

Raspberry Pi-Based Text-to-Speech Reader and Object Detection System for the Visually Impaired

Ashwini Metkar¹, Prathamesh Ligade², Sanika Magar³, Ganesh More⁴, Prof. Sonal Lad⁴

Department of Electronics and Telecommunication Engineering

Kolhapur Institute of Technology's College of Engineering, Kolhapur

Abstract - This paper presents a Raspberry Pi-based assistive system designed to enhance independence and situational awareness for visually impaired individuals. The suggested system combines the ability to read text in real time with the ability to find objects in order to give continuous auditory feedback about the environment. An Optical Intelligent Character Recognition (iOCR) module is used to turn printed and handwritten text into speech, making it easier for users to read documents and text in their surroundings. At the same time, an object detection module based on YOLOv8 is set up to find and sort objects in real time. A text-to-speech interface sends the detected objects and recognized text to the user. The system works well on a cheap, portable hardware platform and shows good accuracy and processing speed during testing. The findings suggest that the proposed solution possesses significant potential for practical implementation as an economical assistive technology for visually impaired individuals.

Key Words: component, formatting, style, styling, insert

1. INTRODUCTION

A person with a visual impairment has a lot of trouble doing everyday things like reading, finding things, and getting around safely in new places. Global health reports say that millions of people around the world have partial or complete vision loss. This means that there is a strong need for assistive technologies that are both cheap and reliable. White canes and guide dogs are examples of traditional assistive tools that help people move around, but they don't give them access to written information or a better understanding of their surroundings. Recent progress in embedded systems, artificial intelligence, and computer vision has made it possible to make smart assistive devices that can help blind people live more independently and improve their quality of life. Thanks to the rise of cheap embedded computing platforms like Raspberry Pi,

Identify applicable funding agency here. If none, delete this.

it is now possible to make portable and efficient assistive systems that can handle complicated real-time processing

tasks. Raspberry Pi has enough processing power, is small enough, and uses little power, making it a good choice for wearable and portable assistive applications. Such platforms can support real-time environmental perception and user interaction by combining machine learning algorithms with sensor-based data collection. This study suggests a Raspberry Pi-based assistive system that incorporates two fundamental features: real-time text reading and object detection. The text reading part uses Intelligent Optical Character Recognition (iOCR) captures printed or handwritten text and turns it into speech output. This lets people who can't see read books, documents, and public signs. At the same time, the object detection module uses a pre-trained deep learning model to find objects nearby and give audio feedback about the user surroundings. This dual function makes users more aware of their surroundings and keeps them safer while they are navigating and doing everyday things. The proposed system is meant to be affordable, easy to use, and portable, which makes it useful for everyday use. The system that was created is meant to be an affordable assistive solution that helps visually impaired people become more independent and mobile. People with disabilities.

2. LITERATURE SURVEY

The review of related literature for a Raspberry Pi-based text reader and object detection system covers multiple research domains, including embedded assistive technologies, Optical Character Recognition (OCR), deep learning-based object detection, and human-computer interaction for visually impaired users. These research areas collectively contribute to the development of intelligent assistive systems that enhance accessibility, safety, and independence for visually impaired individuals. Research on embedded platforms highlights the growing adoption of low-cost computing solutions such as Raspberry Pi in assistive technology applications. Studies conducted by Molleman et al. (2017) demonstrate that Raspberry Pi has been widely used across healthcare, automation, and assistive device development due to its compact size, low power consumption, flexibility, and affordability. The platform enables developers to integrate sensors, cameras, and audio output devices into a single system, making it highly suitable for portable assistive

