

An Auditory System Interface for Augmented Accessibility: Empowering the Visually Impaired [†]

Lalit Kumar, Ajay Shriram Kushwaha *  and Agrim Jain

Department of CSA, Sharda School of Engineering & Technology, Sharda University,
Greater Noida 201310, Uttar Pradesh, India; lalit.kumar1@sharda.ac.in (L.K.);
2020002047.agrim@ug.sharda.ac.in (A.J.)

* Correspondence: ajay.kushwaha@sharda.ac.in

[†] Presented at the 2nd Computing Congress 2023, Chennai, India, 28–29 December 2023.

Abstract: In 2023, global data revealed that approximately 2.2 billion people are affected by some form of vision impairment. By addressing the specific challenges faced by individuals who are blind and visually impaired, this research paper introduces a novel system that integrates Google Speech input and text-to-speech technology. This innovative approach enables users to perform a variety of tasks, such as reading, detecting weather conditions, and finding locations, using simple voice commands. The user-friendly application is designed to significantly improve social interaction and daily activities for individuals who are visually impaired, underscoring the critical role of accessibility in technological advancements. This research effectively demonstrates the potential of this system to enhance the quality of life for those with visual impairments.

Keywords: auditory interface; visual impairment; text to speech; optical character recognition (OCR); assistive technology



Citation: Kumar, L.; Kushwaha, A.S.; Jain, A. An Auditory System Interface for Augmented Accessibility: Empowering the Visually Impaired. *Eng. Proc.* **2024**, *62*, 8. <https://doi.org/10.3390/engproc2024062008>

Academic Editors: Geetha Ganesan, Xiaochun Cheng and Valentina Emilia Balas

Published: 5 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The daily challenges that people who are blind and visually impaired encounter, such as reading, determining their present location, detecting the weather, determining their phone battery condition, and determining the time and date, were taken into account when developing this project. In order to analyze such tasks, we used Google Speech input, which requires the blind user to utter certain phrases. The user must swipe either right or left on the screen to activate the voice assistant and speak in this straightforward application. A text-to-speech option is also provided so that the user can hear how the system works and learn how to use it. It was developed to make social interactions simpler for people who are deaf and blind. It gives the deaf/blind users the ability to carry out some simple everyday tasks, like reading, using a calculator, checking the weather, finding their position, and checking their phone's battery, with only a few touches and taps.

In Figure 1, the data reveal a consistent increment in the percentage of blind individuals from 2010 to 2023, showing an average of around 0.03% annually, with a recent decline to 0.01% in 2022 and 2023, indicating a positive global trend in reducing the burden of blindness [1]. Nonetheless, the distribution of blindness is unequal globally, with the majority in low- and middle-income countries lacking adequate eye care services, leading to higher increments than the global average. The World Health Organization (WHO) is actively addressing this issue through its Action Plan, aiming to diminish avoidable blindness and visual impairment, and offering a roadmap for countries to mitigate this disparity.

The percentages given are based on estimates made by the World Health Organization (WHO) for the year 2020 [2]. These statistics, which are presented in Table 1, provide information on the prevalence of blindness in various geographical areas of the world and highlight the substantial worldwide burden of visual impairment. These estimates

provide a useful frame of reference for comprehending the global impact of blindness and the requirement for measures to address this issue.

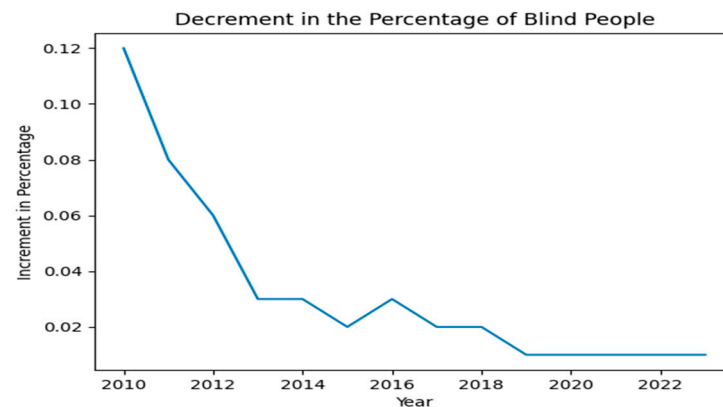


Figure 1. Percentage decrement of blind people in the world.

Table 1. Features of the works in the literature.

