Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Real-Time View Assistance for the Blind using Image Processing

Abhiram M[1], Adarsh K Rajeev[2], Akhil Raj V[3], Anirudh K[4], Manoj M[5]

Computer Science, Jawaharlal College of Engineering and Technology, KTU, Palakkad, India

E-mail: [1]abhirammktp@gmail.com, [2]adarshkrajeev2020@gmail.com, [3]akhilrv031@gmail.com, [4]anirudhedavalath@gmail.com, [5]manojm@jawaharlalcolleges.com

Abstract

Visually impaired people usually have difficulty in doing daily activities. Imagine a future where visually impaired people can seamlessly and independently identify objects and people in their environment. The aim of this research is to increase the independence and mobility of visually impaired people by developing a real-time object and person recognition system. This system uses the power of machine learning and uses computer vision techniques to accurately identify and classify objects and people in the user's environment. Through the integration of speakers or headphones, the system provides auditory feedback to the user and conveys important information about the detected object or person. By combining advanced image processing algorithms with audio output, this solution serves as a valuable tool for visually impaired people, allowing them to effectively perceive and understand their surroundings. This innovative approach demonstrates the potential of technology to bridge access gaps and empower people with visual impairments in their daily lives.

Keywords: Computer Vision, Optical Character Recognition (OCR), Object Recognition, Facial Recognition, Text to Speech, Raspberry Pi.

1. Introduction

In a world of continuous technological innovation, the need for solutions that enable inclusion and accessibility is clearer than ever. This research begins the journey to create a "Personal assistant for the blind", a transformative effort based on the use of advanced

DOI: https://doi.org/10.36548/jiip.2024.2.002
Received: 15.03.2024, received in revised form: 16.04.2024, accepted: 02.05.2024, published: 11.05.2024

technology. Designed to offer real-time support and empowerment to blind people, this personal assistant seeks to redefine the way people with vision problems interact with their environment [9-12]. This research is a response to the enormous challenges faced by blind people navigating a world often designed for the sighted. By harnessing the power of artificial intelligence and machine learning, this initiative aims to improve the overall quality of life for people with visual impairments. With real-time object recognition, facial recognition, and text-to-speech conversion, the personal assistant aims to enhance the quality of life of the blind peoples [13-15].

2. Related Work

Automated Walk-in Assistant for the Blinds [1]

The work aims to develop a walking solution for blind people that detects obstacles and provides warning signals based on the distance of the obstacle. The glasses, embedded with a walking assistant, feature an obstacle detection module and an alarm generator. The ultrasonic sensor in the module can detect objects within 3 meters at a 60-degree angle. The module generates a high-frequency signal, evaluates it against an echo, and measures the distance from the object. This data is then used to generate alarms based on the distance between the objects. Experimental studies show that this approach offers a more affordable and lightweight walk-in aid for blind individuals.

A Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired and Blind [2]

A new system has been developed to assist people with low vision and blindness (VIB) in traveling. The system uses mobile phones and depth camera functions for obstacle avoidance and object recognition. It includes a voice-controlled mobile app and gesture controls for navigation. The system collects depth values from 23 coordinate points and analyzes them to determine obstacles in the head, trunk, ground area, or complete body. The probe detects open objects up to 1.6 meters away, providing a reliable warning system. The object detection function includes an interactive feature that allows users and the device to interact, enabling them to find hidden objects more than 80 percent of the time. This flexible and compact system is designed for use without additional hardware.

Digital Assistant for Blind [3]

Blind individuals and partially sighted individuals often lack access to the internet due to a lack of low-cost solutions, including hardware and software. Current technologies rely on expensive Braille displays and keyboards, and less than 2% of the visually impaired population can interpret Braille. A voice-controlled system has been designed to address this issue, allowing users to receive and send emails, access daily news and weather forecasts, set reminders and alarms, and take notes. This voice-controlled system aims to provide a more accessible way for.

Virtual Assistant for Blind People [4]

In today's advanced technology, people with visual disabilities face social limitations and struggle to help themselves in unfamiliar environments. Inclusive technology, such as Artificial Intelligence, machine learning, and image and text recognition, can help these individuals. A mobile application for Android is being developed to implement these features, including voice assistants, image recognition, currency recognition, e-books, and a chatbot. The software can recognize items in the environment and perform text analysis, making it an effective approach for the blind to engage with the world and use technology features.