applications. Researchers have also emphasized its capability to support real-time image processing and machine learning tasks, which are essential for text recognition and object detection systems. Significant advancements have been observed in Optical Character Recognition (OCR) technologies over the past decade. Studies by Liang et al. (2018) and Luo et al. (2020) focus on improving OCR accuracy, processing speed, and robustness under varying lighting and background conditions. Modern OCR systems use machine learning and deep learning approaches to recognize both printed and handwritten text. Recent research also emphasizes implementing OCR on edge devices rather than cloud-based systems. Edge-based OCR improves privacy, reduces latency, and enables real-time processing, which is critical for assistive devices used in dynamic real-world environments. Object detection research has progressed rapidly with the emergence of deep learning and Convolutional Neural Networks (CNNs). Research studies by Redmon et al. (2018) and Liu et al. (2016) demonstrate the effectiveness of deep learning-based object detection models in achieving high detection accuracy and real-time performance. These studies discuss the development of lightweight yet powerful detection models capable of running on embedded systems. Such models are essential for assistive applications where quick and accurate environmental understanding is required to ensure user safety. In addition to technical performance, user experience and usability play a critical role in assistive technology adoption. Research by Andre et al. (2018) and Brewster et al. (2019) highlights the importance of Human-Computer Interaction (HCI) design principles in assistive systems. These studies emphasize the need for intuitive audio interfaces, minimal user input requirements, and customizable feedback systems to match individual user preferences. Proper interface design improves usability, reduces cognitive load, and increases user acceptance of assistive technologies. Recent research trends also focus on integrating multiple assistive functionalities into a single compact device. Many studies suggest that combining text recognition and object detection into a unified platform significantly improves usability compared to using separate devices. Integrated systems reduce hardware cost, improve portability, and provide a seamless user experience. Additionally, advancements in Text-to-Speech (TTS) systems have improved speech naturalness and clarity, making audio feedback more understandable and comfortable for users. Furthermore, researchers are exploring the use of artificial intelligence to enhance assistive system adaptability. Adaptive learning models can improve recognition accuracy based on user environment and usage patterns. Future research directions include integrating navigation assistance, obstacle detection, and scene description capabilities into assistive devices, creating a more comprehensive support system for visually impaired individuals. Overall, the literature indicates strong progress in embedded computing, OCR, and deep learning-based object detection technologies. However, there

remains a need for cost-effective, portable, and integrated assistive systems that can operate reliably in real-time conditions. The proposed system aims to address these gaps by combining real-time text reading and object detection using an embedded platform, thereby providing an affordable and practical assistive solution for visually impaired users.

3. PROBLEM STATEMENT

To design and develop an intelligent object detection and text reading system using Raspberry Pi, a camera module, and audio output. The system should be able to recognize commonly encountered objects, extract text from images using IOCR, and convey the information audibly to assist blind or visually impaired individuals in navigating and interacting with their environment more independently.

4. METHODOLOGY

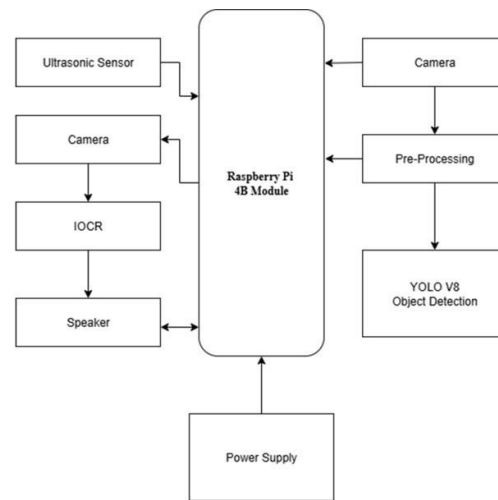


Fig. 1. block diagram

The proposed system, “Raspberry Pi-Based Text-to-Speech Reader and Object Detection System for the Visually Impaired,” aims to improve the autonomy and safety of visually impaired individuals by delivering real-time auditory information regarding their environment. The Raspberry Pi 4B module is the main part of the system. It is the central processing unit that handles image processing, sensor data, and audio output. A processing unit is connected to a high resolution camera that takes pictures of the environment in real time. These pictures are used for both recognizing text and finding objects. The system uses Intelligent Optical Character Recognition (IOCR) to read text. It can read printed or handwritten text from images taken in different lighting conditions and with different backgrounds. different kinds of light and busy backgrounds. A Text-to-Speech (TTS) engine processes the extracted text and turns it into natural speech.