Region	Percentage of Blind People
World	0.62%
Africa	0.95%
Asia	0.57%
Europe	0.41%
Latin America and the Caribbean	0.40%
North America	0.26%
Oceania	0.35%

This initiative seeks to address the daily obstacles faced by individuals with visual impairments by offering practical solutions for tasks such as reading, determining location, checking weather conditions, monitoring phone battery levels, and obtaining the current time and date. The integration of Google Speech input allows users to activate specific functionalities effortlessly through simple vocal commands. User interaction with the system involves swiping right or left on the screen to initiate communication with the voice assistant. The incorporation of a text-to-speech technique enhances the overall user experience, providing individuals who are both deaf and blind with auditory cues to understand the system's functionality. The proposed system emphasizes touch and screen interactions, enabling users to perform essential daily activities, including reading, using a calculator, checking weather forecasts, determining location, and monitoring time, date, and phone battery levels. A noteworthy feature of the system lies in its responsiveness to voice commands, streamlining user engagement. For instance, uttering the command "Read" promptly initiates the corresponding activity. This research aims to improve communication and accessibility for individuals with visual and auditory impairments by offering a comprehensive and user-friendly interface for performing essential tasks through a combination of touch and voice commands.

2. Background and Related Works

People who are blind may travel more freely and comfortably thanks to the smart walker stick. The sensor is used to detect obstacles in regular canes and sticks. However, it is ineffective for those who are visually challenged. Due to the fact that a blind person cannot tell what kinds of things or objects are in front of them, a blind individual is unable to determine the size of an object or their distance from it [3].

People who are visually challenged still have a lot of difficulties navigating their surroundings safely. Along with their regular directing assistance, they also significantly rely on speech-based GPS. However, veering problems, which hinder the ability of blind people to keep a straight course, are not resolved by GPS-based devices. Some research systems offer input meant to rectify veering, although they frequently make use of large, specialized hardware [2]. Many people currently have compromised senses, such as foot and eye impairments. Blindness is one of the many impairments that are so often neglected in our culture. Blind people are people who have lost their ability to see [4].

Designing a navigation system for individuals who are blind poses significant challenges. Those with visual impairments face numerous difficulties in their daily lives, with navigating unfamiliar environments being one of the most daunting tasks. Unlike sighted individuals, who rely on their vision to orient themselves, people with visual impairments lack this crucial sense and must find alternative means of familiarizing themselves with their surroundings [5].

A voice-based email system is suggested in a paper [6] to improve accessibility for blind users. The proposed architecture addresses the current deficiency of audio feedback in blind user technology by incorporating speech recognition, interactive voice response, and mouse click events. During user authentication, voice recognition offers an additional degree of protection. The mailbox, login, and registration modules make up the system. While the login module uses speech commands, including a voice sample, for authentication, the registration module gathers user data. Once they log in, users can use voice commands to navigate the mailbox and carry out operations including writing, evaluating sent mail, monitoring the inbox, and taking care of the trash.

A unique email system is presented in reference [7], which departs from conventional approaches by including voice instructions. The system uses speech-to-text technology, asking users to give orders to perform different tasks. It has a main activity screen with a full-sized button for beginning tasks, and it makes use of IMAP (Internet Message Access Protocol). The system is more accessible and user-friendly—especially for those who have visual impairments—because users may use voice commands to carry out tasks like writing emails and viewing their mailboxes. Users can navigate activities without visual clues when a standard graphical user interface is removed, improving efficiency and intuitiveness. Accessing email messages using a voice-based interface is more efficient and safer when IMAP is used.

The core technologies employed in the system described in reference [8] are speech to text, text to speech, and interactive voice response. Upon their first visit to the website, users must register using voice commands, which will also be captured and stored in the database along with their voiceprint. Subsequently, users will be assigned a unique ID and password, which they can use to access the mail feature post-login. The website's user interface was developed using Adobe Dreamweaver CS3, with a focus on optimizing comprehension and ease of use. To further improve the system's functionality, a "Contact Us" page is available to users, where they can make suggestions or request assistance. This feedback mechanism enables the system's developers to receive constructive feedback from users and address any concerns they may have, leading to continuous improvements and enhancements to the system.

