International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I8P107

Volume 74 Issue 8, 99-108, August 2026


*Original Article*

# Development of a Smart Wearable Assistive Device with Object Detection and Text-to-Speech Functionality for Visually Impaired Individuals using Raspberry Pi

Chona R. Dagatan[1], Jestoni P. Tan[2], Joy Stiffany A. Amodia[3], Lei Jo C. Bacu[4], Brett H. Callanga[5], Jeff Niño G. Caparoso[6], Avril Mae O. Malaga[7], Jedelou B. Pepito[8]

[1,3,4,5,6,7,8]*Computer Engineering, Cebu Technological University-Danao Campus, Philippines.*
[2]*Electrical Engineering, Cebu Technological University-Danao Campus, Philippines.*

[1]*Corresponding Author : chona.dagatan@ctu.edu.ph*

Received: 02 December 2025 Revised: 10 June 2026 Accepted: 24 June 2026 Published: 29 August 2026

***Abstract -*** *Visually impaired individuals struggle in terms of their mobility, awareness, and overall quality of life. Traditional aids, such as canes, offer only limited assistance. This capstone project aims to develop a smart wearable device to enhance environmental awareness and assistance for visually impaired individuals in Danao City. The smart wearable assistive device is composed of a Raspberry Pi 4, ultrasonic sensors, a web camera, and text-to-speech technology that is used to deliver audio feedback that helps in better comprehension of information for these individuals. Experimental tests were conducted using percent error and paired t-tests to evaluate the overall performance of the wearable device. Survey questionnaires and interviews were also conducted to test the effectiveness of the device based on participants' experiences. The quantitative results show that the smart wearable device consistently underperformed in its different functions. In distance range detection, the system has an average distance of 2.5 meters from the desired range, with a statistical analysis ($t = 6.74$, $p < 0.05$) indicating a significant difference. Object detection accuracy identifies an average of 9 out of 15 objects with a statistical analysis of ($t = -8.96$, $p < 0.001$) and has an effect size of 2.31. The system shows an average of 1-2 missed words per trial and a statistical analysis of ($t = -5.13$, $p < 0.001$). Overall, the findings of this study confirm that the smart wearable device's potential, specifically in text-to-speech systems, helps to enhance independence and confidence in common environments.*

***Keywords -*** *Computer vision, Object detection, Text-to Speech, YOLO, Wearable device.*

## 1. Introduction

As of 2020, global estimates indicate that approximately 43.3 million individuals were living with blindness,295 million people were affected by Moderate to Severe Visual Impairment (MSVI) [3]. Women comprised 60% of the blind population and 59%of individuals experiencing MSVI.

Vision loss greatly impacts personal independence in doing everyday tasks. Although individuals without sight often rely on their senses of touch and hearing, it is not enough.

However, the development of real-time object detection and recognition systems has significantly improved the ability of visually impaired individuals to improve mobility and independence. People with blindness or visual impairment frequently encounter challenges in distinguishing objects. While traditional mobility aids like the white cane offer important support, they are limited in range. Modern versions address this limitation by machine learning with sensor integration. However, there are cost and a technology gap among old generation. Hence, the study aims to develop a computing assistive device that is a user-centric feedback loop validated by a professional demographic integrated with Raspberry Pi 4, Ultrasonic and Camera.

This study was evaluated by technical accuracy, users' perceived Text-to-Speech Clarity and Navigational Confidence within a specialized setting. This approach emphasizes that the study not merely focuses on detecting objects but also focuses on environmental interaction of the users.

## 2. Review of Related Literatures

### *2.1. Related Literatures*

#### *2.1.1. Object Detection*

Many studies show that researchers are creating smart devices to help people with visual impairments. These devices use technologies such as GPS, ultrasonic sensors, computer vision, and artificial intelligence to assist users in moving around, recognizing objects, and staying safe. Most of these studies aim to create devices that are low-cost, easy to use, and useful in everyday life. Features like real-time feedback and obstacle detection are also important.

These technologies can assist with basic tasks such as moving around and more complex activities like recognizing objects and providing emergency help.

Machine learning-based object detection systems are being designed to support visually impaired individuals [6]. A wearable device has been developed, powered by a Raspberry Pi and CNNs that can detect objects and read text in real time [5]. The system is able to recognize everyday items and printed materials, especially road signs and Philippine currency. It also has a battery alert that tells users what the battery is doing.

The study suggests that this type of technology can help visually impaired individuals become more independent and less reliant on others. [7] uses the YOLO algorithm to detect indoor objects and give immediate audio feedback to the user, and was showed effective. [8] uses the YOLO algorithm and is converted into Torch Script so it can run on mobile devices.

Another study developed a method for turning textual input into audible speech by concentrating on the design and implementation of a Text-to-Speech (TTS) synthesizer [11].

Their study emphasizes how important TTS systems are for improving accessibility. The AI-powered smart blind stick, which uses object detection and speech recognition to provide real-time feedback. This gadget can identify and categorize objects, giving users auditory information and empowering them to navigate on their own.

They developed this system by combining a Raspberry Pi, ultrasonic sensors, a water level sensor, a GPS module, a vibration motor, and a PIC18F4525 microcontroller [15]. The device improves user confidence and independence. Real-time object detection has become a vital function in wearable gadgets 25 designed for visually impaired people.

