centered around the Raspberry Pi Zero 2W single-board computer, as shown in Figure 4. This board was selected due to its compact dimensions (65 mm × 30 mm × 13 mm), low weight (16 grams), and energy efficiency. Its quad-core Broadcom BCM2710A1 processor (64-bit, clocked at 1 GHz) provides sufficient computational power for operations such as real-time text recognition and scene analysis. The board integrates 2.4 GHz Wi-Fi and Bluetooth 4.2, enabling seamless wireless communication. Compared to alternatives such as the Raspberry Pi 4B or 3A+, the Zero 2W offers adequate computational power for OCR, Text-to-Speech (TTS) synthesis, and real-time image analysis, while being significantly smaller and more energy-efficient.

The Arducam IMX519 autofocus camera, displayed in Figure 4, was selected as the primary imaging module due to its high-resolution capabilities and compatibility with

Raspberry Pi systems. The 16-megapixel sensor ensures high-quality image acquisition, making it well-suited for applications requiring detailed visual analysis. With an 80° field of view, the camera aligns optimally with the user's line of sight, facilitating effective capture of environmental and textual data. Additionally, the motorized autofocus feature enhances its utility by ensuring sharp images at varying distances. Compared to alternatives such as the Raspberry Pi Camera v1 and v2, the Arducam IMX519 offers superior autofocus capabilities, making it the preferred choice for advanced applications requiring precision and adaptability.

The audio system utilizes an external USB sound card, which provides clear audio feedback and supports voice interactions. The tactile input system consists of four strategically placed buttons, allowing users to seamlessly switch between operating modes and trigger specific functions without requiring visual input.

For powering the device, a HY 18650 lithium-ion battery system with a capacity of 7800 mAh and an output voltage of 3.7V was selected. This battery is integrated with a UPS module that includes a DC/DC boost converter to step up the voltage to 5V while incorporating protective circuitry. This configuration ensures stable operation and sufficient battery life for daily use.

All hardware components with a connections are illustrated in Figure 4. Hardware configuration of the proposed system, featuring the Raspberry Pi Zero 2W, a HY 18650 lithium-ion battery stepped up to 5V using a UPS 18650 module, an Arducam IMX519 camera connected via an FFC cable, a sound card interfaced via USB, and buttons connected through GPIO pins.

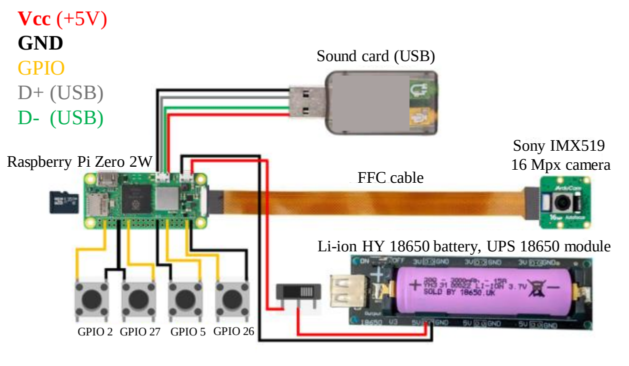

FIGURE 4. Hardware configuration.

B. SOFTWARE DESIGN

The development of the software for the assistive glasses was guided by a modular architecture, ensuring scalability, flexibility, and ease of updates. The software was primarily implemented in Python, leveraging its extensive libraries for computer vision, artificial intelligence (AI), and hardware interfacing. This section outlines the core components of the software design. The system operates through five primary modes: text reading, audio file playback, audio recording, remote assistance, and AI assistance.

The text reading mode employs OCR to extract text from images captured by the camera, which is then vocalized using a Text-to-Speech (TTS) engine. This functionality allows users to access printed information, such as street signs, product labels, and documents, with minimal effort.

The scene description function, available in AI assistance mode, provides real-time information about the user's surroundings. This feature utilizes computer vision algorithms to identify objects, people, and spatial layouts in the captured images, generating audio descriptions to enhance situational awareness.

Button mapping for text reading, audio file playback and audio recording mode is shown in Table I.

TABLE I
BUTTON MAPPING FOR TEXT READING, AUDIO FILE PLAYBACK AND AUDIO RECORDING MODES

Buttons	Text reading mode	Audio files mode	Audio record mode
Button 1 (GPIO2)	Read extracted text aloud	Play current audio.	Start/stop recording
Button 2 (GPIO27)	Rotate & read text	Play next audio.	N/A
Button 3 (GPIO5)	Translate text to Slovak	Change audio directory.	N/A
Button 4 (GPIO 26)	Step back/change mode (Applies to all modes)		

The AI assistant supports voice commands, enabling hands-free interaction with the system. It incorporates convolutional neural networks (CNNs), which are fundamental in image analysis tasks such as object recognition, image classification, and segmentation. The CNN architecture, characterized by convolutional layers applying filters to input data, allows the system to extract hierarchical features, capturing both low-level details (e.g., edges) and high-level concepts (e.g., object shapes).