Corridor-Walker: Mobile Indoor Walking Assistance for Blind People to Avoid Obstacles and Recognize Intersections [5]

The Corridor-Walker is a smartphone-based system designed to assist the blind in navigating interior corridors. It uses a LiDAR sensor to create a 2D occupancy map of the environment, generates obstacle avoidance paths, and detects upcoming intersections on the grid map. The system then navigates the user to the generated path, providing information about the existence and shape of each intersection through vibrations and sound feedback. A study with 14 blind participants showed that the Corridor-Walker helped them avoid obstacles, reduce reliance on walls, and recognize intersections.

"Siri Talks at You": An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind [6]

Voice-activated personal assistants (VAPAs) offer potential benefits to blind individuals, but studies have not explored their usage and potential barriers. A study of fourteen legally blind adults found that while they appreciated VAPA's accessibility, they faced challenges with input, responses, and control of information. The study suggests that design should consider privacy, situational factors, trust concerns, and synchronize visual and non-visual cues to promote inclusivity and ease of use.

EyeSpy - A Personal Assistant for the Blind and Visually Impaired [7]

The study aimed to evaluate the effectiveness of EyeSpy, a navigation device prototype with a companion lifestyle application. EyeSpy uses echolocation to detect objects and emits signals when they approach at a faster pace. The application has four functions: instructions, location, points, and fitness. The device works independently and is covered in PVC hospital fabric. The application uses Kotlin and XML to convert location data into an audible form with a human voice. Participants found EyeSpy to be practical and easy to use, increasing their walking pace and confidence in blindly traversing spaces. Despite some detection issues with an ultrasonic sensor, the device received a 5-star rating. Further testing and development are needed to improve the device's usability and non-invasive measurement of heart rate, blood pressure, and oxygen saturation using photoplethysmography at two wavelengths.

Evaluating AI Techniques for Blind Students Using Voice Activated Personal Assistants [8]

The research is centered on developing an AI system that would better the process of academic registration for blind students. The aim of this study is to create a model that will enable visually impaired learners to have access to academic service requests and activities. In order to combine theoretical constructs with practical applications, the model works based on empirical data from an archive approach. Its real-world effectiveness was demonstrated by integrating it into existing institutional portals, where it was partly validated. Among other classifiers, this one achieves better outcomes by working together with voice-activated personal assistants (VAPAs), text mining techniques, bag-of-words methodologies, and case-based reasoning that are employed in the system.

3. Proposed Work

This innovative system has been meticulously crafted to serve as a ground breaking aid for individuals with visual impairments, empowering them to comprehend their environment and navigate effortlessly. It consists of a structured framework that consists of hardware components, software modules, and user interfaces. The trained model, now equipped with a profound understanding of its input data, undergoes deployment onto a single-board computer seamlessly connected to a camera and a speaker. As the camera captures real-time images of the surroundings, the model works tirelessly to discern and identify objects and individuals within the frame. The auditory component of the system comes to life as the speaker converts the textual data derived from the image recognition into clear and comprehensible speech, thereby enabling a visually impaired individual to navigate their surroundings with ease.