The writers of paper reference [9] present Picture Sensation, a cutting-edge smartphone app made specifically for those with visual impairments. This software makes use of sophisticated audification algorithms, which are in line with the exploratory strategies that individuals with visual impairments typically apply. With Picture Sensation, blind people can explore pictures independently and with real-time audio feedback thanks to an intuitive swipe-based speech-guided interface. This instantaneous input provides a more immersive experience by improving the effectiveness and depth of picture comprehension. By incorporating cutting-edge technologies specifically designed to meet the needs of visually impaired people, the system enables them to engage with and comprehend visuals more efficiently. The project [10] uses an Arduino UNO board to develop an accurate and

affordable tool or handheld gadget for obstacle recognition, greatly enhancing the quality of life for blind people.

In order to help visually impaired people become more integrated into the rapidly changing digital landscape, improve their online communication, and make their lives easier, a method is introduced in reference [11]. By addressing issues like email accessibility, this technology helps those who are blind or visually impaired participate in the creation of a digital India. As demonstrated by Karamad in [12], the accomplishment of this project could act as a source of inspiration for developers, encouraging the development of additional solutions for people who are blind or visually impaired. With the purpose of empowering people in impoverished areas, Karamad is a voice-based crowdsourcing platform that lets users complete tasks with simple phones and earn mobile airtime balance. Reference [13] comes to the conclusion that modern object identification techniques, especially those that make use of Google's text-to-speech API and the quick YOLO Object Detection, show how technology may be used effectively. With a processing speed of 45 frames per second, YOLO is a potential tool for people who are visually handicapped due to its quickness and ability to recognize objects in general. The goal of the article is to turn this technology into an offline navigation aid that guides users with 3D sounds.

The author of [14,15] describes a successful mobile application for mobility aid that gives users access to a portable, adaptable navigational assistive device that combines object detection and obstacle detection. It can be operated with gestures and a voice-command-enabled user interface. In order to produce an obstruction map and alert users to approaching danger, the obstacle detection system uses depth values from several locations. The Automatic Voice Assistant, which allows more powerful and natural interaction, is implemented by the author in reference [16], considerably boosting the experience of interacting with a graphical user interface as per Table 2 mentioned below.

Table 2. Challenges of the work in the literature.

Feature/ Category	Voice-Based System For People Who Are Visually Impaired	Existing Systems
Accessibility	Specifically designed to cater to the needs of users who are visually impaired, providing audio-based interactions and feedback, making them more accessible and inclusive.	May or may not have accessibility features for users who are visually impaired, which can pose challenges to their usability.
Input Method	Relies primarily on voice commands and audio feedback for input and output, making them accessible for users who may have limited or no visual acuity.	Typically relies on visual input methods, such as touch or visual cues, which can be challenging for users who are visually impaired to interact with.
Output Method	Provides audio output in the form of synthesized speech or auditory cues, making it easier for users who are visually impaired to receive information and feedback.	Provides visual output, such as text or graphical interfaces, which may not be accessible to users who are visually impaired without customization or adaptation.
Navigation	Can provide audio-based navigation instructions and guidance, allowing users who are visually impaired to navigate through menus, options, and content using voice commands.	Relies on visual cues, gestures, or touch-based navigation, which can be challenging for users who are visually impaired to follow or interact with.

Table 2. Cont.

Feature/Category	Voice-Based System For People Who Are Visually Impaired	Existing Systems
User Interface	May have minimal or no visual interface, with an emphasis on audio-based interactions, making them more user-friendly for users who are visually impaired.	Usually has a visual interface, which may require customization or adaptation for users who are visually impaired, posing challenges to usability.
Dependency on Visual Information	Relatively low dependency on visual information, as the primary mode of interaction is through voice commands and audio feedback, making them more accessible for users who are visually impaired.	Relies on visual information for input, output, and navigation, which can pose challenges for users who are visually impaired who may have limited or no visual acuity.
Independence	Can promote independence for users who are visually impaired by providing them with a means to interact with technology and access information without relying heavily on visual cues or assistance from others.	May require assistance or customization to be accessible for users who are visually impaired, limiting their independence in using the application effectively.

3. Proposed Methodology

Adding external libraries or modules to an Android project is a crucial step in the development process as it provides access to additional functionality and features that are not available in the standard Android framework. By incorporating these libraries, developers can streamline the development process, reduce code complexity, and enhance the overall functionality of their applications. Once the necessary dependencies are added, the user interface can be designed using XML. This enables developers to create visually appealing and intuitive applications and provides a more seamless user experience. A user may read the app's feature or function by swiping left on the screen. Voice input may be started on the screen by swiping right. The voice command will automatically redirect to that specific activity when the user issues it.

Let us assume that when a user says "read", the "read" activity starts up automatically. In order to read the text in the photo aloud, the user only needs to tap on the screen to take a picture.

