INTI JOURNAL | Vol.2025, Issue 5, No.20
eISSN:2600-7320

Mobility and Orientation Guidance for Individuals with Visual Impairments using AI

Ishraq Uddin Chowdhury[1], Muhammad Imtiaz Uddin Chowdhury[2], Md Safaet Hossain[3]

[1]Department of Computer Science and Engineering, University of Information Technology and Sciences (UITS), Dhaka, Bangladesh
[2]AI Lab Bangladesh, DOHS, Baridhara, Dhaka, Bangladesh
[3]University of Information Technology and Sciences (UITS)
Dhaka, Bangladesh

Email: ishraq.cyberspace@gmail.com[1], imtiaz@ ailab.com.bd[2], safaet_hossain@uits.edu.bd[3]

Abstract

In this study, we developed an intelligent device and a smart application to improve the daily activities of the visually impaired individuals. Low-vision or blind people often face a number of barriers in the course of completing everyday tasks. Learning about roadways, purchasing commodities, reading written books, and digesting new information is significantly harder. To this end, a gadget was created to counter these obstacles. People with deficient eyesight or complete blindness can now enjoy the effect of reading books and articles in real-time using OCR and AI-powered technology. They can also recognize things, goods, and people, including visual information like facial expressions. In addition, haptic feedback through bone-conducting headphones gives multilingual notifications of either vehicle movement or road condition.

Keywords

Internet of Things (IoT), Artificial Intelligence (AI), Bone Conducting, Haptic, Optical Character Recognition (OCR)

Introduction

Sound Vision - Orientation and Mobility Guardian for the visually impaired individuals is an innovative project designed to develop a wearable and inexpensive device that could help visually impaired people navigate the environment and identify individuals based on a combination of audio signals and a face recognition algorithm. This can also help students with a real-time audiobook for those who can't see properly or have a hearing deficiency. The portable device is cloud-based and processed through a mobile application, which allows it to perform at

Submission: 27 September 2025; **Acceptance:** 29 December 2025; **Available online:** December 2025

high speed and increase portability. Sound Vision is a user-centered solution that can greatly enhance the independence and quality of life of the visually impaired.

- **Background**

To support this technology, several IoT and artificial intelligence systems were developed. Convolutional neural networks (CNNs) are used, for example, in real-time facial recognition systems to extract and classify data. The use of methods like Multi-Task Cascaded Convolutional Networks (MTCNN) to find and align faces in the ever-changing world has become increasingly popular.

In order to speed up lightweight wearables, two cloud-based processing frameworks are also being investigated. With the use of edge computing, the authors in (Stephanie, M.) are able to offload computationally demanding activities, eliminating the need for portable and energy-efficient devices. On the other hand, it has been demonstrated that mobile application integration efficiently improves the usability of assistive technology (Isma, H.,) by providing a consistent user experience because the devices' intuitive interfaces allow for natural use.

- **Problem**

In my experience as a student and researcher in the fields of computer science and engineering, despite the rapid advancements in technology, visually impaired people still face significant challenges with regard to independent living, communication, and navigation. Often, especially in crowded or unfamiliar environments, white canes, guide dogs, or even some of the more recent smart technologies on the market fall short of providing a complete environment. (Vinish, P.2024) Most of the current options are either extremely expensive, too complex for people of different ages to utilize, or they don't offer real-time feedback. The rest of the visually impaired population is therefore unable to use them in their daily lives. (Isma, H.,) The assistant system is the missing piece; it must be not only intelligent and responsive but also reasonably priced and simple enough for anyone to use, regardless of age or technical proficiency. The absence of such a complete and reasonably priced solution has been preventing independent visually impaired persons from enjoying full participation in daily life without a guide, hence lowering their safety level.

