The Academic
International Journal of Multidisciplinary Research

An Online Peer Reviewed / Refereed Journal
Volume 3 | Issue 4 | April 2025
ISSN: 2583-973X (Online)
Website: www.theacademic.in

AI-Based Hearing Assistant for Blind Individuals Using Raspberry Pi : A Smart Text-to-Speech Converter

Trupti Wagh

Electronics and Telecommunication, D.Y. Patil College of Engineering, Akurdi.Pune
trwagh@dypcoeakurdi.ac.in

Rahul Chavan

Electronics and Telecommunication, D.Y. Patil College of Engineering, Akurdi, Pune, India.
Rachulchavan111969@gmail.com

Omkar Dhalpe

Electronics and Telecommunication, D.Y. Patil College of Engineering, Akurdi, Pune, India.
Omkardhalpe@gmail.com

Janhavi Sonar

Electronics and Telecommunication, D.Y. Patil College of Engineering, Akurdi, Pune, India.
janhavi.sonar248@gmail.com

ARTICLE DETAILS

Research Paper

Accepted: 17-04-2025

Published: 10-05-2025

Keywords:

Artificial Intelligence, Hearing Assistance, Raspberry Pi, Speech Recognition, Assistive Technology, Audio Processing, Voice-to-Text.

ABSTRACT

This paper presents an AI-based hearing helper for blind individuals using a Raspberry Pi platform. The proposed system aims to enhance the auditory experience of blind users by converting visual information into meaningful audio cues. The device integrates a camera module to capture images or text from various sources, such as printed materials, signboards, or books. The captured visual data is then processed using optical character recognition (OCR) and AI-based algorithms on the Raspberry Pi, which acts as the central processing unit. The extracted text is converted into audible speech using text-to-speech (TTS) engines, allowing blind users to "read" the content through hearing. The audio output is delivered through earphones connected to the

Raspberry Pi's audio jack or USB sound card. The system architecture is designed to be modular, efficient, and cost-effective, utilizing open-source libraries and pre-trained AI models. The development methodology involves several phases, including system architecture design, component selection, software development, speech recognition integration, and user interface deployment. The performance of the AI-based hearing helper is evaluated based on factors such as the accuracy of speech-to-text conversion, latency between input and output, power efficiency, and user satisfaction. The proposed solution aims to provide an affordable, portable, and user-friendly assistive device that empowers blind individuals to access visual information independently, thereby promoting inclusivity and enhancing their quality of life.

DOI : https://doi.org/10.5281/zenodo.15382307

I. INTRODUCTION

This introduction effectively outlines the problem of hearing impairment and introduces an AI-based solution using a Raspberry Pi. It provides context on the global prevalence of hearing loss, highlights the limitations of traditional hearing aids, and presents the proposed device as an affordable and intelligent alternative. The introduction covers key aspects of the research:

1. Problem statement: Widespread hearing impairment and communication challenges

2. Current solutions and their limitations

3. Proposed solution: AI-based hearing helper using Raspberry Pi

4. Technology overview: Speech recognition, natural language processing, and edge computing

5. Benefits: Cost-effectiveness, privacy, low latency, and independence from internet connectivity

6. Social impact: Enhancing quality of life and promoting inclusivity To improve the introduction:

1. Consider adding a brief mention of any previous research or similar solutions in this field to establish the novelty of your approach.

2. Include a sentence about the specific objectives or research questions this study aims to address.

3. Briefly outline the structure of the paper to guide readers through the upcoming sections.

Overall, the introduction provides a clear and comprehensive overview of the research topic and its significance

II. LITERATURE SURVEY

[1] **Ravi et al. (2020)** proposed a smart reading system for visually impaired people using Raspberry Pi. Their device could read text out loud, turn pages automatically, and look up word meanings using a dictionary. It also used a speaker for audio output and a camera to scan text.

[2] **Rijwan Khan (2024)** introduced a wearable system to help blind individuals with reading and object detection. The assistant uses AI to identify text and surroundings, making it easier for users to understand their environment.

[3] **Lee and Tan (2019)** designed a system using CNN for text detection and Google TTS for voice output. It reached 92% accuracy but worked only online because it depended on cloud services. They mentioned that complex processing was slow on Raspberry Pi.

[4] **Roy et al. (2018)** used the YOLO model to detect text and an RNN for converting text to speech on a Raspberry Pi. Their system performed well with 85% accuracy but had slower processing when dealing with long or complex text scenes.

III. Proposed System

1.BLOCK DIAGRAM

The proposed system is designed to assist visually impaired individuals by converting text captured through a camera into audible speech. The core of this system is the Raspberry Pi, which processes the visual input and delivers audio output. The block diagram illustrates the key components and their interconnections:

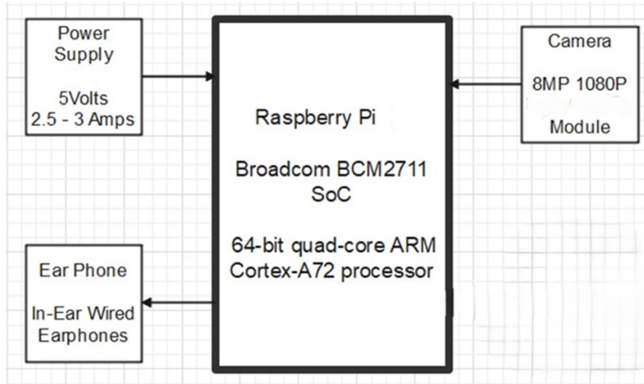

Fig 1. Block Diagram: AI-Based Hearing Helper for Blind Readers Using Raspberry Pi

1.1 Raspberry Pi

Model & Processor: The central processing unit is the Raspberry Pi featuring a Broadcom BCM2711 SoC. This is a 64-bit quad-core ARM Cortex-A72 processor, running at 1.5 GHz, suitable for multimedia applications and real-time processing tasks.