4. Methods Applied

- Text to speech (TTS) is a technique for turning written text into spoken speech. For speech output and voice feedback for the user, TTS is crucial. TTS is used in applications when audio support is necessary. TTS will translate a voice command the user enters into text and carry out the specified activity.
- Voice to text (STT): Android comes with a built-in function called "speech-to-text" that enables a user to provide a speech input. The speech input will be translated to text in the background while TTS activities are carried out.

In Figure 2, we can clearly observe that this flowchart serves as a foundational representation of the application's major features and operational functions. It describes the sequential processes and interactions that users can expect to encounter while using the application. It provides a clear and comprehensive description of how the program works by breaking down the process into a visual diagram.

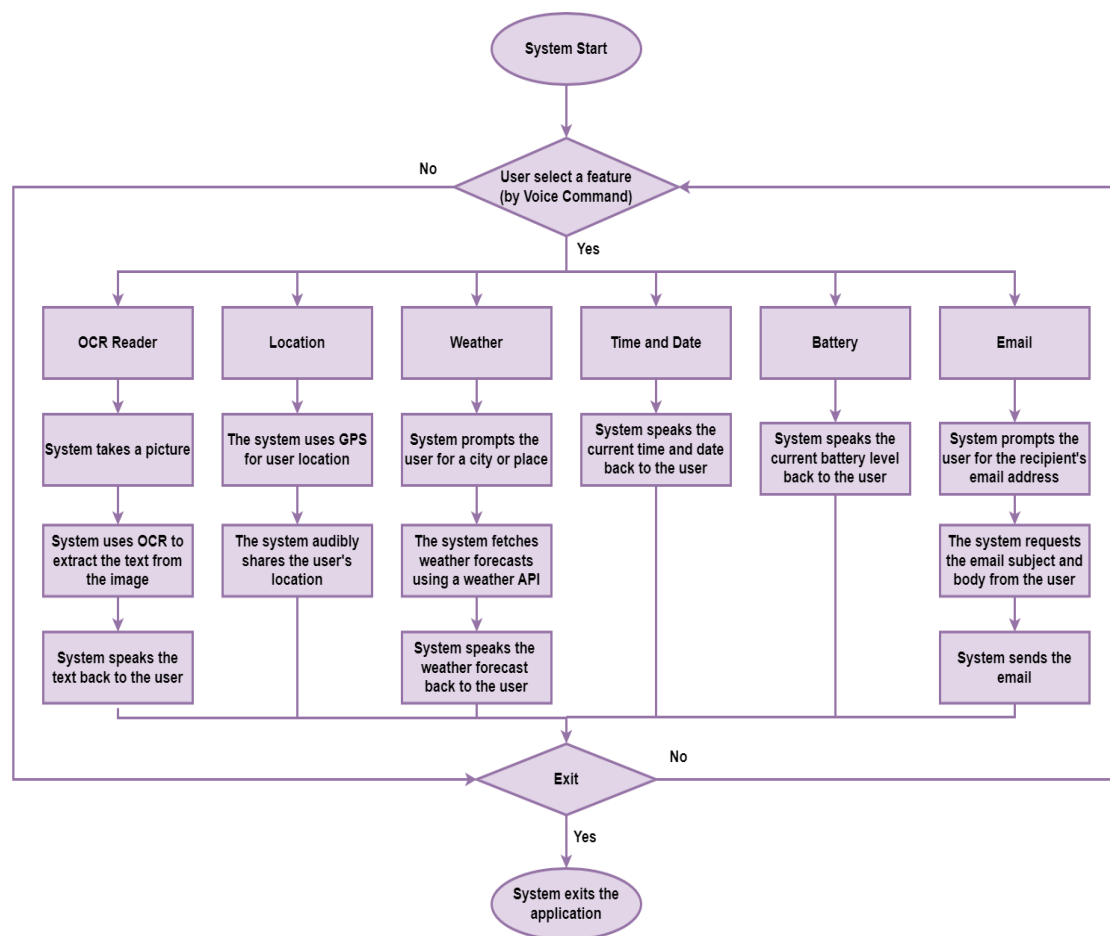


Figure 2. Flowchart illustrating the application's features.