The volunteer assistance mode enables users to connect with remote volunteers through a cloud-based server, allowing them to receive real-time guidance and support. By sharing images and audio messages, users can get personalized assistance suited to their specific needs. The system captures and transmits real-time images of the user's surroundings, helping volunteers provide relevant and accurate guidance based on the current environment. In addition to image sharing, the system supports two-way communication, where volunteers send text-based instructions that are automatically converted into audio output for the user. This ensures clear and effective assistance, especially in complex or unfamiliar situations where immediate support is needed.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3554366

The communication process between a visually impaired user and a volunteer assistant follows these sequential steps:

1. The system retrieves a list of available volunteers currently connected.
2. The user browses the list and selects a preferred volunteer for assistance.
3. A connection request is sent to the selected volunteer.
4. The volunteer accepts the request, confirming availability.
5. Upon acceptance:
   - A message interface is displayed to the volunteer.
   - The volunteer can send text-based messages or request additional images from the user.
   - The user can transmit voice messages (via the microphone) and real-time images (via the camera).
   - All incoming messages are converted into speech output, ensuring accessibility for the user.

Table II presents the button mappings for the remote assistance and AI modes. Since the GPIO assignments remain the same as those in Table I, they are not repeated here to avoid redundancy.

TABLE II
BUTTON MAPPING FOR REMOTE ASSISTANCE AND AI MODES

Buttons	Remote assistant selection	Remote assistance action	AI
Button 1	Switch to next volunteer.	Send an image of the environment to the volunteer	Describe scene
Button 2	Select current volunteer	Ask the volunteer a question	Activate AI voice assistant
Button 3	Load the list of available volunteers	N/A	N/A
Button 4	Step back/change mode (Applies to all modes)		

That table outlines how each button functions depending on whether the user is in remote assistant selection mode, remote assistance action mode, or AI mode.

The user interface for the assisting volunteer is presented in Figure 5, showing the overall system layout.

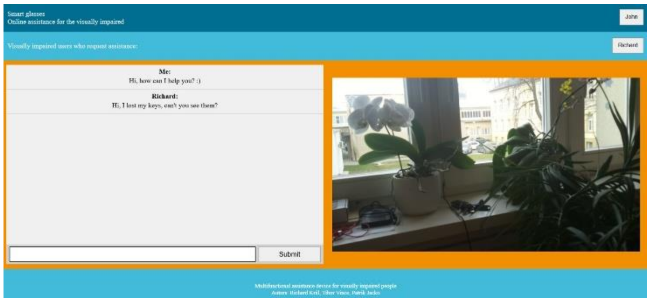

FIGURE 5. Volunteer UI overall layout, showing both the text communication and graphical components.

The interface is designed with two main components: a text-based communication system and a graphical interface. The text-based system allows volunteers to send messages that are converted into audio output for the user, ensuring clear and effective communication. Meanwhile, the graphical interface provides additional support by displaying real-time images captured by the user's device. Along with the visual feed, it also shows the names of both the volunteer and the user, helping volunteers provide more precise and context-aware guidance.

Figure 6 illustrates the text-based communication interface for volunteers. This section enables volunteers to exchange messages with visually impaired users. While only the volunteer sees the text, the user communicates using a microphone and speaker. The device automatically converts the user's voice into text, facilitating seamless interaction.

FIGURE 6. Text-based communication interface for message exchange between the assisting volunteer and the visually impaired user.

The graphical interface, shown in Figure 7, enhances the volunteer's ability to assist by displaying real-time images captured by the user's device. In addition to providing visual context, it presents the names of both the volunteer and the user in the upper right corner. This layout ensures that volunteers can deliver accurate and timely guidance based on the user's surroundings.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

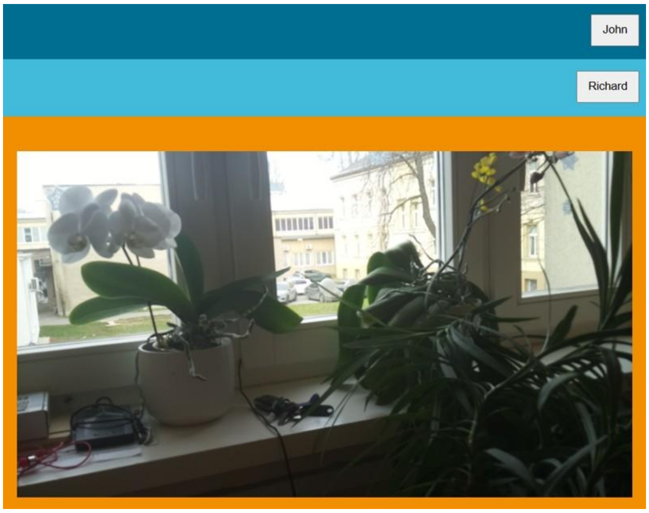

FIGURE 7. **Graphical interface displaying real-time images from the visually impaired user's environment including volunteer's and a client name.**