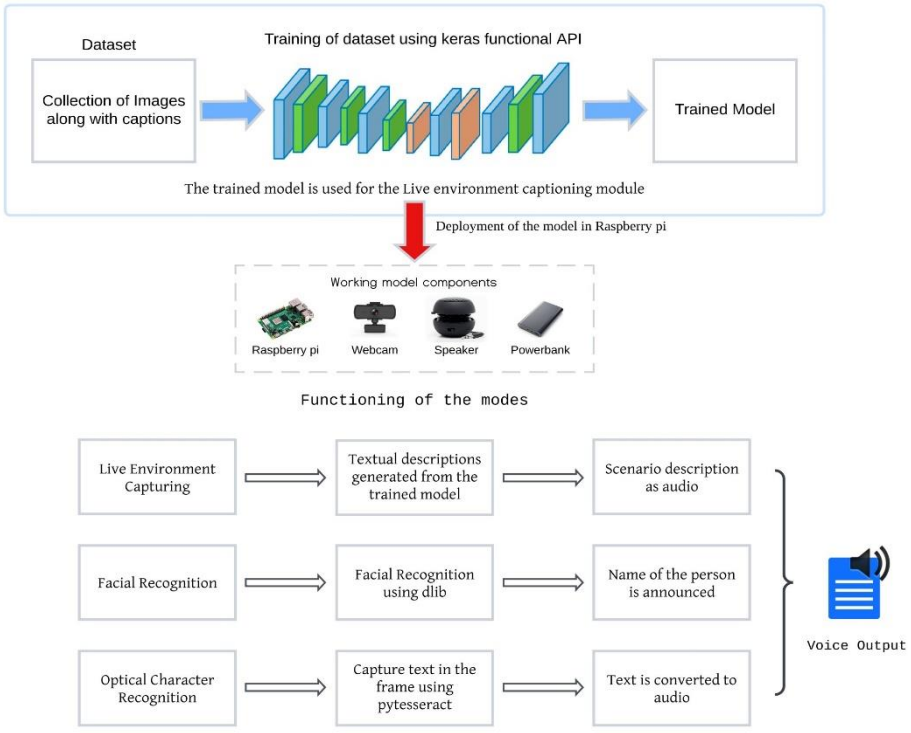


Figure 1. System Architecture

The System architecture depicts the overall view of the system. The system consists if the following components, Raspberry Pi, Webcam, Speaker, Power bank. All the required dependencies for the proper working of the system is installed on the Raspberry Pi . The dataset contains collection of images along with captions. This dataset is trained using keras functional

API which is basically used for easy building and training models. This trained dataset is used for the working of the live environment capturing module in the system. When we connect power to the system the system will start working. The system has three modes: Live environment capturing, Facial recognition, Optical Character Recognition. The live environment capturing module will take input frame from the camera about the object present or the current scene and generate output audio according to the textual descriptions generated according to the dataset. The Facial recognition module will take the frame of person present in front of the user and validate it from the stored image in the system and generate the name of the person as output. This face recognition is done using dlib. The dlib is a popular toolkit for machine learning that is used primarily for computer vision and image processing tasks, such as face recognition, facial landmark detection, object detection, and more. The Optical Character Recognition module will take a frame and analyses texts present in it. It is done using Pytesseract library in python. Pytesseract is a Python library that provides an interface to the Tesseract optical character recognition (OCR) engine. OCR is a technology used to recognize and extract text from images, scanned documents or other visual media. The output from the system will be audio output which is generated by gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate text-to-speech API. The audio output will be played through the speaker.

'Roboflow' is an open-source platform that facilitates the management and preprocessing of image datasets for machine learning projects, particularly in the domain of computer vision. It provides a suite of tools and workflows to streamline the process of preparing image data for training machine learning models.

It helped in developing the datasets. An extensive dataset (images of objects required for the Live Environment Captioning module) around **'15000'** samples are used, featuring images accompanied by informative captions. Each current frame will be captured from the camera, and will be validated from the dataset and will produces output.

3.1 Data Flow Diagram

The Figures 2 to 6 illustrates the data flow diagram of the proposed work.