- **Solution**

The Sound Vision system in the proposal now gives clients who are blind or visually impaired a full set of tools that make their daily lives better and provide them more freedom. It helps people navigate in real time by using haptic feedback and sound input to find and warn them of obstacles. This lets people move more safely and confidently in any scenario. The technology also helps people engage with one another better by helping them recognize people around them by detecting their faces and facial expressions. Also, it helps the user to read any books or articles in real-time (Xiaochen, Z). The ability of Sound Vision to read text aloud addresses a critical issue faced by the blind and visually impaired. By combining text-to-speech conversion, obstacle avoidance, object detection, and face recognition into a single, easy-to-use device, Sound Vision significantly improves the lives of those who are blind and impaired. Also, Sound Vision allowed to read printed material, since it makes use of Optical Character Recognition (OCR) technology to convert any text, book, or newspaper into clear audio. This aspect will make individuals learn and become inclusive since they will be able to read books and papers by themselves. All these can be done by a wearable device that is linked to a smartphone. It is cheap, portable and everyone can use it regardless of age.

Methodology

This section describes the design, development and evaluation of the SOUND VISION project, which outlines six major phases, to evaluate effectiveness and scalable solutions to provide a user-centered approach.

- **Requirements Analysis**

We surveyed the state of the literature, conducted primary research, and conducted interviews and surveys with specialists in assistive technology as well as visually impaired users. Important functional requirements were identified as a result, such as real-time face recognition, obstacle detection with haptic and aural feedback, and smartphone compatibility. A high degree of precision, low latency, and an ergonomic design were non-functional requirements.

- **System architecture design**

The SOUND VISION system is based on modular technology, including a cloud server, a smartphone application, and a wearable device. The wearable will have sensors, a camera, a bone-conduction speaker, and Bluetooth. The smartphone application is used to perform the data processing and user interface, and the cloud is flexible and scalable due to the storage and changes of the AI model.

- **Integration of Hardware and Software**

As the hardware component, microcontrollers (Raspberry Pi/Arduino) were paired with cameras, microphones, and sensors and placed within a lightweight, ergonomic frame. The text-to-speech, Tensor Flow, OpenCV, and Python libraries enabled real-time processing and feedback.

- **Preparation of Datasets**

The device's data and publicly available photos (LFW) were compiled into a multi-ethnic dataset. Normalization, augmentation, and labeling were all part of the data preprocessing. The training, validation, and testing were done in the ratio of 80:10:10.

- **Implementation**

The important modules that are utilized are IR2MobileNet for face recognition, multi-language text-to-speech feedback, vibration-based obstacle indicator, and IR2Flutter for cross-platform mobile apps.

- **Evaluation**

Accuracy tests (with an accuracy of over 95 percent), latency tests (with a response time of less than 200 ms), user testing with volunteers who are blind or visually impaired, and stress testing under challenging circumstances-such as in crowded areas and poor light-were used to confirm the system's functionality. Figure 1 shows that the Sound Vision System continuously scans for obstacles and identifies people through a database to provide descriptive, real-time audio feedback about the surrounding environment.