C. VISUAL DESIGN OF THE ASSISTIVE GLASSES

The visual design of the glasses was modeled using Autodesk Fusion 360, emphasizing modularity and user comfort. Initial prototypes were manufactured using Fused Deposition Modeling (FDM) technology. However, due to the complex geometry of the frames, the final iteration was produced using Selective Laser Sintering (SLS), which allows for higher precision and structural integrity.

The camera module is integrated into the frame of the glasses, providing a fixed viewpoint aligned with the user's field of vision. Tactile buttons are placed along the temple arm, ensuring easy accessibility and seamless operation. By carefully selecting these components, the system achieves an optimal balance of performance, functionality, and user comfort, ensuring feasibility as a practical assistive device for visually impaired users. The modular hardware architecture allows for cost-effective upgrades and repairs, ensuring long-term usability. Additionally, the open-source software framework enables customization based on user feedback, making it adaptable to diverse needs. The final visual design is shown in Figure 8.

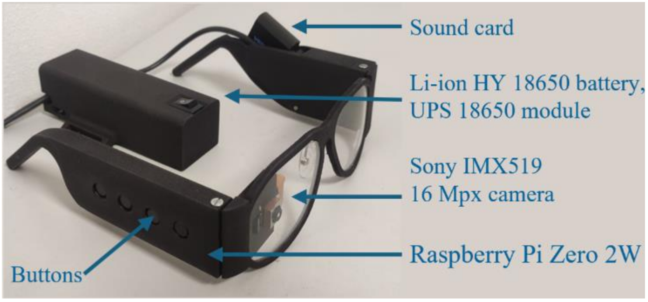

FIGURE 8. **Final design of the assistive glasses.**

The figure illustrates that the Raspberry Pi Zero 2W and control buttons are mounted on the right side, while the external sound card is positioned on the left. The camera is integrated into the frame, ensuring alignment with the user's field of vision. Due to weight constraints, the battery module is housed separately, optimizing both balance and comfort for the user.

V. USABILITY AND FUNCTIONAL ASSESMENT OF THE SYSTEM

The usability and functional assessment of the assistive glasses was conducted to evaluate their functionality, usability, and effectiveness in real-world scenarios. A combination of controlled laboratory experiments and practical field tests was employed to ensure a comprehensive evaluation of the device. The participant group consisted of five visually impaired individuals, all male, with varying ages and degrees of blindness, selected to reflect diverse user needs and capabilities.

The testing process began with an orientation session to familiarize participants with the device. During this session, users were introduced to the tactile button layout, audio feedback system, and various operational modes, including text reading, object detection, and scene description. The objective of this session was to minimize the learning curve and provide participants with confidence in independently operating the device during testing. The structured evaluation included a series of real-world tasks designed to simulate common challenges experienced by visually impaired individuals. Participants were asked to read printed text from documents and product labels, navigate an indoor environment with obstacles, and identify objects in a cluttered space. Additionally, they used the remote assistance mode to connect with volunteers for real-time guidance.

Each task was carefully selected to assess specific functionalities of the device. The evaluation focused on measuring the accuracy of the OCR module in recognizing and vocalizing text, analyzing the performance of the obstacle detection system in identifying hazards, and testing the reliability of the remote communication features in facilitating seamless interaction between users and volunteers.

In addition to technical performance, the usability of the device was evaluated based on user feedback. Participants rated their experience using the device based on ease of use, comfort, and interface intuitiveness. Additionally, qualitative feedback was collected through follow-up interviews, providing insights into user preferences and identifying potential areas for improvement. To ensure a rigorous evaluation, quantitative performance metrics were recorded throughout the testing process. The assessment focused on measuring the accuracy of text recognition, analyzing the response time of the obstacle detection system, and evaluating the efficiency of the remote assistance feature in facilitating seamless interaction. These metrics provided valuable insights into the overall performance and reliability of the assistive device.

The testing framework was designed to provide a complete analysis of both technical performance and user experience. By combining measurable data with direct user feedback, the study provided a clear understanding of the device's strengths and limitations. This approach ensured that the results would not only confirm the current design but also guide future improvements to better meet the needs of visually impaired users.

A. TEXT RECOGNITION PERFORMANCE

The performance of the assistive glasses was compared to a commercially available assistive device, the Envision Glasses [33], with a particular focus on text recognition accuracy. A well-known product in assistive technology was selected as a benchmark for comparison. This section outlines the key findings from the comparative analysis.