5. Result and Discussion

As a result, the system suggests the following uses:

- OCR reader: After swiping right on the screen, the user must say “read”, at which point the system will ask if the user wants to read; they must select yes to proceed or no to return to the menu. An illustration of how to convert a picture to text is shown in Figure 3.
- Email: “Voice Command” enables spoken commands for email composition and inspection, allowing hands-free email control.
- Location: In this scenario, the user must first say “position” before tapping on the screen to see their present location.
- Calculator: “Calculator” must be said by the user. Subsequently, the user must tap the screen and indicate what needs to be calculated; the application will then provide the result.
- Date and time: The user must state current date and time in order to check them.
- Battery: The user must say “battery” in order to view the phone's current battery state.
- Weather: The user in this scenario would first say “weather”, followed by the name of the city. The weather for that particular city will be shown after that application. Figure 4 depicts an example of a weather report.

Some of the main features of the proposed system are highlighted in the bullet points above. This facilitates a blind person's ability to carry out their regular responsibilities.

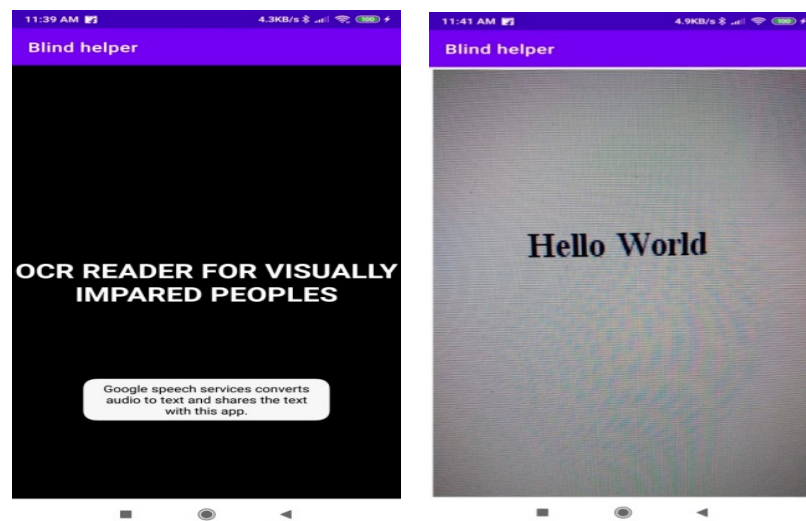


Figure 3. Conversion of images to text.

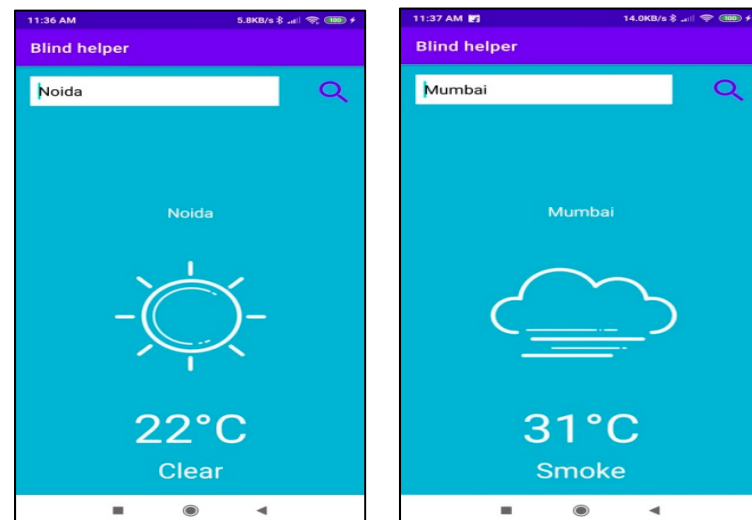


Figure 4. The application's weather report.

6. Conclusions

The majority of our everyday tasks are currently completed through mobile apps on smartphones. However, using mobile phones and tablets to access these mobile apps requires support for those with vision difficulties. Google has been creating a variety of mobile apps for persons with vision impairments using Android applications. However, it still has to adapt and include appropriate artificial intelligence algorithms to provide the app with additional useful features. Two ecologically friendly ideas for blind persons were shown in this paper. We provided details on the application for blind people. For blind people, this proposed system will be more useful. This application has to be developed for the future. The technology is utilized by blind individuals, but it is also accessible to sighted people. The proposed approach offers a safe and independent means of mobility for individuals with visual impairments, which can be used in any public setting. By using a voice-based system, users can easily specify their destinations and navigate their surroundings autonomously without the need to acquire assistance from others. This innovative approach has the potential to significantly improve the mobility and independence of individuals who are visually impaired, allowing them to navigate public spaces with greater ease and confidence. This system was created quickly, cheaply, and with minimal power use. The proposed system is dynamic and takes up less room. This method is

more effective than others that are already in place. It may be improved much more by including object identification capabilities utilizing machine learning technologies and the TensorFlow model, and by adding a reminder feature.