This lets people listen to written information from books, documents, labels, and public signs in real time. This feature helps people who are blind or have low vision read text on their own, without help from others. For environmental awareness, the system integrates a YOLOv8-based object detection model developed by Ultralytics. In object detection mode, captured images undergo preprocessing operations such as noise removal, resizing, and contrast enhancement before being passed to the deep learning model. The YOLOv8 model then identifies and classifies objects present in the scene in real time. The detected object names are converted into audio output and delivered to the user, helping them understand their surroundings and avoid potential hazards. Additionally, an ultrasonic sensor is incorporated into the system to enhance navigation safety. The sensor continuously measures the distance between the user and nearby obstacles. When an object is detected within a predefined threshold distance, the system generates an immediate audio warning, helping prevent collisions and improving safe movement. The overall system is designed to be low-cost, portable, and energy-efficient, making it suitable for daily real-world usage.

5. PYTHON SIMULATION

The output of the Python Text-to-Speech (TTS) simulation code will be both text printed to the console and spoken audio from your computer’s speakers

note:only English text

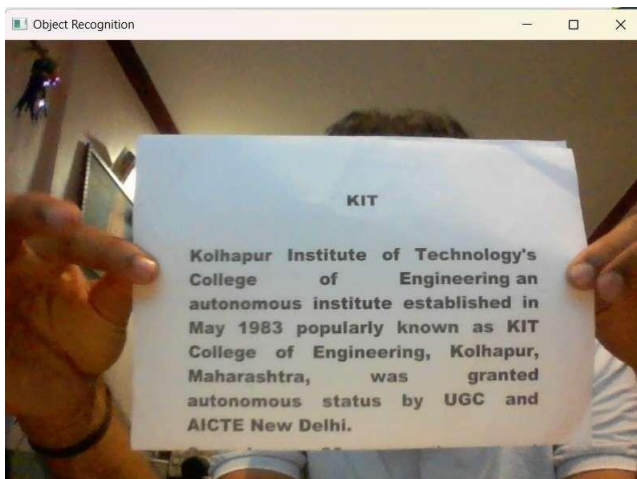


Fig. 2. text to speech simulation

The output of the python Object Detection Python simulation using OpenCV for e.g. : Mouse



Fig. 3. object detection simulation

6. WORKING OF THE PROJECT

The suggested assistive system works by using a high-resolution camera connected to the Raspberry Pi 4B to constantly take pictures of the user’s surroundings in real time part The Raspberry Pi is the main controller. It controls the acquisition of images, the monitoring of sensors, the processing of data, and the generation of audio output. The system is meant to work in real time and can automatically switch between different modes of operation based on user commands or preset system settings. The camera module starts the working process by taking pictures of the environment all the time. These pictures are sent right away to the processing unit for analysis. The captured images are processed using either the Text Reading Mode or the Object Detection Mode, depending on the chosen operation mode. This makes the system flexible and adaptable to different real-world situations.

A. text to speech

The Intelligent Optical Character Recognition (IOCR) module processes the images that were taken in Text Reading Mode. The IOCR module looks at pictures taken in different lighting conditions. Lighting and complicated backgrounds make it hard to get printed or handwritten text right. The text that was taken out is then turned into a digital format and sent to the Text-to-Speech (TTS) engine. The TTS engine makes speech that sounds natural and clear, and it sends it to the user through a speaker or headphones. This lets people who can’t see read books, documents, medicine labels, product packaging, and public signs without help.