In text recognition tasks, the proposed assistive glasses demonstrated high accuracy when processing standard fonts such as Calibri and Times New Roman. The OCR module achieved an error rate of 9%, which was slightly higher than the 5% error rate recorded by the Envision Glasses. However, the performance gap increased when recognizing complex or artistic fonts, such as Freestyle Script, where the error rate for the proposed assistive glasses rose to 22.2%, compared to 8.94% for the Envision Glasses.

These results emphasize the need for further optimization of the OCR algorithm, particularly for non-standard text styles, to enhance the overall usability and effectiveness of the assistive system.

B. AI-BASED SCENE DESCRIPTION

The scene description functionality, powered by artificial intelligence (AI), was another key area of evaluation. Using the integrated camera, the system provided real-time audio descriptions of the environment, enabling users to develop a better understanding of their surroundings.

For example, when presented with a kitchen setup, the system accurately described the locations of objects, such as the sink, countertop items, and kitchen appliances. Participants reported that this feature significantly improved their spatial awareness, particularly in unfamiliar environments.

However, some users noted the need for improvements in object recognition speed and specificity, as certain descriptions were overly general or delayed in response.

C. REMOTE ASSISTANCE AND VOLUNTEER SUPPORT

The remote assistance mode of the prototype proved to be a valuable feature during usability tests. Participants successfully used the system to connect with remote volunteers, transmitting images and audio queries to receive real-time guidance. This feature was particularly beneficial in situations requiring detailed visual interpretation.

User feedback highlighted the need for faster image transmission and enhanced two-way communication options, such as voice-to-text conversion to facilitate more seamless interaction between users and volunteers.

D. USER FEEDBACK AND SYSTEM IMPROVEMENT RECOMMENDATIONS

Overall, participants were highly satisfied with the usability and design of the assistive glasses. The lightweight frame and intuitive button layout allowed for comfortable, prolonged use without causing significant fatigue. However, several areas for improvement were identified.

Participants reported challenges in navigating audio files and switching between modes, which occasionally led to confusion. Additionally, they expressed a need for enhanced navigation features, such as GPS-based outdoor guidance and real-time obstacle detection. Participants also suggested several refinements to enhance the user experience. They emphasized the need for a more natural-sounding voice in the audio output to improve speech clarity and reduce the robotic tone. Additionally, they recommended incorporating distance estimation for individual objects, enabling users to better gauge proximity in their surroundings. Another key suggestion was to provide more detailed descriptions of people, including clothing color, estimated age, or facial expressions, to enhance environmental awareness and interaction.

Experts from the rehabilitation center reinforced the potential of the prototype while identifying areas for further refinement. They recommended the integration of ultrasonic sensors to enhance real-time obstacle detection, improving safety and spatial awareness. Additionally, they emphasized the importance of incorporating GPS functionality to support outdoor navigation, enabling users to move independently in unfamiliar environments. Another key suggestion was the adoption of bone conduction headphones, which would allow users to receive auditory feedback without obstructing environmental sounds, ensuring better situational awareness. Experts also suggested creating a database of commonly used products and services, allowing users to retrieve detailed information by scanning barcodes or QR codes. Experts also proposed enhancements to the AI module to improve overall system performance. One key recommendation was the implementation of RH-VOICE speech synthesis, which would enhance the naturalness of the audio output, making spoken feedback clearer and more human-like. Additionally, they suggested expanding scene analysis capabilities to provide more detailed descriptions of people, including their appearance and distinguishing features, further improving the device's ability to assist users in recognizing individuals in various environments.

Despite its current limitations, the system was positively received as a promising advancement in assistive technology. Participants emphasized the importance of continued development, particularly in addressing navigation challenges and real-time information retrieval. The feedback from this evaluation has been instrumental in shaping the next iteration

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3554366

of the device, ensuring that future enhancements align closely with the needs and expectations of visually impaired users.

VI. DISCUSION

The findings from usability testing and functional assessments indicate both the potential and challenges associated with implementing such a system in real-world scenarios. Users identified several limitations and provided valuable feedback for future improvements.

One of the most promising aspects of the developed device is its ability to integrate multiple functionalities within a compact and lightweight form factor. The combination of real-time text recognition, obstacle detection, scene description, and remote volunteer assistance represents a comprehensive approach to assistive technology. By adopting a hands-free, wearable format, the system eliminates the need for users to hold or manually operate additional hardware, which can interfere with mobility and ease of use. This feature was particularly well-received by participants, reinforcing the importance of usability and ergonomics in assistive devices.

However, the assistive glasses form factor remains a design challenge, as the hardware required for the proposed improvements could result in a bulkier device, which some visually impaired users may find undesirable to wear.