Figure 2. Level 0

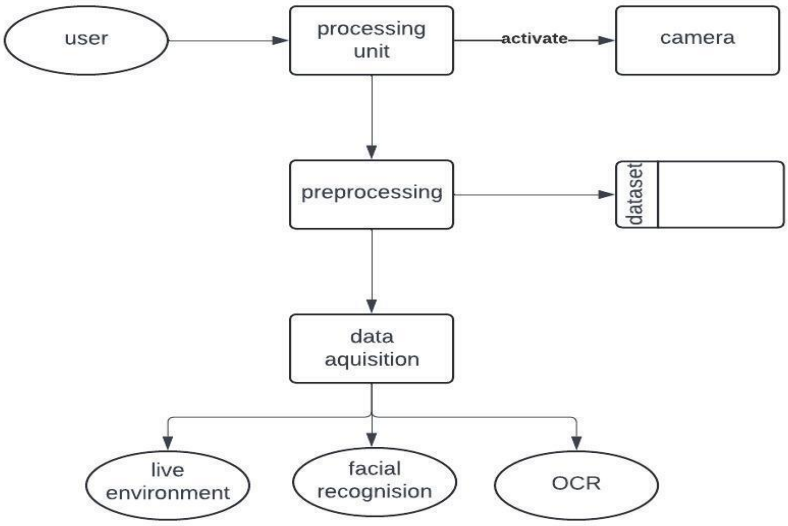


Figure 3. Level 1

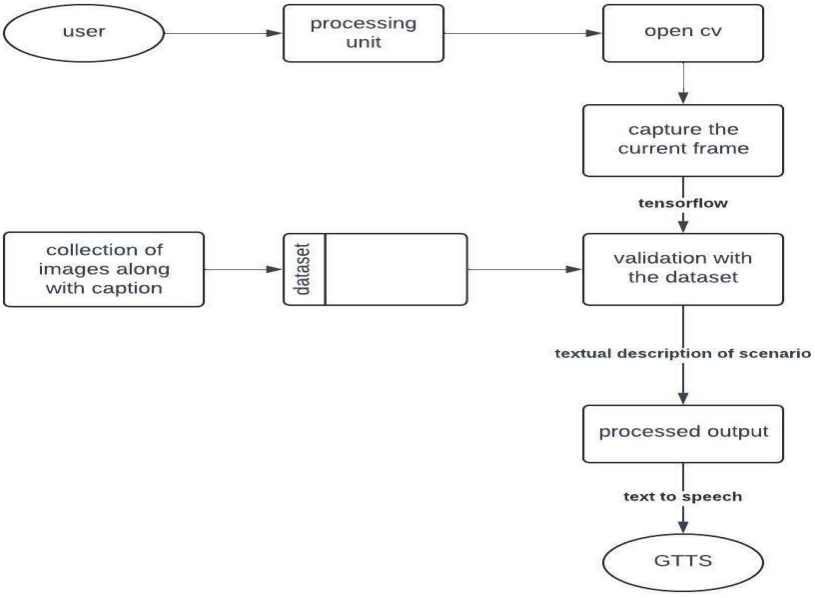


Figure 4. Level 2.1

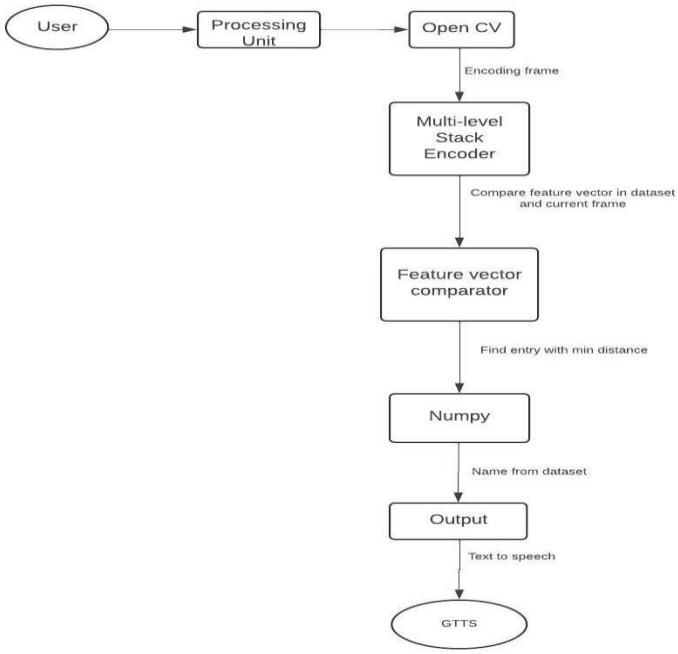


Figure 5. Level 2.2

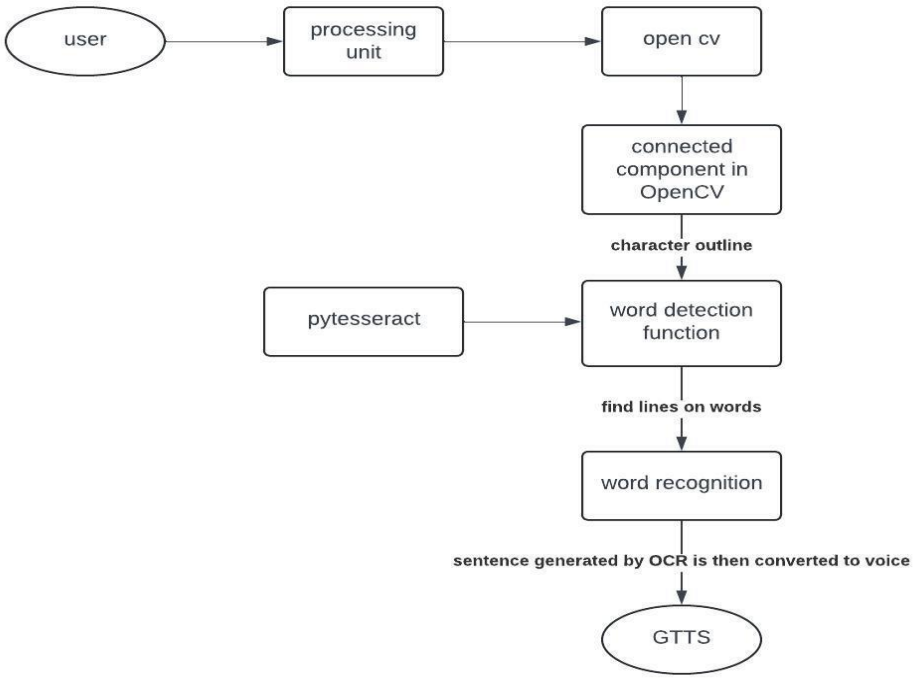


Figure 6. Level 2.3

4. Results and Discussion

The personal assistant for the blind underwent extensive testing in various environments to evaluate its performance. Results indicate working of almost all of the features mentioned. The additional modes such as Facial recognition and Optical Character Recognition was working perfectly.

- **Live Environment Capturing Module:** The live environment capturing module will takes a frame of the current scenario or objects and provides a description of the scenario as audio output. This helps the blind people to recognize their surrounding and objects within them.
- **Face Recognition Module:** The facial recognition module will identify the person in front of the user and announce the name of the person. Unrecognized persons will be announced as “Unknown”.