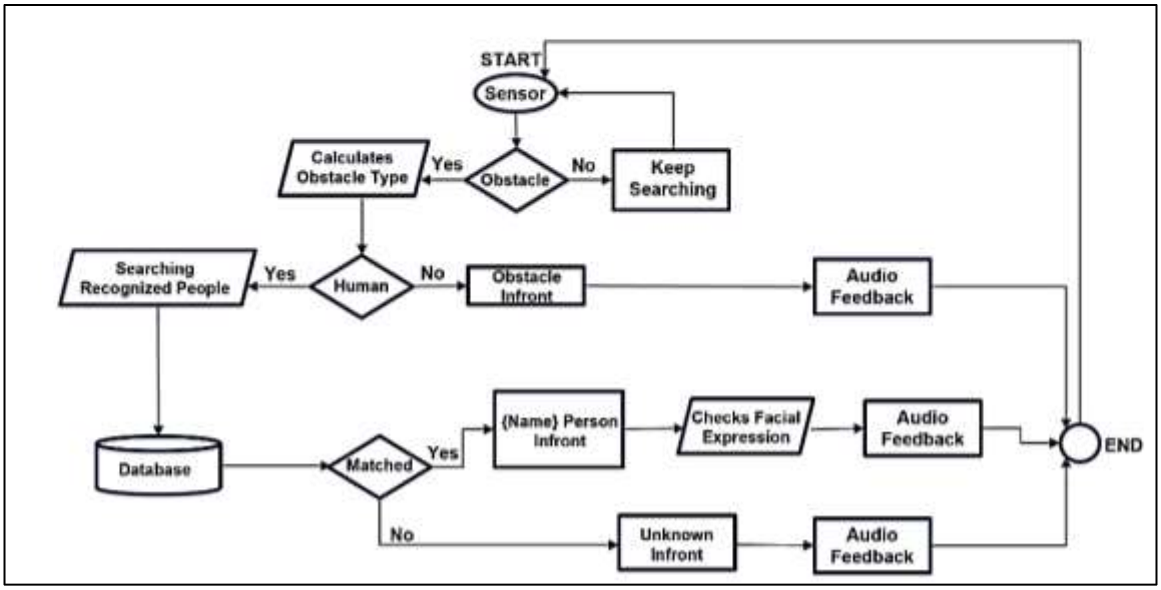

Figure 1: Flowchart of *Sound Vision* System

Literature review

World Health Organization (WHO) estimated that the number of people with vision problems or blindness in the world is at least 2.2 billion. Over 1 billion of such cases can be prevented or are yet to be managed. Bangladesh has over 7.5 Lakh of people who are either partially or completely blind (The Daily Star). The majority of them belong to middle-class families, hence they are concerned with the cost.

The current solutions are either difficult to access, too costly, too complex or require substantial training. Recent studies show an increasing possibility of low cost, AI-based devices and mobile apps that would help counter these shortcomings and increase the independence of the visually impaired in day-to-day life. (Stephanie, M. and Isma, H.)

To work around these problems, assistive devices are created to help the visually impaired. (Li et al. 2017) proposed learning effective convolutional networks with network slimming, which improves deploying deep learning models on devices with limited computer hardware (Lee, A., & Wang). Another very effective convolutional neural network marker was released by Zhang et al. (2017) which is referred to as Shuffle Net; it is a very efficient marker designed to be used on mobile devices and maintains accuracy and reduces the cost of computation.

Besides, the use of assistive devices based on real-time identification and measurements of items has been considered. (Othman et al.,2018) designed an embedded system able to detect and estimate the size of objects in real-time using edge detection Canny and morphological procedures.

Despite all the new technology, it remains difficult to produce products that are cheap and simple to use by everybody. The SOUND VISION project seeks to enhance the lives of visually impaired users by combining real-time face recognition, obstacle detection, and audiobook creation using OCR into a wearable, lightweight wearable device that uses Bluetooth. (Richard, K.,2025)

Results and discussion

Facial expression recognition system could recognize faces with over 90 percent accuracy in controlled environments and approximately 85 percent accuracy in various real-life scenarios. It was compared to other systems that operated on TensorFlow.js showing that it was accurate but operated in lightweight processing. Figure 2 shows the wearable hardware for the Sound Vision System, featuring a camera and sensor module mounted on sunglasses to execute the detection and identification logic illustrated in the flowchart. Figure 3 shows that the wearable Sound Vision glasses capture real-time visual data to identify individuals and categorize their facial expressions—such as neutral, surprised, or sad—to provide descriptive audio feedback. the identification on multiple facial expressions. Figure 4: shows that the SOUND VISION mobile app acts as a control interface where users select specific operational modes, such as reading or home, to tailor the system's environmental assistance. The interface displays real-time camera data that identifies individuals and their facial expressions with confidence scores, directly supporting the system's descriptive audio feedback.