A. PERFORMANCE OF CORE FEATURES

The real-time text recognition feature exhibited high accuracy with standard fonts but struggled with artistic or highly stylized fonts, suggesting the need for further optimization of the OCR algorithm. The delay in processing text highlights a trade-off between computational efficiency and recognition accuracy, which could be mitigated by integrating more advanced deep learning models trained on a broader range of typographic variations.

Similarly, the scene description function was beneficial in enhancing situational awareness, yet users reported delays in object identification and occasional overgeneralized scene descriptions. Participants also expressed a desire for enhanced functionality to improve overall usability. They emphasized the need for faster object recognition to provide quicker feedback in real-time scenarios. Additionally, they suggested incorporating more detailed descriptions of people, including clothing color, estimated age, and facial expressions, to improve identification and interaction. Another key improvement involved adding distance estimation for objects, allowing users to better assess spatial relationships and navigate their surroundings more effectively.

These insights suggest that improving processing speed and object specificity could significantly enhance the effectiveness of the device.

Obstacle detection, a critical feature for ensuring user safety, performed adequately in detecting common environmental hazards. However, limitations in speed and response time indicate the need for additional sensory inputs. Several potential improvements could enhance the functionality and usability of the assistive device. One approach involves integrating ultrasonic or LiDAR-based sensors to improve distance measurement, allowing for more precise detection of obstacles in the user's environment. Another enhancement focuses on implementing tactile feedback mechanisms, providing non-auditory guidance that would be particularly useful in noisy environments or situations where audio feedback is impractical. Additionally, adding GPS-based navigation support would significantly enhance outdoor usability, enabling users to navigate unfamiliar locations with greater confidence and independence.

Currently, the absence of GPS functionality restricts the device's utility in outdoor environments, where real-time positioning would significantly improve independent mobility.

B. REMOTE ASSISTANCE AND COMMUNICATION

The remote assistance mode was one of the most positively reviewed features, as it allowed users to connect with volunteers for real-time guidance and support. However, some participants reported that image transmission speed was slow, causing delays in receiving assistance. To address this issue, several optimizations could be explored. Improving data transmission protocols would help reduce latency and ensure faster communication between users and volunteers. Incorporating low-latency streaming solutions could further enhance image processing speed, allowing for quicker visual feedback. Additionally, introducing real-time voice interaction support instead of relying solely on text-based responses would create a more seamless and efficient communication experience, enabling users to receive immediate guidance in critical situations.

While real-time voice interaction would improve communication efficiency, it would also increase computational demands on the hardware, potentially leading to greater delays.

C. HARDWARE AND COST CONSIDERATIONS

Despite its promising capabilities, the system has several areas requiring refinement before it can be considered a fully viable assistive solution. The main limitation of the current prototype is its slow response time. A key design question is whether to process all functions locally, which would require more powerful hardware, or to offload processing to external servers, requiring faster communication protocols.

Another usability challenge identified was difficulty navigating between modes and managing audio outputs. To enhance accessibility, several solutions could be explored. Streamlining the interface with intuitive voice commands would allow users to interact with the device more efficiently, reducing the need for manual inputs. Additionally, incorporating gesture-based controls would enable hands-free navigation, offering a more seamless and user-friendly experience, particularly for individuals with limited dexterity or those who prefer non-verbal interaction methods.

Affordability remains a major concern, as many advanced assistive technologies are often priced beyond the financial reach of visually impaired individuals. While incorporating higher-end components or additional sensors could address many of the identified limitations, it would also lead to an increase in price, weight, size, and power consumption. To accommodate these changes, the device would either require a larger battery with a higher capacity, which would add extra weight, or operate with a shorter battery life, potentially limiting usability and daily functionality. Balancing performance with energy efficiency remains a key challenge in further optimizing the design.

The current version of the assistive glasses prototype was developed with a material cost of approximately $100 using commercially available components. Massive production would the price significantly decrease. In comparison, a commercial product such as Envision Glasses is available for approximately $3500.

D. ALTERNATIVE APPROACHES FOR ASSISTIVE TECHNOLOGY

Most referred research presents standalone assistive systems with dedicated AI models. This paper explores an alternative approach by considering thin-client-based assistive technologies.

With the rapid advancement of AI and 5G connectivity, offloading processing to remote AI models presents a cost-effective alternative. This approach would significantly reduce hardware requirements, leading to lower device costs while maintaining functionality. By leveraging powerful cloud-based AI models, it could also improve overall system performance, enabling faster and more accurate real-time processing. Additionally, this method would enhance accessibility, making assistive technology more affordable and widely available to individuals who may not have access to high-end hardware.

There are already AI models capable of processing continuous video inputs in real-time [34]. Developing dedicated AI modules for assistive technologies would further enhance the effectiveness and accessibility of such systems.