[12] developed a smart headgear that uses a Single-Shot Multibox Detector MobileNet v2 model to efficiently interpret visual data while reducing processing needs.

This technology, using advanced machine learning techniques, provides a unique way to increase road safety and situational awareness for visually impaired people.

*2.1.2. IOT-Based Device for the Visually Impaired*

One design utilizes specialized infrared, ultrasonic, and puddle sensors to detect stairs, obstacles within four meters, and wet surfaces, alerting users via voice prompts [11]. Another system integrates machine learning, cameras, and GPS tracking to identify obstacles in real time while allowing guardians to monitor the user's location and receive emergency alerts via a mobile app [13].

Smart blind sticks were designed to seamlessly integrate a GPS tracking system with SOS capabilities that are based on the Internet of Things. According to [14], this system uses a GPS module and microcontroller to continuously track user locations and send data to registered users and a central server.

In an emergency, the SOS feature, which notifies adjacent devices, is extremely helpful in guaranteeing timely notifications to family members.

[9] utilizes the Tesseract OCR tool on a Raspberry Pi and camera accessory in order to recognize and convert the written English language into text.

Then, through the application of the Google Text to Speech (GTTS) tool. Another benefit is the ability to use this device to perform translations to the regional languages of Telugu and Hindi, among others.

## 3. Materials and Methods

### *3.1. System Overview*

The Smart Wearable Device is designed to help visually impaired people move safely and independently. The Raspberry Pi 4B is used as the main control unit of the device, controlling and connecting all parts of the system. An Ultrasonic Sensor is also used to detect the distance of objects and help avoid any obstacles and collisions.

A web camera is used to determine different types of objects in the surroundings, providing more detailed information aside from the distance alone.

The system converts the detected information into voice messages using the Text-to-Speech (TTS) technology, which users can hear through earphones. A buzzer is also added to send warnings when obstacles are detected nearby.

Additionally, the device also is integrated with a rechargeable battery that powers the system, allowing it to work for long hours while using it in low energy designed in a way for daily activities, enabling users to move freely and comfortably.

Overall, the system helps the user to understand their environment and move more confidently.

Table 1 details the primary hardware components of the proposed device based on cost, weight and suitability for edge deployment.

**Table 1. Hardware configuration and justification**

| Component | Specification | Justification |
|---|---|---|
| Raspberry Pi 4 | Quad-core ARM Cortex-A7 2, up to 4GB RAM | Selected for its affordability, but still capable enough for edge AI deployment |
| INPLAY C1080PRO USB Camera | 1080p resolution, USB interface | Selected for its USB plug-and-play compatibility and low cost and lightweight |
| SanDisk 64GB Extreme Pro | MicroSD, high-speed read/write | Selected for its high read and write speed as well as for its proven reliability. This component directly reduces system boot time and model loading time on the Raspberry Pi. The 64GB storage capacity covered the operating system, YOLO model weights, label map files, and system logs without storage constraints. |

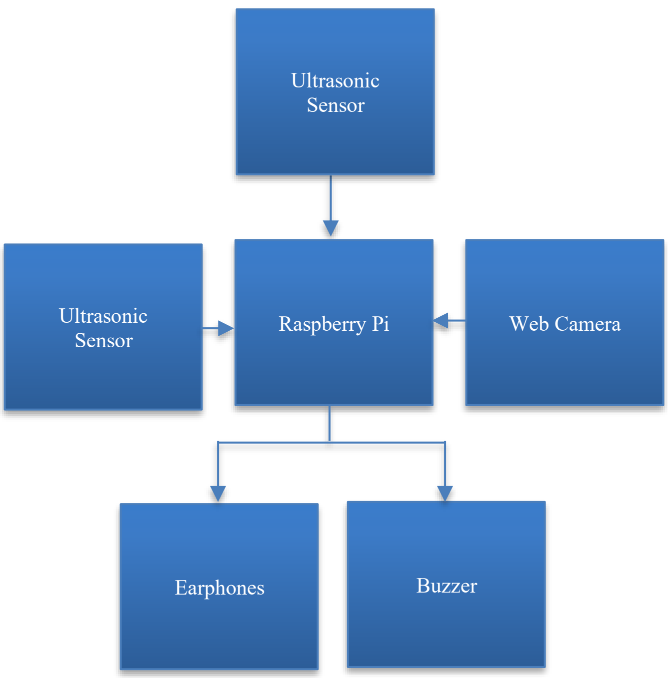


**Fig. 1 Block diagram of the wearable device**

The block diagram in Figure 1 illustrates a sophisticated system centered around a Raspberry Pi 4 as our microcontroller. Researchers use the RPI 8MP Camera Board Raspberry Pi V2 in capturing the necessary images or videos, and an Ultrasonic Sensor is integrated into the system for precise distance measurement.

The flow chart in Figure 2 demonstrates the overall operation of the system. The operation starts by initializing the system, preparing the components to employ. The system then scans frames and images and processes it to determine whether an object is detected or not. If not, it continuously scans frames and images.