- **Optical Character Recognition (OCR) Module:** The optical character recognition module will identify the texts in the frame and it will read it to the user. This helps in reading paragraphs from books.

4.1 Requirements and Analysis

The software and the hardware requirements of the proposed work is illustrated in Table 1.

Table 1. Requirements

H/W or S/W Requirements	Specification
Raspberry pi	Raspberry Pi serves as a compact and versatile computing platform. It is used as the processing unit.
Camera	Camera is the visual sensory input in the system, playing a crucial role in real-time environmental perception for blind individuals.

Speaker	The speaker is a vital component that serves as the auditory output interface. The gtts output generated by the system will be announced to the user through the speaker.
Python	Python plays a pivotal role in this, serving as the primary programming language. Various python libraries are used for the working of the system features.

4.2 Prototype

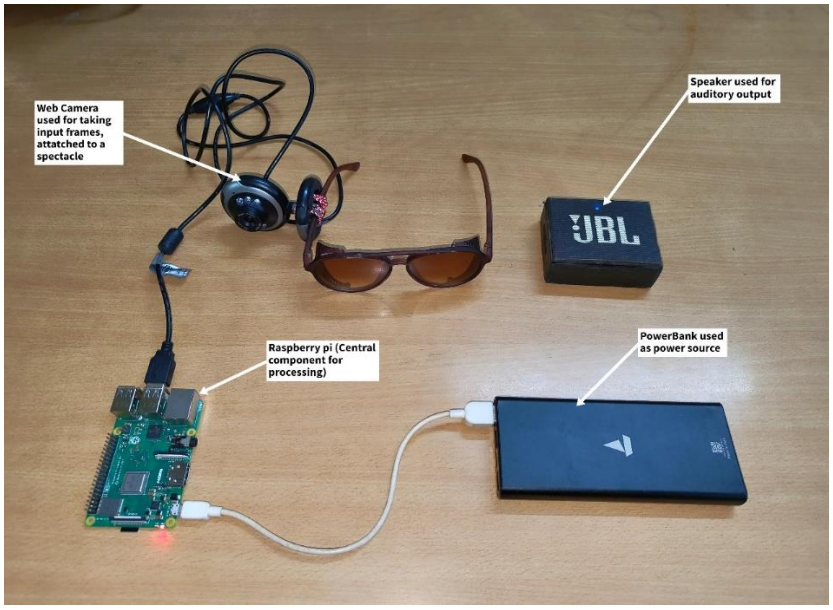


Figure 7. Prototype Model

The model consists of Raspberry Pi, Camera, Speaker and a Power source.