VII. FUTURE WORK

Future improvements to the proposed assistive device will focus on optimizing processing efficiency, feedback mechanisms, and usability. Enhancing real-time text recognition remains a priority, particularly in reducing processing delays and increasing accuracy with complex or artistic fonts. Future iterations will explore the integration of more advanced deep-learning models trained on a broader range of typographic variations.

Additionally, efforts will be directed toward alternative feedback mechanisms beyond audio output. The development of a haptic feedback vest will be investigated, allowing users to receive tactile notifications for navigation assistance, obstacle warnings, and direction cues. This alternative approach would improve accessibility for users in noisy environments or situations where audio feedback is impractical.

To enhance outdoor mobility, GPS-based navigation could be incorporated, offering real-time positioning and environmental awareness. The integration of ultrasonic or LiDAR-based sensors would be explored also to improve distance estimation and obstacle detection capabilities. These enhancements will enable users to navigate safely and independently in diverse environments.

Another key aspect of future work is hardware miniaturization and power efficiency. Reducing the overall weight and size of the device without compromising performance would ensure greater user comfort. Alternative form factors, such as a vest-mounted assistive system, will be considered as a potential solution for accommodating bulkier hardware components while maintaining ease of use.

Finally, the long-term goal of this research is to develop a cost-effective, cloud-based assistive technology model. By leveraging remote AI processing and high-speed networks, the device could offload computationally intensive tasks to external servers, reducing on-device hardware requirements and improving affordability. Future research could investigate the feasibility of real-time AI-driven video processing to enhance scene description and object recognition capabilities.

VIII. CONCLUSION

This study on the daily mobility activities of visually impaired individuals identified several key challenges. Among the participants surveyed, the most frequently reported difficulties included independent movement (90%), object recognition (40%), and reading text (20%). Traditional assistive tools, such as white canes and smartphone applications, were found to be helpful but insufficient in providing a comprehensive solution. Participants emphasized the need for an integrated device that offers real-time navigation, text recognition, and scene description while maintaining ease of use.

In response to these findings, a multifunctional wearable assistive device was developed in the form of smart glasses, designed to address the mobility and accessibility challenges faced by visually impaired individuals. The system integrates real-time text recognition, allowing users to read printed materials through an OCR module. It features obstacle detection, enhancing spatial awareness by identifying and alerting users to nearby objects. Additionally, the device includes remote assistance functionality, enabling users to connect with volunteers who can provide real-time guidance based on transmitted images and audio queries.

The device is built around a Raspberry Pi Zero 2W for processing, an Arducam IMX519 autofocus camera for capturing visual data, and a tactile button interface for simplified control. It provides real-time audio feedback, enhancing situational awareness and enabling users to navigate independently in various environments.

REFERENCES

[1] B. J. Lee and N. A. Afshari, “The global burden of blindness,” *Current Opinion in Ophthalmology*, vol. 36, no. 1. Ovid Technologies (Wolters Kluwer Health), pp. 1–3, Dec. 05, 2024. doi: 10.1097/icu.0000000000001099.

[2] M. Robby, E. Hartoyo, E. Wydiamala, H. Husaini, and S. Arifin, “Meta Analysis: The Relationship of Looking Distance and Long Time of Smartphone use with Myopia,” *Jurnal Berkala Kesehatan*, vol. 10, no. 2. Center for Journal Management and Publication, Lambung Mangkurat University, p. 157, Nov. 30, 2024. doi: 10.20527/jbk.v10i2.10947.

[3] J. Madake, S. Bhatlawande, A. Solanke, and S. Shilaskar, ‘‘A qualitative and quantitative analysis of research in mobility technologies for visually impaired people,’’ IEEE Access, vol. 11, pp. 82496–82520, 2023, doi: 10.1109/ACCESS.2023.3291074.

[4] F. Ashiq, M. Asif, M. B. Ahmad, S. Zafar, K. Masood, T. Mahmood, M. T. Mahmood, and I. H. Lee, ‘‘CNN-based object recognition and tracking system to assist visually impaired people,’’ IEEE Access, vol. 10, pp. 14819–14834, 2022.

[5] J. Zhang, K. Yang, A. Constantinescu, K. Peng, K. Müller, and R. Stiefel hagen, ‘‘Trans4Trans: Efficient transformer for transparent object and semantic scene segmentation in real-world navigation assistance,’’ IEEE Trans. Intell. Transp. Syst., vol. 23, no. 10, pp. 19173–19186, Oct. 2022.

[6] T. Vince, et al., “Android application supporting developed web applications testing,” in Proc. 2017 Int. Conf. Modern Electr. Energy Syst. (MEES), 2017, pp. 392–395.

[7] S. Busaeed, I. Katib, A. Albeshri, J. M. Corchado, T. Yigitcanlar, and R. Mehmood, ‘‘LidSonic V2.0: A LiDAR and deep-learning-based green assistive edge device to enhance mobility for the visually impaired,’’ Sensors, vol. 22, no. 19, p. 7435, Sep. 2022.