If yes, the system identifies the object detected and determines their relevance. The system then triggers an output to alert the user using speech or a buzzer. The cycle will keep on repeating, allowing continuous monitoring and operation.

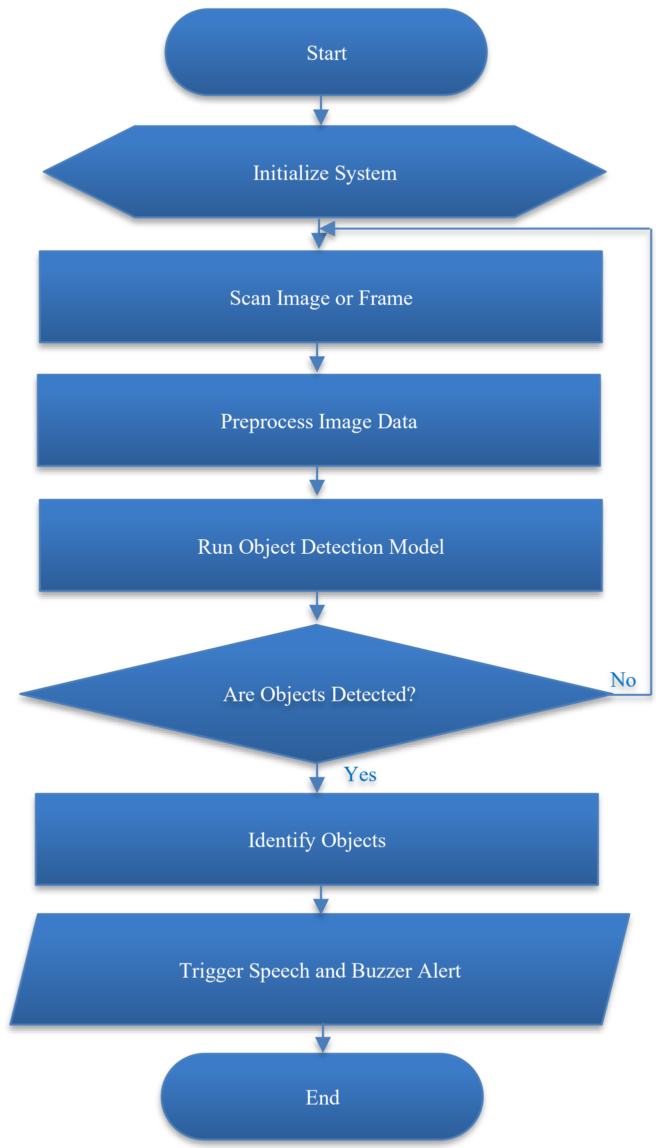


**Fig. 2 Flowchart of the device**

### *3.2. System Design*

Overall, the presented studies provide important insights into the current advances and challenges in the field of assistive technologies used by visually impaired individuals. However, there is the need for ongoing innovation in this area despite the presence of challenges associated with usability, accessibility, and cognitive workload.

Nevertheless, future advances hold many promising opportunities for visually impaired individuals due to the application of artificial intelligence, real-time feedback, and social networking. Therefore, there are many promising opportunities for the creation of efficient, affordable, and user-friendly solutions in the future, implying that visually impaired individuals can travel more independently and safely.

### 3.3. Object Detection Model and Dataset

The object detection component makes use of YOLOv8 from Ultralytics and was pre-trained on the Microsoft COCO 2017 dataset [20], which comprises 118,287 training images across 80 annotated object categories. No custom training was performed, and the pre-trained weights were deployed directly on the Raspberry Pi 4, identical with the system integration method established by Vaishnav et al. [7] and Das & Roy [8] on comparable edge hardware platforms. Out of the 80 available COCO categories, only 15 classes were retained for evaluation, which were the most relevant to the target navigation environment of visually impaired users in Philippine commercial settings.

This includes a person, bottle, chair, vehicle, a keyboard, monitor, laptop, jacket, mug, vase, bench, potted plant, phone, refrigerator, and slipper. A minimum confidence threshold of 0.5 was applied during inference, indicating that only detections scoring at or above 50% confidence were passed to the text-to-speech pipeline. This value was established to optimize detection sensitivity against false positive rates since lower values led to numerous false alerts during initial testing, while higher values resulted in frequent difficulties to detect smaller objects.

### 3.4. Research Instrument

The researchers used questionnaires and conducted interviews to collect information and feedback to evaluate the overall performance and effectiveness of the device. The questionnaire was divided into different categories that assessed the following: Basic Information of Users, Problems in Moving Around, Desired Features of the Device, Accuracy of Detecting Objects, Quality of Voice Alerts, and Overall Quality of the Device. Interviews were also conducted to allow participants to explain their answers in more detail and share their personal experiences. It also measured the overall usefulness of the device, as caregivers were included as respondents to provide recommendations and feedback regarding the quality of the device. Overall, these tools helped the researchers understand how effective the device was and identify the improvements needed.

### 3.5. Test Conditions and Experimental Protocol

#### 3.5.1. Testing Venues

Assessment of the system was conducted across four indoor commercial locations in Danao City: Gaisano Capital - Shoppers Outlet, City Mall Danao, Metro Danao, and Sands Gateway Mall. Locations were selected as visually impaired individuals were present in the area, which creates a unique setting for investigating the usage of smart wearables for the blind, promoting a more detailed evaluation for the system.