- **Raspberry pi** is the core component that acts as the processing unit. It runs the code fragments and provides output.
- **Camera** is used to take the input frames to the processing unit to process and produce output.
- **Speaker** is the output device used to get the voice outputs.
- **Power Source** is the necessary item for starting and working of the system

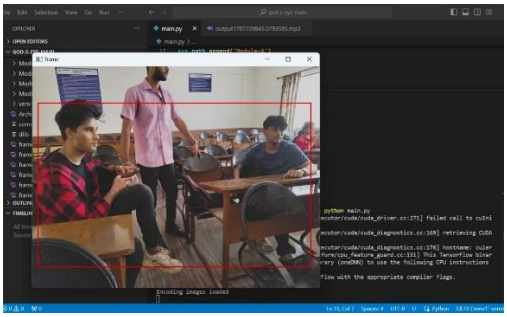
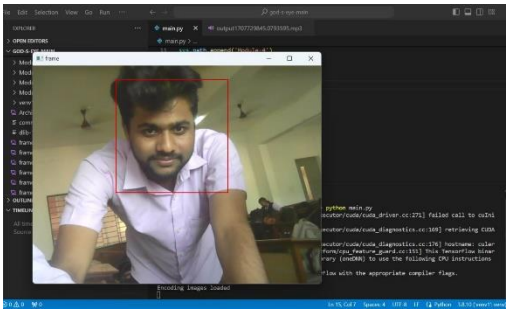
4.2.1 Working

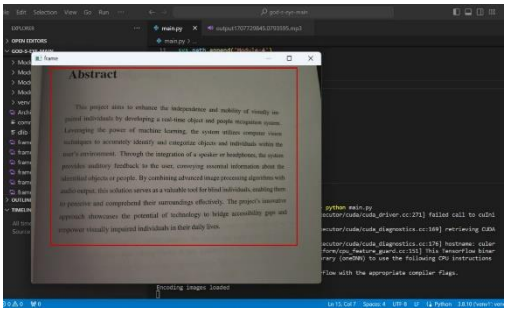
1. The system will start working when the power source is connected to the central component Raspberry Pi.
2. The camera will get activated and it will start taking input frames.
3. The system works in 3 modes: Live environment capturing, facial recognition and OCR.
4. The voice output generated will be played through the speaker.

4.3 Test Cases

The Table 2 shows illustrates the test case of the proposed work.

Table 2. Test Cases

Test Case Type	Input Frame	Expected Audio Output	Obtained Audio Output
Live Environment Capturing		The system should produce auditory output of the recognized objects	Group of people are sitting around table
Facial Recognition		The name of the person is produced as output	Anirudh

Optical Character Recognition		Texts from the input frame will be produced as auditory output	Texts in the frame will be played as audio.

5. Conclusion

In conclusion, the culmination of machine learning and computer vision in this serves as an invaluable aid for the visually impaired community. By seamlessly integrating these technologies, the system transcends mere object and people recognition; it becomes a beacon of assistance for those with visual impairments. This innovative solution narrates the environment for the visually impaired, effectively transforming their surroundings into a tangible, accessible reality. By providing detailed descriptions and guidance, it empowers individuals to navigate their world with increased independence and confidence. In essence, this ground breaking system isn't just about recognizing elements; it's about bridging the gap between the sighted and visually impaired, enriching lives and fostering a more inclusive society. The focus now shifts to the implementation phase, where hardware components will seamlessly come together, new modules will augment the system's capabilities, and rigorous testing will validate its real-world effectiveness. This phase is not just a continuation but a realization of the vision to empower and transform the lives of individuals with visual impairments through a sophisticated and user-friendly personal assistant.

The journey doesn't end here; it evolves. The "Personal Assistant for the Blind" is not merely a technological innovation; it represents a step towards a more inclusive and accessible future. As moving forward, the commitment remains unwavering, driven by the belief that technology can be a powerful force for positive change in the lives of those who need it most.

5.1 Future Scope

The future scope of the project is expansive, with opportunities for further development, enhancements, and broader impact. It may include:

- Continuous improvement and integration of more advanced machine learning algorithms can enhance the accuracy and efficiency of real-time object recognition and other intelligent features.