[8] A. Lo Valvo, D. Croce, D. Garlisi, F. Giuliano, L. Giarré, and I. Tinnirello, “A Navigation and Augmented Reality System for Visually Impaired People,” Sensors, vol. 21, no. 9. MDPI AG, p. 3061, Apr. 28, 2021. doi: 10.3390/s21093061.

[9] S. Bhatlawande, R. Borse, A. Solanke and S. Shilaskar, "A Smart Clothing Approach for Augmenting Mobility of Visually Impaired People," in *IEEE Access*, vol. 12, pp. 24659-24671, 2024, doi: 10.1109/ACCESS.2024.3364915.

[10] A. I. Apu, A.-A. Nayan, J. Ferdaous, and M. G. Kibria, “IoT-Based Smart Blind Stick,” Lecture Notes on Data Engineering and Communications Technologies. Springer Singapore, pp. 447–460, Dec. 04, 2021. doi: 10.1007/978-981-16-6636-0_34.

[11] Y.-L. Poy, S. Darmaraju, C.-H. Goh, and B.-H. Kwan, “Standalone Smart Glass System for the Blind and Visually Impaired,” 2024 IEEE 14th Symposium on Computer Applications & Industrial Electronics (ISCAIE). IEEE, pp. 239–244, May 24, 2024. doi: 10.1109/iscaie61308.2024.10576410.

[12] A. A. A. Alshawi, J. Tanha and M. A. Balafar, "An Attention-Based Convolutional Recurrent Neural Networks for Scene Text Recognition," in *IEEE Access*, vol. 12, pp. 8123-8134, 2024, doi: 10.1109/ACCESS.2024.3352748.

[13] Z. Li, B. Li, S. G. Jahng and C. Jung, "Improved VGG Algorithm for Visual Prosthesis Image Recognition," in *IEEE Access*, vol. 12, pp. 45727-45739, 2024, doi: 10.1109/ACCESS.2024.3380839.

[14] Z. Liu, Y. Li, F. Ren, W. L. Goh, and H. Yu, ‘‘SqueezedText: A real-time scene text recognition by binary convolutional encoder–decoder network,’’in Proc. 32ndAAAIConf.Artif. Intell., 2018, vol. 32, no. 1, pp. 7194–7201, doi: 10.1609/aaai.v32i1.12252.

[15] Y. Gao, Z. Huang, Y. Dai, C. Xu, K. Chen, and J. Guo, ‘‘DSAN: Double supervised network with attention mechanism for scene text recognition,’’in Proc. IEEE Vis. Commun. Image Process. (VCIP), Dec. 2019, pp. 1–4, doi: 10.1109/VCIP47243.2019.8965779.

[16] A. Perez-Yus *et al*., "RASPV: A Robotics Framework for Augmented Simulated Prosthetic Vision," in *IEEE Access*, vol. 12, pp. 15251-15267, 2024, doi: 10.1109/ACCESS.2024.3357400.

[17] M. Bamdad, D. Scaramuzza, and A. Darvishy, “SLAM for Visually Impaired People: a Survey,” 2022, arXiv. doi: 10.48550/ARXIV.2212.04745.

[18] M. Kim, Y. Park, K. Moon and C. Y. Jeong, "Impact of Device and Environment on Visual-Auditory Sensory Substitution: A Comprehensive Behavioral Analysis Using the vOICe Algorithm," in IEEE Access, vol. 12, pp. 90501-90510, 2024, doi: 10.1109/ACCESS.2024.3419102.

[19] Z. Yu and M. Hu, "Real Environment Warning Model for Visually Impaired People in Trouble on the Blind Roads Based on Wavelet Scattering Network," in *IEEE Access*, vol. 12, pp. 82156-82167, 2024, doi: 10.1109/ACCESS.2024.3412328.

[20] M. Alghamdi, H. A. Mengash, M. Aljebreen, M. Maray, A. A. Darem and A. S. Salama, "Empowering Retail Through Advanced Consumer Product Recognition Using Aquila Optimization Algorithm With Deep Learning," in *IEEE Access*, vol. 12, pp. 71055-71065, 2024, doi: 10.1109/ACCESS.2024.3399480.

[21] K. Minatani, “Proposal for a Versatile Smart White Cane Infrastructure Using a 3D Printer: As a R&D Platform and a Practical Product Design,” Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive

[22] Environments. ACM, pp. 21–26, Jun. 26, 2024. doi: 10.1145/3652037.3652064.

[23] Z. Cheng, Y. Xu, F. Bai, Y. Niu, S. Pu, and S. Zhou, ‘‘AON: Towards arbitrarily-oriented text recognition,’’ 2017, arXiv:1711.04226.