Moreover, locations also reflect more typical navigation environments usually visited by blind people, where a lot of obstacles are present. All testing was done under controlled fluorescent indoor lighting, and for that reason, limitations are considered regarding the light conditions, as the study did not cover outdoor environments or situations with direct sunlight.

#### 3.5.2. Device Testing Protocol

**Table 2. Device testing protocol**

| Test | Setup | Trials | Success Criterion |
|---|---|---|---|
| Distance Range Detection | Each of the 15 target objects positioned at a fixed desired distance of 5 meters from the device | 1 trial per object | Actual detection distance recorded at confidence score ≥ 0.5 |
| Object Detection Accura cy | All 15 objects arranged simultaneously within the camera's field of view | 15 trials of 60 seconds each | Number of correctly identified objects recorded per trial |
| Text-to -Speech Accuracy | 10 target object labels verified against system spoken output | 15 trials | Counted as correct only if the spoken word exactly matched the detected object label. |

#### 3.5.3. Test Participants

Ten participants with no prior wearable technology experience participated in the study; six reported total blindness, and four reported severe low vision. Each completed a standardized 15-minute navigation task involving three planted obstacles and two target objects. Performance data was paired with post-task structured questionnaires.

### 3.6. Metrix for Evaluation

#### 3.6.1. Distance /Range Detection Accuracy Formula

With the use of this test, the hardware was methodologically scrutinized in order to evaluate the system's capability in detecting obstacles, terrain identification, and adaptability to different factors in terms of distance accuracy through this formula:

$$\text{error} = \left(\frac{\text{Actual Distance}-\text{Desired Distance}}{\text{Desired Distance}}\right)\text{x}100 \quad (1)$$

$$Difference = Actual\,Distance - Desired\,Distance$$

*3.6.2. Object Detection and Text-to-Speech Accuracy*

This test is utilized to evaluate how the system converts written text into spoken words and how accurately it detects object underscoring the importance of pronunciation, pacing and clarity of the speech generated through this formula:

$$\%\text{error} = \left(\frac{\text{Actual Distance}-\text{Desired Distance}}{\text{Desired Distance}}\right)\text{x}100 \quad (2)$$

$$Difference = Actual\,Distance - Desired\,Distance$$

*3.6.3. Data Analysis*

T -test statistical tool was employed in the study in order to assess the difference between the real and desired measurements concerning Distance/Range Object Detection, Accuracy, and Text-to-Speech Accuracy through this formula:

$$t = \frac{d}{SD\sqrt{n}} \quad (3)$$

## 4. Results and Discussion

### *4.1. Data for Distance Range Object Detection*

As can be seen from Table 3, there is a statistically significant difference between the actual and desired detection distances. Moreover, the values obtained are on average 2.5 meters less than the desired 5-meter target. Additionally, the t-test statistic yielded t=6.73 for df=14, where $p < 0.05$. Since the t-value is higher than the critical point, ±2.14, it is imperative to deduce that there is enough evidence to refute the hypothesis.

Furthermore, detection shortfall is not uniform across object types. Large objects with broad surface areas had been detected at or near the desired 5-meter range, while small objects, including keyboards, phones, and slippers, registered actual detection distances of only 0.5 meters.

The discrepancy was caused by two factors: First is the HC-SR04 Ultrasonic Sensor's sound wave becomes feeble whenever it hits small surfaces, deliberately make it unreliable for detecting tiny objects beyond 3 meters; and second is the YOLO model's threshold score becomes less than 0.5 much faster in detecting small objects rather than large ones at an increased distance.

The results were similar to the study conducted by the group of Vaishnav. Their study reported [7] that there is a distance discrepancy for small objects utilizing comparable sensor configurations on Raspberry Pi platforms.

While reducing confidence threshold below 0.5 may elevate the small-object detection issue, it costs an increased false positive announcement.

**Table 3. Distance/Range object detection**

| Trials | Actual Distance | Desired Distance | Percent Error | Difference |
|---|---|---|---|---|
| Person 1 | 5 | 5 | 0% | -0 |
| Person 2 | 3 | 5 | 0.4% | -2 |
| Bottle | 3 | 5 | 0.4% | -2 |
| Chair | 1 | 5 | 0.8% | -4 |
| Vehicle | 4 | 5 | 0.2% | -1 |
| Keyboard | 0.5 | 5 | 0.9% | -4.5 |
| Monitor | 1 | 5 | 0.8% | -2 |
| Laptop | 3 | 5 | 0.4% | -2 |
| Jacket | 3 | 5 | 0.4% | -2 |
| Mug | 1 | 5 | 0.8% | -4 |
| Vase | 4 | 5 | 0.2% | -1 |
| Bench | 5 | 5 | 0% | -2 |
| Potted Plant | 3 | 5 | 0.4% | -2 |
| Phone | 0.5 | 5 | 0.9% | -4.5 |
| Slippers | 0.5 | 5 | 0.9% | -4.5 |
| **Mean difference** | **2.5** | | | |
| **t-statistic** | **6.73** | | | |
| **df** | **14** | | | |
| **p-value** | **< 0.05** | | | |

### *4.2. Data for Object Detection Accuracy*

The accuracy of object detections was indicated in Table 4 through 15 trials. The analysis yielded a t-test statistic value of -8.96 with a degree of freedom of 14 and a p-value < 0.001. Significantly, these values indicate that there is a difference between the actual number of objects being detected and the desired target of 15.