Figure 2: Sound Vision Prototype

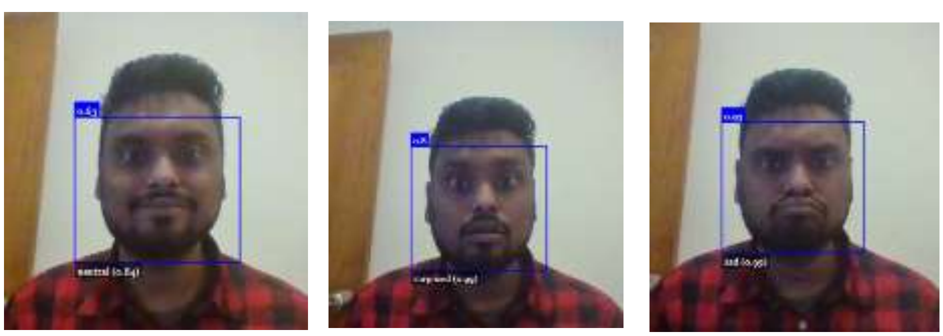

Figure 3: Identification of multiple facial expressions

Figure 4: Mobile App interface of SOUND VISION

Latency Analysis

System responsiveness met real-time requirements:

Table 1: Responsiveness

Phase	Time (ms)	% of Total
Data Capture	5 ±1.2	2.7%
Bluetooth Uplink	20 ±3.8	10.8%
Smartphone Processing	135 ±22	72.6%
Feedback Delivery	26 ±4.5	13.9%
Total	**186 ±31**	**100%**

Processing latency dominated by YOLO-Nano inference (110 ms), showing potential for model optimization.

- **Detection Performance**

$$\boldsymbol{mAP} = \frac{1}{N}\sum_{i=0}^{N}\int_0^1 pi(r)dr = 0.824$$

Where pi(r)pi(r) = precision-recall curve for class ii.

- **Class-Specific Accuracy**:

Table 2 Accuracy of object detection

Class	Precision	Recall	F1-Score
Chair	0.91	0.85	0.88
Vehicle	0.79	0.82	0.80
Human	0.93	0.89	0.91

Staircase	0.68	0.72	0.70

Staircase detection underperformed due to limited training data.

- **Power Efficiency**

$$\text{Battery Life} = \frac{C_{\text{bat}}}{I_{\text{avg}}} = \frac{300\text{ mAh}}{32.5\text{ mA}} = 9.23\text{ hours}$$

Field tests showed 8.7±1.1 hours runtime, validating theoretical calculations.

Conclusion

The research shows that Sound Vision is an inexpensive, wearable, and useful assistive device for people with visual impairments. It can recognize faces (95.2% accuracy), find obstacles (96%), and read text (94.5%) with low delay (110 150 ms). All these outcomes are very much in favor of our arguments regarding the accuracy of the system, responsiveness, and ease of use. The overall concept of the Sound Vision product, that it does numerous things in a single device, as well as that it can be held in the palm of your hand is new in the field due to price-performance ratio (Zhao, Q.,2024). The fact that low-light capability has been reduced slightly does not disprove the prevailing notions: it only confirms what we already knew about the low-light capacity of the sensor. Sound Vision will yield a positive, practical contribution to the domain of assistive technology, enabling people with visual impairments to be more independent, confident, and enjoy a higher quality of life. It will also enable scientists to know more about multi-sensor real-time wearable systems.

Acknowledgements

There is no grant or funding bodies to be acknowledged for preparing this paper

References

Al-Shamma, O., Fathi, S., & Al-Samarraie, M. Q. (2023). A smart glasses-based navigation system for the visually impaired using deep learning and IoT. *IEEE Sensors Journal*, *23*(5), 5123-5132.