B. Object detection

In Object Detection Mode, the images that are taken go through preprocessing steps like resizing, noise reduction,

contrast enhancement, and normalization before they are saved. These preprocessing steps make detection more accurate and lighten the load on processing. Ultralytics made the YOLOv8 deep learning model, which is used to look at the processed images. The model finds many things in the scene and sorts them in real time. The names of the detected objects are turned into audio messages and sent to the user. This helps them understand what's around them and find important objects or possible obstacles. At the same time, the ultrasonic sensor keeps an eye on the distance between the user and any nearby obstacles. The sensor sends out ultrasonic waves and measures the signals that bounce back to figure out how far away the object is. If the system sees something within a set safety range, it immediately makes an audio warning sound. This feature makes things safer and helps people feel more sure of themselves when they are in crowded or new places. The system works in a loop that never ends. It takes pictures, processes them, detects or recognizes them, and gives audio feedback. This system for giving feedback in real time makes sure that users get accurate and timely information about their surroundings. Combining several assistive features into a single embedded platform makes the hardware less complicated while keeping performance high

7. FUTURE SCOPE

The proposed assistive system shows a lot of promise for helping people who are blind or have low vision by combining text reading and object detection. In future versions, improvements and technological upgrades can be made to make the program work better, easier to use, and more useful. The following areas show where the system could go in the future.

1) Recognizing text in more than one language

Right now, the system may only support one language for recognizing text. The Intelligent Optical will be able to The Character Recognition (IOCR) module can be improved to work with more than one language. This will let people from different language groups read written information in their own language. Combining advanced language models can also make it easier to read regional scripts and handwritten text.

2) Advanced Detection of Obstacles

The current ultrasonic sensor can be integrated with supplementary sensing technologies, including infrared sensors and LiDAR systems. or cameras with stereo vision. These extra sensors can make it easier to find obstacles, especially in dimly lit areas, when objects are clear, or when there are a lot of people around. You can use sensor fusion techniques to make navigation help more reliable.

3) AI processing in the cloud or at the edge

The current system processes data locally, but future versions could use cloud computing or edge AI devices. to do tasks that require a lot of computing power, like advanced object

detection and complex text recognition. You can speed up your system, lower latency, and improve its overall performance by offloading heavy processing tasks. This can be done while keeping hardware requirements low.

4) Managing batteries and optimizing energy

Future research may concentrate on enhancing power consumption through efficient processing methodologies and intelligent power solutions. circuits for managing. This will make the battery last longer and make the system easier to use all the time.

5) Better Assistive Features

The system can be expanded to include more helpful features, like facial recognition for recognizing people you know. Recognizing people, recognizing money for financial transactions, reading QR codes to get to digital information, and helping people find their way around indoors using beacon-based positioning. These features can turn the system into a full-fledged help platform.

8. FINAL RESULT

A. Text to Speech

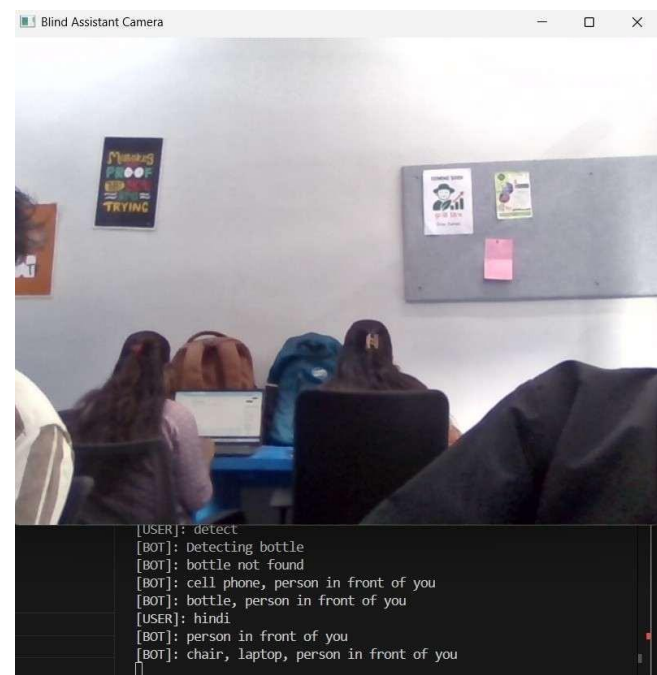


Fig. 4. text to speech simulation

The experimental findings validate the efficacy of the proposed assistive system in real-time scenarios. Recognition happens at the same time, and the information that is found is turned into audio feedback. The system can correctly identify things like people, cell phones, and accuracy.