Moreover, the effect size of 2.31 indicates that underperformance is not caused by statistical irregularity but rather reflects reality, where only 9 out of 15 objects were successfully identified per trial.

The detection gap is aroused by two primary factors. One reason is that the YOLO model was pre-trained on the general-purpose COCO dataset without fine-tuning on locally specific object categories — items such as slippers and jackets are underrepresented in COCO training data, resulting in lower confidence scores that fall below the 0.5 detection threshold.

The second reason is that the existence of all 15 objects at the same time in one camera image resulted in a stoppage phenomenon, thereby affecting the algorithm to decide confident detections.

The findings of the study confirm the study conducted by the group of Bongao, which stated that poor detection outcomes happen when multiple object categories were to be evaluated simultaneously in a cluttered area. Improving the detection accuracy could be made by fine-tuning the YOLO model on a locally collected dataset.

**Table 4. Accuracy of object detection**

| No. of Trials | No. of Actual Objects Detected | No. of Desired Object Detected | Percent Error | Difference |
|---|---|---|---|---|
| 1 | 9 | 15 | 0.4% | -6 |
| 2 | 7 | 15 | 0.53% | -8 |
| 3 | 5 | 15 | 0.67% | -10 |
| 4 | 6 | 15 | 0.6% | -9 |
| 5 | 5 | 15 | 0.67% | -10 |
| 6 | 8 | 15 | 0.87% | -7 |
| 7 | 9 | 15 | 0.4% | -6 |
| 8 | 7 | 15 | 0.53% | -8 |
| 9 | 8 | 15 | 0.47% | -7 |
| 10 | 11 | 15 | 0.27% | -4 |
| 11 | 9 | 15 | 0.4% | -6 |
| 12 | 10 | 15 | 0.33% | -5 |
| 13 | 13 | 15 | 0.13% | -2 |
| 14 | 11 | 15 | 0.27% | -4 |
| 15 | 14 | 15 | 0.06% | -1 |
| **Mean difference** | **6.2** | | | |
| **t-statistic** | **- 8.96** | | | |
| **df** | **14** | | | |
| **p-value** | **< 0.001** | | | |

### *4.3. Data for Text-to-Speech Accuracy*

Results of the TTS were provided in Table 5, where it reveals a statistically significant difference between the spoken words and the target words because it yielded a t-test statistic of -5.13 with 14 degrees of freedom and a p-value less than 0.001. However, the error level is, on average, the system correctly pronounces 9 words out of 10 throughout the whole process. TTS performed successfully as compared to the object detection and is even better as the number of trials increases, where it could achieve a perfect 10 targets in four out of seven last tests, emphasizing stability. The limitations identified by the users were related to timing, delay between object detection and latency. In real-world assistive use, one single missed word per interaction does not impair user comprehension confirmed by 74% user satisfaction rating for TTS clarity. Moreover, the findings of the study collectively indicate the strength of the TTS component in performing elements of the proposed system as compared to object detection and distance range detection components.

**Table 5. Accuracy of Text-to-Speech**

| Trials | No. of correct words spoken | No. of desired correct words spoken | Percent error | Difference |
|---|---|---|---|---|
| 1 | 8 | 10 | 0.2% | -2 |
| 2 | 9 | 10 | 0.1% | -1 |
| 3 | 8 | 10 | 0.3% | -2 |
| 4 | 9 | 10 | 0.1% | -1 |
| 5 | 9 | 10 | 0.1% | -1 |
| 6 | 8 | 10 | 0.2% | -2 |
| 7 | 9 | 10 | 0.1% | -1 |

| | | | | |
|---|---|---|---|---|
| 8 | 8 | 10 | 0.2% | -2 |
| 9 | 10 | 10 | 0% | 0 |
| 10 | 9 | 10 | 0.1% | -1 |
| 11 | 10 | 10 | 0% | 0 |
| 12 | 9 | 10 | 0.1% | -1 |
| 13 | 9 | 10 | 0.1% | -1 |
| 14 | 10 | 10 | 0% | 0 |
| 15 | 10 | 10 | 0% | 0 |
| **Mean difference** | **1** | | | |
| **t-statistic** | **-5.13** | | | |
| **df** | **14** | | | |
| **p-value** | **< 0.001** | | | |

### *4.4. Instrument Results*

The hardest problem task was identifying objects and people (85%), followed by avoiding obstacles (78%), crossing streets (68%) last as shown in Figure 3.

These results confirm the basic design choices of the proposed system: the use of YOLO-based object recognition directly addresses the main problem reported by participants, while the use of an ultrasonic proximity sensor addresses the second most important problem of avoiding obstacles.