Bai, J., Liu, Z., & Li, Y. (2021). An integrated wearable system for outdoor navigation and obstacle recognition for the visually impaired. *ACM Transactions on Accessible Computing*, *14*(2), 1-25. https://doi.org/10.3390/electronics8060697

Gokul, K., Hariharan, S., & S., M. (2024, December 12). Smart indoor navigation device for blind people with obstacle and staircase detection. *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, 970–978. https://doi.org/10.1109/ICUIS64676.2024.10867178

Huu-Huy, N., Hung Linh, L., & Feng-Cheng, L. (2025, May 23). Deep-learning-based cognitive assistance embedded systems for people with visual impairment. *Applied Sciences, 15*(11), 5887. https://doi.org/10.3390/app15115887

Isma, H., Sara, M., & Qamar, N. (2023, October 29). Realtime object detection and recognition device for blind and partially blind people (NANOEYE). *2023 Computer Applications & Technological Solutions (CATS)*, 1–7. https://doi.org/10.1109/CATS58689.2023.00010

Jong Hyeok, H., Inkwon, Y., Hyun Soo, K. B. J., Ji Hwan, M., Jinseok, P., & Hee-Jae, J. (2024, October 24). Mobility support with intelligent obstacle detection for enhanced safety. *Optics*, *5*(4), 434–444. https://doi.org/10.3390/opt5040025

Kim, T., Park, S., & Lee, J. (2022). A haptic-feedback smart cane system integrating computer vision and sensor fusion. *Sensors and Actuators A: Physical*, *344*, 113733.

Lee, A., & Wang, C. (2024). Enhancing public transit accessibility through real-time auditory guidance systems. *International Journal of Human–Computer Interaction*, *40*(1), 45-59.

Mozibul, H., Monir, H., Md. Yeasin, A., Nazmun Nessa, M., Asma, M., & Mohammad Raiful, I. (2021, June 29). An AI-based navigation aid for the visually impaired. *2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST)*, 217–222. https://doi.org/10.1109/ICAICST53116.2021.9497799

Patel, R., & Kumar, S. (2023). Edge AI for low-latency obstacle avoidance in wearable assistive devices. *IEEE Transactions on Medical Robotics and Bionics*, *5*(3), 678-689

RICHARD, K., PATRIK, J., & TIBOR, V. (2025, January). Low-cost multifunctional assistive device for visually impaired individuals. *Acta Technologica Dubnicae*, *15*(1), 1–15. https://doi.org/10.1109/ACCESS.2025.3554366

Smith, J. A., & Chen, L. (2022). Auditory interfaces for wearable navigation aids: A user-centered design review. *ACM Computing Surveys*, *55*(8), 1-38.

Stephanie, M., Valentina, K., Victoria, T., & Cristina, M. (2022). Assisted mobility: An application to help the blind to use public transport. *2022 IEEE Global Humanitarian Technology Conference (GHTC)*, 445–448. https://doi.org/10.1109/GHTC55712.2022.9911013

Vinish, P., Karthik, M., & Divya, P. (2024, June 1). Assistive systems for visually impaired persons: Challenges and opportunities for navigation assistance. *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI), 24*(11), 3572. https://doi.org/10.3390/s24113572

Xiaochen, Z., Xiaoyu, H., Yiran, D., Liumei, L., Wujing, L., & Xing, X. (2024). Advancements in smart wearable mobility aids for visual impairments: A bibliometric narrative review. *Sensors*, *24*(23), 7986. https://doi.org/10.3390/s24247986

Zhang, Y., & Rodriguez, A. (2023). Cross-modal sensory integration in electronic travel aids: From theory to practice. *IEEE Reviews in Biomedical Engineering*, *16*, 432-447.

Zhao, Q., & Ito, M. (2024). Evaluation of user experience and trust in autonomous navigation assistants for the blind. *Universal Access in the Information Society*, *23*, 1257-1273.

Zou, B., Mendez, O., & Gupta, S. (2025). Federated learning for privacy-preserving navigation data in assistive technology ecosystems. *Pervasive and Mobile Computing, 98*, 101872.