B. object detection

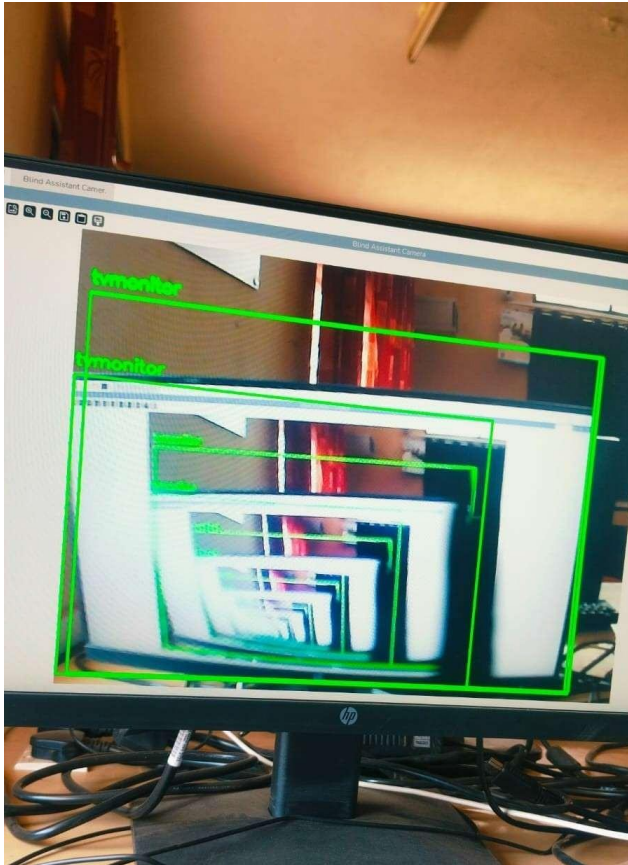


Fig. 5. object detection simulation

illustrates the real-time implementation of the YOLOv8 object detection model, where multiple objects are detected in a dynamic environment. The output is displayed along with corresponding labels, and audio feedback is generated for user assistance. These results confirm that the system provides reliable performance under practical conditions.

ACKNOWLEDGEMENT

We would like to express our deep gratitude to Prof. S. R. Lad Dr. M. D. Sontakke, Head of Department for their constant encouragement and belief in us. Their guidance and attention throughout the project work have been of immense help to us. We express our sincere thanks to all the teaching and non-teaching staff and all those who have directly or indirectly helped in making the project a success. We also thank the administration of KIT's College of Engineering Kolhapur for providing a conducive learning environment and fostering a culture of innovation and excellence that enabled us to pursue and successfully complete this project.

REFERENCES

[1] S. Koley and r. Mishra voice operated outdoor navigation system for visually impaired persons international journal of engineering trends and technology, vol.3, issue 2,2012.

- [2] S.Dhambre and a. Sakare smart stick for blind: obstacle detection, artificial vision and real-time assistance via gps,"2nd international conference on information and communication technology (necit)2011 .
- [3] Ajantha devi and santhosh baboo's "embedded optical character recognition on tamil text image using raspberry pi", international journal of computer science trends and technology (ijest), volume. 2, issue 4, july august 2014
- [4] Vasanthi. G and ramesh babu y india. Vision based assistive system for label detection with voice output, department of ece, dmi college of engineering, chennai, jan 2014.
- [5] . Vikram shirol, abhijit m, savitri a et al. "drashti an android reading aid" international journal of computer science and information technologies vol.6 (july 2015)
- [6] Raspberry pi os <https://www.raspberrypi.com/software/operating-systems/>