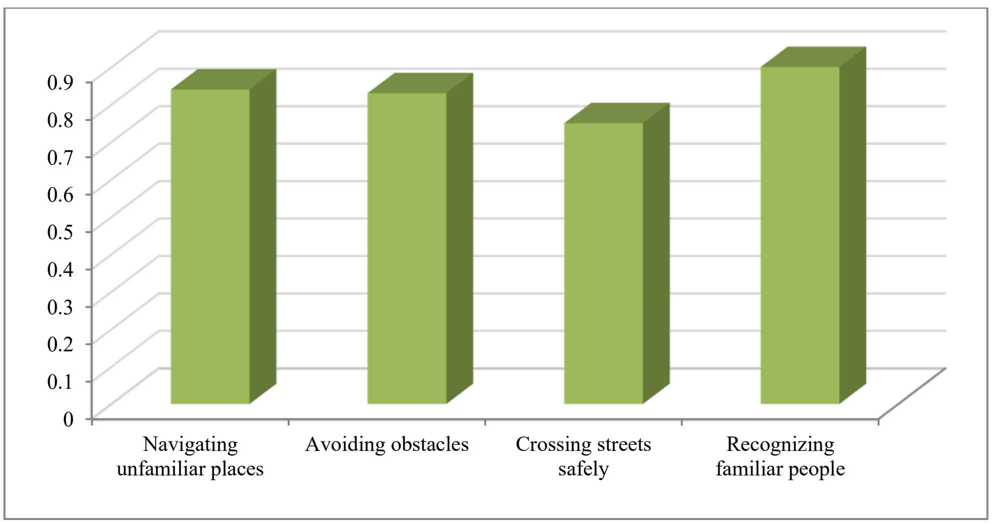


**Fig. 3 Challenges faced by visually impaired individuals**

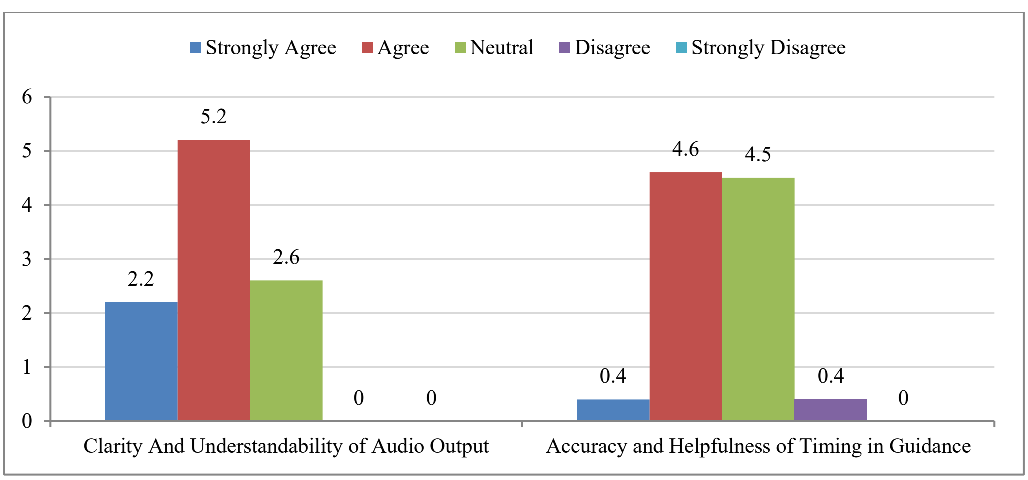


**Fig. 4 Performance of the smart wearable device in terms of object detection**

Figure 4 illustrates participant feedback on the device's object detection efficacy. While sensor accuracy and timely alerts were well received due to the effective dual-alert ultrasonic and buzzer system, feedback on detection distances was highly critical, reflecting a 2.5-meter mean shortfall shown in Table 3. Similarly, most participants expressed high dissatisfaction regarding performance in spatially constrained environments, likely caused by the narrow beam angle of the HC-SR04 sensor and its restricted detection coverage in confined spaces.