[24] D. Zhong, S. Lyu, P. Shivakumara, B. Yin, J. Wu, U. Pal, and Y. Lu, ‘‘SGBANet: Semantic GAN and balanced attention network for arbitrarily oriented scene text recognition,’’ in Proc. Eur. Conf. Comput. Vis., 2022, pp. 464–480.

[25] W. Yu, M. Ibrayim, and A. Hamdulla, ‘‘Scene text recognition based on improved CRNN,’’ Information, vol. 14, no. 7, p. 369, Jun. 2023.

[26] Q.-H. Nguyen et al., “A Visual SLAM System on Mobile Robot Supporting Localization Services to Visually Impaired People,” Lecture Notes in Computer Science. Springer International Publishing, pp. 716–729, 2015. doi: 10.1007/978-3-319-16199-0_50.

[27] M. Yamagishi, M. Okamura and Y. Fujisawa, "Development of English-compatible Braille learning device," *2024 IEEE 13th Global Conference on Consumer Electronics (GCCE)*, Kitakyushu, Japan, 2024, pp. 689-690, doi: 10.1109/GCCE62371.2024.10760949.

[28] S. K. Ghosh, E. Valveny, and A. D. Bagdanov, ‘‘Visual attention models for scene text recognition,’’ in Proc. 14th IAPR Int. Conf. Document Anal. Recognit. (ICDAR), vol. 1, Nov. 2017, pp. 943–948, doi:10.1109/ICDAR.2017.158.

[29] Y. A. Nanehkaran, J. Chen, S. Salimi, and D. Zhang, ‘‘A pragmatic convolutional bagging ensemble learning for recognition of Farsi handwritten digits,’’ J. Supercomput., vol. 77, no. 11, pp. 13474–13493, Nov. 2021.

[30] Y. A. Nanehkaran, D. Zhang, S. Salimi, J. Chen, Y. Tian, and N. Al-Nabhan, ‘‘Analysis and comparison of machine learning classifiers and deep neural networks techniques for recognition of Farsi handwritten digits,’’ J. Supercomput., vol. 77, no. 4, pp. 3193–3222, Apr. 2021.

[31] B. Su and S. Lu, ‘‘Accurate scene text recognition based on recurrent neural network,’’ in Proc. Asian Conf. Comput. Vis., in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 9003, 2015, pp. 35–48, doi: 10.1007/978-3-319-16865-4_3.

[32] S. Bai, H. Tang, and S. An, ‘‘Coordinate CNNs and LSTMs to categorize scene images with multi-views and multi-levels of abstraction,’’ Expert Syst. Appl., vol. 120, pp. 298–309, Apr. 2019, doi:10.1016/j.eswa.2018.08.056

[33] CNIB Smartlife, “Envision Glasses Lux Frames - Boxed SP,” CNIB Smartlife, [Online]. Available: https://cnibsmartlife.ca/products/envision-glasses-lux-frames-boxed-sp. [Accessed: Feb. 4, 2024].

[34] CHEN, Joya, et al. VideoLLM-online: Online Video Large Language Model for Streaming Video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. p. 18407-18418.

RICHARD KRÁL finished his M.Eng. degree study in 2024 at the Technical University of Košice, at the Department of Theoretical and Industrial Electrical Engineering of Faculty of Electrical Engineering and Informatics. He is currently pursuing his Ph.D. at the same department, focusing on the research and development of assistive devices for visually impaired people. The title of his dissertation is "Design and Implementation of a Device for Supporting Visually Impaired Individuals with a Focus on Spatial Orientation." His previous research includes the development of a multifunctional assistive device for visually impaired individuals as part of his master's thesis and a text scanner and reader for visually impaired users in his bachelor's thesis.

PATRIK JACKO finished his PhD. degree study in 2020 at the Technical University of Košice, at the Department of Theoretical and Industrial Electrical Engineering of Faculty of Electrical Engineering and Informatics. The title of his thesis was Application smart DSP processors in industrial electrical engineering. Before the doctoral study, he achieved bachelor and master degree at the same department, which he finished in 2016. The title of his master thesis was Design and implementation of a stepper motor control system based on Arduino. Currently, he is an assistant professor at the Department of Theoretical and Industrial Electrical Engineering, where is primary focused to high performance microcontrollers applied in digital signal processing and electromagnetism.

TIBOR VINCE received his M.Eng. degree in Industrial Engineering from the Technical University of Košice in 2003. He earned his Ph.D. in Theoretical Electrical Engineering from the same university in 2011. Since 2008, he has been affiliated with the Technical University of Košice, initially serving as an assistant professor in the Department of Theoretical and Industrial Electrical Engineering. In 2017, he was promoted to associate professor in the same department. He is the author of eight patents and serves as a scientific board member and reviewer for several international journals. His research focuses on the Internet of Things (IoT), Machine-to-Machine (M2M) communication, Internet communication, and server systems.