- Consideration of wearable devices, such as smart glasses or haptic feedback devices, can provide a more seamless and unobtrusive user experience, further enhancing the personal assistant's functionality.

- Expanding the range of functionalities, such as integrating additional services like navigation, public transportation information, or even social interaction features, can make the personal assistant a more comprehensive tool.

References

[1] Muhammad Sheikh Sadi, Saifuddin Mahmud, Md Mostafa Kamal, and Abu Ibne Bayazid. Automated walk-in assistant for the blinds. In 2014 International Conference on Electrical Engineering and Information & Communication Technology, pages 1–4. IEEE, 2014.

[2] aron Raymond See, Bien Grenier Sasing, and Welsey Daniel Advincula. A smartphone based mobility assistant using depth imaging for visually impaired and blind. Applied Sciences, 12(6):2802, 2022.

[3] Prince Bose, Apurva Malpthak, Utkarsh Bansal, and Ashish Harsola. Digital assistant for the blind. In 2017 2nd international conference for convergence in technology (I2CT), pages 1250–1253. IEEE, 2017.

[4] Avanish Vijaybahadur Yadav, Sanket Saheb Verma, and Deepak Dinesh Singh. Virtual assistant for blind people. International Journal, 6(5), 2021.

[5] Masaki Kuribayashi, Seita Kayukawa, Jayakorn Vongkulbhisal, Chieko Asakawa, Daisuke Sato, Hironobu Takagi, and Shigeo Morishima. Corridorwalker: Mobile indoor walking assistance for blind people to avoid obstacles and recognize intersections. Proceedings of the ACM on Human-Computer Interaction, 6(MHCI):1–22, 2022.

[6] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. ” siri talks at you” an empirical investigation of voice-activated personal assistant (vapa) usage by individuals who are blind. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pages 249–258, 2018.

[7] Ian Korovinsky and Sam Dabir. Eyespy-a personal assistant for the blind and visually impaired. 2022.

[8] Tariq S Almurayziq, Gharbi Khamis Alshammari, Abdullah Alshammari, Mohammad Alsaffar, and Saud Aljaloud. Evaluating ai techniques for blind students using voice-activated personal assistants. IJCSNS, 22(1):61, 2022.

[9] Al-Muqbali, Fatma, Noura Al-Tourshi, Khuloud Al-Kiyumi, and Faizal Hajmohideen. "Smart Technologies for Visually Impaired: Assisting and conquering infirmity of blind people using AI Technologies." In 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC), pp. 1-4. IEEE, 2020.

[10] Khan, Muiz Ahmed, Pias Paul, Mahmudur Rashid, Mainul Hossain, and Md Atiqur Rahman Ahad. "An AI-based visual aid with integrated reading assistant for the completely blind." IEEE Transactions on Human-Machine Systems 50, no. 6 (2020): 507-517.

[11] Hsu, Ling-Yuan, Hwai-Tsu Hu, and Hsien-Hsin Chou. "A high-capacity QRD-based blind color image watermarking algorithm incorporated with AI technologies." Expert Systems with Applications 199 (2022): 117134.

[12] Kumar, Naveen, Sanjeevani Sharma, Ilin Mariam Abraham, and S. Sathya Priya. "Blind assistance system using machine learning." In International Conference on Image Processing and Capsule Networks, pp. 419-432. Cham: Springer International Publishing, 2022.

[13] Strumillo, Pawel, Michal Bujacz, Przemyslaw Baranski, Piotr Skulimowski, Piotr Korbel, Mateusz Owczarek, Krzysztof Tomalczyk, Alin Moldoveanu, and Runar Unnthorsson. "Different approaches to aiding blind persons in mobility and navigation in the “Naviton” and “Sound of Vision” projects." Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies (2018): 435-468.

[14] Asokan, Adithya, N. S. Afsal, Aswathy Nandakumar, Aswathy Rajan, and R. M. Sherin Rappai. "Reading Assistance for Blind." International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering 5, no. 4 (2017).

[15] Gharat, Madhura, Rizwan Patanwala, and Adithi Ganaparthi. "Audio guidance system for blind." In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), vol. 1, pp. 381-384. IEEE, 2017.