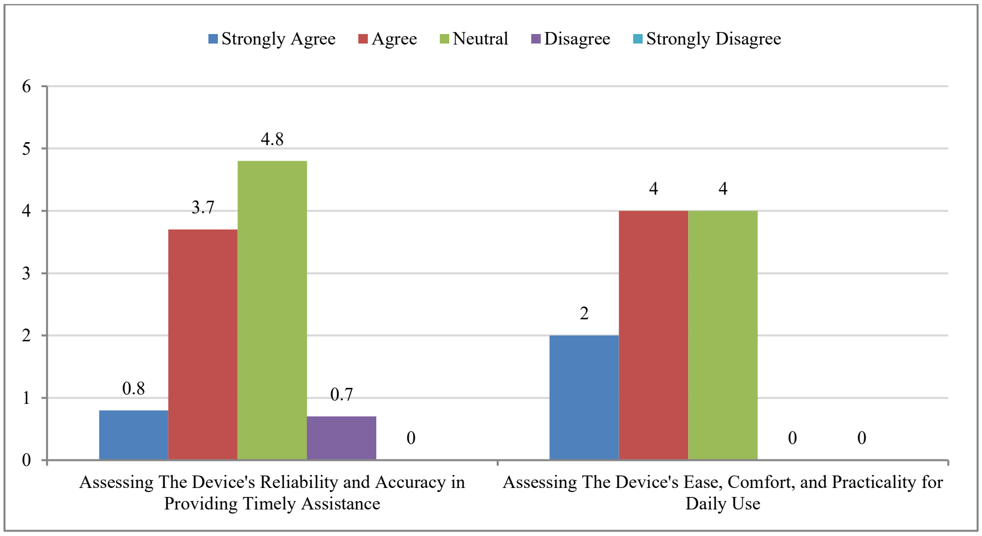


**Fig. 5 Text-to-Speech functionality effectiveness**

The graph encapsulates the user’s feedback with regard to the functionality and effectiveness of text-to-speech, focusing on two aspects: (1) clarity and understandability of audio output, and (2) accuracy and helpfulness of timing in guidance. Moreover, it indicates that text-to-speech functionality has made two generalizations: (1) effective in terms of clarity, while (2) the timing requires improvement. As for the clarity and understandability of audio output, the majority of users or 52% ‘, agreed, indicating that the audio output of the study is generally clear and understandable. Followed by 26% of which were ‘neutral’, and the remaining 22% voted for ‘strongly agreed.’

Overall, there was no negative reception and dissatisfaction in this area. Furthermore, in terms of Accuracy and Helpfulness of Timing in Guidance, 46% of the users voted for ‘agree,’ followed by 45% who chose neutral and 40% who chose ‘strongly agree’ while only 4% chose to disagree. Although this has a slightly greater variation in perception compared to audio clarity, it has gained a prevalent positive outlook, reflecting that the functionality of the study is deemed effective by the users. Finally, the graph indicates that the Test-to-Speech (TTS) Functionality is generally effective, especially in terms of clarity, while the timing may require some tweaks for improvement.

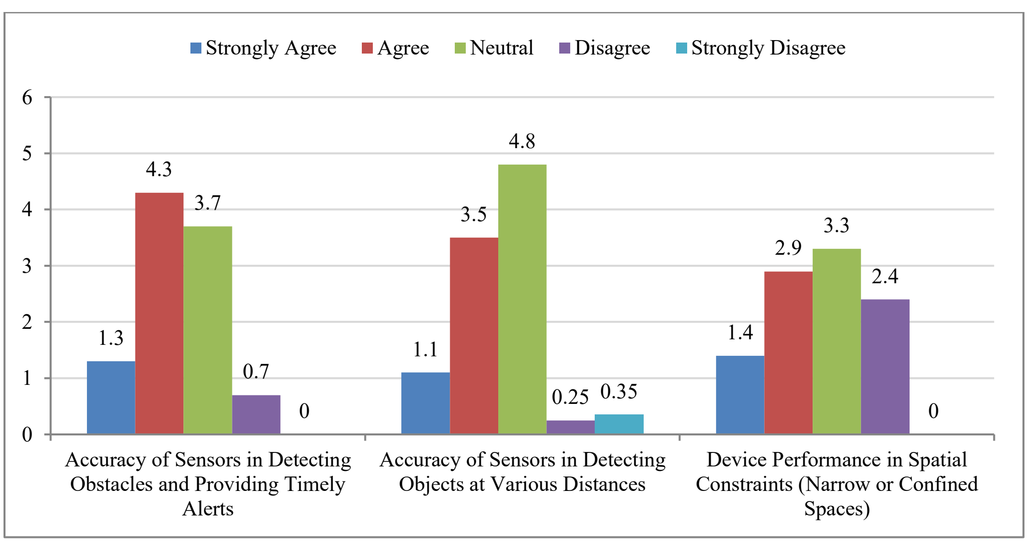


**Fig. 6 Assessing the helpfulness of the smart wearable device for visually impaired individuals**

The graph captures the evaluation on the functionality of the device involving two main criteria: the device's (1) reliability and accuracy in providing timely assistance, and (2) ease, comfort, and practicality for daily usage. The data shows that most of the respondents deemed this device as cautious (neutral score); it still has a favorable opinion with regard to its performance. The result of this paper indicates that the system provides effective support. The integration of object detection and text-to-speech (TTS) technologies has helped researchers establish a good result in enhancing users' spatial awareness and safety.

## 5. Conclusion

The result of this paper indicated that the system provides effective support to visually impaired individuals, especially in overcoming difficulties related to object recognition and navigation. The integration of object detection and Text-to-Speech (TTS) technologies has helped researchers establish a good result in enhancing users' spatial awareness and safety by providing timely and informative alerts.

Despite the system's good performance, several imperfections have been shown, including its incapability to detect objects at varying distances as well as its operation in confined or cluttered spaces. Nonetheless, implementing the system in an actual setting would surely contribute to the improvement of society, reducing accident rates and increasing users' self-reliance in navigating public areas. Overall, the system highlighted the device's potential as an assistive tool for impaired individuals to enhance their independence and mobility.