Author Contributions: In this study, L.K. handled the concept, design, data collection, and analysis. A.S.K. focused on drafting the paper, conducting the literature review, writing, and revising the manuscript. A.J. supervised the research, critically reviewed the content, and provided valuable feedback throughout the publication process. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This edition contains no newly produced data.

Acknowledgments: The authors extend their gratitude to Arun Prakash Agrawal and Rajneesh Kumar Singh for their valuable advice and mentorship and appreciate their insightful criticism. Thanks are also given to the DC members for their cooperation and discussions that shaped the paper's key ideas. Sharda University is acknowledged for providing necessary facilities and a conducive research environment. Lastly, heartfelt thanks are extended to the authors' friends and family for their unwavering support and encouragement throughout the research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Burton, M. The Lancet Global Health Commission on global eye health: Vision beyond 2020. *IHOPE J. Ophthalmol.* **2022**, *1*, 16–18. [CrossRef]
2. Blindness and Vision Impairment. Available online: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> (accessed on 10 August 2023).
3. Vijayalakshmi, N.; Kiruthika, K. Voice Based Navigation System for the Blind People. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2019**, *5*, 256–259. [CrossRef]
4. Rizky, F.D.; Yurika, W.A.; Alvin, A.S.; Spits, W.H.L.H.; Herry, U.W. Mobile smart application b-help for blind people community. *CIC Express Lett. Part B Appl.* **2020**, *11*, 1–8. [CrossRef]
5. Available online: <https://www.wbur.org/npr/131050832/a-mystery-why-can-t-we-walk-straight> (accessed on 10 August 2023).
6. Suresh, A.; Paulose, B.; Jagan, R.; George, J. Voice Based Email for Blind. *Int. J. Sci. Res. Sci. Eng. Technol. (IJSRSET)* **2016**, *2*, 93–97.
7. Badigar, M.; Dias, N.; Dias, J.; Pinto, M. Voice Based Email Application For Visually Impaired. *Int. J. Sci. Technol. Eng. (IJSTE)* **2018**, *4*, 166–170.
8. Ingle, P.; Kanade, H.; Lanke, A. Voice Based email System for Blinds. *Int. J. Res. Stud. Comput. Sci. Eng. (IJRSCSE)* **2016**, *3*, 25–30.
9. Souman, J.L.; Frissen, I.; Sreenivasa, M.N.; Ernst, M.O. Walking Straight into Circles. *Curr. Biol.* **2009**, *19*, 1538–1542. [CrossRef]
10. Guth, D. Why Does Training Reduce Blind Pedestrians Veering? In *Blindness and Brain Plasticity in Navigation and Object Perception*; Psychology Press: Mahwah, NJ, USA, 2007; pp. 353–365.
11. Khan, R.; Sharma, P.K.; Raj, S.; Verma, S.K.; Katiyar, S. Voice Based E-Mail System using Artificial Intelligence. *Int. J. Eng. Adv. Technol.* **2020**, *9*, 2277–2280. [CrossRef]
12. Randhawa, S.M.; Ahmad, T.; Chen, J.; Raza, A.A. Karamad: A Voice-based Crowdsourcing Platform for Underserved Populations. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI'21). Association for Computing Machinery, New York, NY, USA, 8–13 May 2021; pp. 1–15. [CrossRef]
13. Annapoorani, A.; Kumar, N.S.; Vidhya, V. Blind—Sight: Object Detection with Voice Feedback. *Int. J. Sci. Res. Eng. Trends* **2021**, *7*, 644–648.
14. See, A.R.; Sasing, B.G.; Advincula, W.D. A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind. *Appl. Sci.* **2022**, *12*, 2802. [CrossRef]
15. Devika, M.P.; Jeswanth, S.P.; Nagamani, B.; Chowdary, T.A.; Kaveripakam, M.; Chandu, N. Object detection and recognition using tensorflow for blind people. *Int. Res. J. Mod. Eng. Technol. Sci.* **2022**, *4*, 1884–1888.
16. Shivani, S.; Kushwaha, V.; Tyagi, V.; Rai, B.K. Automated Voice Assistant. In *ICCCE 2021*; Springer: Singapore, 2022; pp. 195–202.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.