### *5.1. Limitations*

The study findings and their generalizability are limited by several constraints. The YOLO model is deployed as the first approach, which uses COCO pre-trained weights without category fine-tuning for more localised objects. The COCO training dataset directly affects the low detection rates observed for certain categories, such as slippers, which are found in many places native to the Philippines. Second, all tests were performed in 4 commercial locations of Danao City using a controlled fluorescent light setting. The device has limited performance in high-glare and low-light environments. Third, the HC-SR04 ultrasonic sensor creates lateral blind spots due to its narrow beam angle, as reflected in the user dissatisfaction reported for confined space performance. Fourth, system latency concerns on fast paced application. Fifth, the small sample size of participants.

Finally, percent error was applied as the primary evaluation metric in place of standard information retrieval metrics such as precision, recall, and F1-score, which represents a methodological limitation acknowledged for future work.

### *5.2. Future Works*

The future works about the proposed system would be undertaken to surpass the limitations mentioned in this study and develop better accuracy, responsiveness and usability when used under practical conditions. This includes:

- Model Improvement - Fine-tune the YOLO model for improving detection accuracy on objects that are common in the Philippine environment using a locally compiled dataset.
- Expanded Detection Range - Adding or using an additional sensor, or a more depth camera with wider vision in comparison to the HC-SR04 sensor to extend the device's detection range.
- Extended Testing Condition - better testing condition by evaluating the device in an outdoor environment or during lowlight, and also with different weather conditions to enhance the strength of the system.
- Multilingual Support - enhance text-to-speech capabilities by including Filipino and Cebuano language to help support local users better.
- User study expansion - Conduct studies beyond different types of visually impaired participants and by increasing sample size to assess real-world effectiveness.

## Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

## Funding Statement

This research received no external Funding.

## Acknowledgement

The researchers extend their heartfelt gratitude to all of those who contributed to the completion of this capstone project. Primarily, the researchers extend their warm and heartfelt appreciation and gratitude to the Almighty Father for his unceasing blessings and guidance he bestows upon us. The researchers also wish to acknowledge the guidance and support of their research adviser, Engr. Chona R. Dagatan.

## References

[1] Sonja Alimović, "Benefits and Challenges of using Assistive Technology in the Education and Rehabilitation of Individuals with Visual Impairments," *Disability and Rehabilitation: Assistive Technology*, vol. 19, no. 8, pp. 3063-3070, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Envision - Enabling Vision for Visually Impaired, Envision, 2026. [Online]. Available: https://www.letsenvision.com/

[3] Eye, *Eye News*, Eye, 2026. [Online]. Available: https://www.eyenews.uk.com/

[4] Verbit Editorial, Advances in AI are Promoting Greater Accessibility, Verbit.ai, 2024. [Online]. Available: https://verbit.ai/accessibility-hub/advances-in-ai-are-promoting-greater-accessibility/
[5] Melchiezedhieck J. Bongao et al., “SBC-based Object and Text Recognition Wearable System using Convolutional Neural Network with Deep Learning Algorithm,” *International Journal of Recent Technology and Engineering*, vol. 10, no. 3, pp. 198-205, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[6] Junxin Chen, Wei Wang, and Gwanggil Jeon, “Real-Time Machine Learning Based Object Detection and Recognition System for the Visually Impaired,” *Proceedings of the 2023 Workshop on Advanced Multimedia Computing for Smart Manufacturing and Engineering*, Association for Computing Machinery, New York, United States, pp. 31-35, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[7] Devashree Vaishnav, B. Rama Rao, and Dattatray Bade, “Wearable Assistance Device for the Visually Impaired,” *Advances in Machine Learning and Computational Intelligence: Proceedings of ICMLCI*, Springer, Singapore, pp. 667-676, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[8] Dibyadarsan Das, and Sritama Roy, “Object Detection with Voice Output for Visually Impaired,” *2024 International Conference on Communication, Computing and Internet of Things (*IC3IoT), Chennai, India, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[9] Reddy Sushma Sree et al., “Empowering Independence: Raspberry Pi OCR for Visually Impaired User,” *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, India, pp. 1711-1716, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[10] Aryan Singh et al., “Design and Implementation of Text to Speech Synthesizer,” *International Journal of Futuristic Innovation in Engineering, Science and Technology (IJFIEST)*, vol. 1, no. 1, pp. 10-12, 2022. [Google Scholar]
[11] Matshehla Konaite et al., “Smart Hat for the Blind with Real-Time Object Detection using Raspberry Pi and Tensorflow Lite,” *Proceedings of the International Conference on Artificial Intelligence and its Applications*, Association for Computing Machinery, New York, United States, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[12] Abhijit Pathak et al., “An IoT based Voice Controlled Blind Stick to Guide Blind People,” *International Journal of Engineering Inventions*, vol. 9, no. 1, pp. 9-14, 2020. [Google Scholar]
[13] Salam Dhou et al., “An IoT Machine Learning-based Mobile Sensors Unit for Visually Impaired People,” *Sensors*, vol. 22, no. 14, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[14] Reuben O Jacob et al., “IoT based GPS Tracking System with SOS Capabilities,” *2022 International Mobile and Embedded Technology Conference (MECON)*, Noida, India, pp. 72-75, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[15] Shubham Sumaet al., “Vision Navigator: A Smart and Intelligent Obstacle Recognition Model for Visually Impaired Users,” *Mobile Information Systems*, vol. 2022, no. 1, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]