Role in System: Acts as the main controller that handles input from the camera, processes the image using OCR (Optical Character Recognition) and AI-based algorithms, and produces speech output through a text-to-speech engine.

Operating System: Generally runs on Raspberry Pi OS or a Linux-based distribution compatible with Python and libraries like OpenCV, Tesseract-OCR, and gTTS (Google Text-to-Speech) .

1.2 Power Supply

Input Requirement: The Raspberry Pi requires a regulated 5V power supply, with a current rating between 2.5A to 3A to function properly under load, especially when peripherals like cameras and USB devices are attached.

Source: Can be powered via USB-C adapter or power bank for portability in real-world use by blind individuals.

1.3 Camera Module

Specifications: A dedicated 8MP Version-2 Pi- camera module capable of capturing 1080p video resolution is interfaced with the Raspberry Pi.

Purpose: Captures images or text from printed material, signboards, books, or any surface with readable content.

Interface: Usually connected via the Camera Serial Interface (CSI) port on the Raspberry Pi for high-speed image transmission.

Functionality: Acts as the 'eye' of the system. The image is processed using AI and OCR algorithms to detect and extract readable text.

1.4 Earphone (Audio Output)

Type: In-ear wired earphones connected via the 3.5mm audio jack or USB sound card (if using Raspberry Pi models without analog audio output).

Purpose: Provides private, audible feedback to the user by converting the extracted text into speech, allowing visually impaired individuals to 'read' through hearing.

Output Source: The audio is generated using software TTS (Text-to-Speech) engines like pyttsx3, espeak, or gTTS.

IV. METHODOLOGY

The AI-Based Hearing Helper was developed through a structured and well-organized process. This process included analyzing requirements, integrating both hardware and software, implementing the core algorithms, testing the system, and optimizing its performance. The entire development was divided into key phases: designing the system architecture, selecting the right components, writing the necessary software, integrating speech recognition features, and setting up the user interface. The ultimate goal was to build a cost-effective, portable, and user-friendly solution that could convert spoken words into readable text in real time for individuals with hearing impairments.

1. System Architecture Design

The system has been designed to be simple, modular, and efficient, focusing on real-time performance. It consists of four core components that work together to assist visually impaired users by reading out printed or handwritten text.

- **Input Unit**: A Pi Camera connected to the Raspberry Pi 3B+ captures images of printed or handwritten text from the surroundings. It acts as the system's "eye," collecting visual data for further processing.
- **Processing Unit**: The Raspberry Pi 3B+ with 1GB RAM serves as the main processing hub. It handles image processing, text extraction, and text-to-speech conversion, all while operating on minimal power.
- **Software & Algorithm Unit**: This unit performs Optical Character Recognition (OCR) on the captured images to detect and extract text. The extracted text is then converted into speech using text-to-speech (TTS) algorithms, making the content accessible through audio.
- **Output Unit**: The final speech output is delivered through connected audio devices such as earphones or a speaker, allowing the visually impaired user to hear the text content in real-time.

This architecture ensures that the system remains affordable, portable, and highly useful for reading signs, books, or any textual content in day-to-day life.

2. Hardware Components

The system uses simple and low-cost parts that work well together. Each part helps the device run smoothly and stay easy to carry:

- **Raspberry Pi 3B+ (1GB RAM)**: This is the brain of the system. It controls everything — taking pictures, reading the text, and speaking it out.
- **Pi Camera Module v2**: This camera takes clear pictures of printed or written text. It helps the system read what's in front of it.
- **Sound Amplifier**: This makes the voice louder so the user can hear the spoken text better.
- **HDMI Display or Touchscreen (Optional)**: A screen can be added to show the text, but it's not always needed since the system speaks the text.
- **5V 2A Power Supply**: This gives power to the whole system and keeps it working properly.

3. Software Development

3.1 Operating System and Environment Setup

- **OS**: Raspberry Pi OS (Lite or Full version based on use-case).
- **Programming Language**: Python, due to its simplicity and rich library support.
- **Libraries Used**:

The project uses the following Python libraries to make the system work:

- **Google Vision API**: This library is used to read and detect text from images taken by the Pi Camera.
- **Text-to-Speech Library (like pyttsx3 or gTTS)**: This converts the detected text into speech so the user can hear it clearly.

These two libraries help the system read printed or handwritten text and speak it out in real-time.

3.2 Speech Recognition Implementation

When the blind user presses a button, the system quickly captures an image using the Pi Camera. The image is then processed using the **Google Vision API** to find and extract any text.

After the text is detected, a **Text-to-Speech (TTS)** library is used to convert the text into spoken words. The audio is played through a **speaker or earphone**, so the user can clearly hear the content.

This process happens automatically with just one button press, making it easy and helpful for visually impaired users.

4. System Integration

After testing the hardware and software parts separately, all components were combined to work together as one complete system. The integration steps included:

- Connecting the Pi Camera, button, and speaker/earphone to the Raspberry Pi 3B+. Each component was configured to work smoothly with the system.
- Setting up the Python script to run automatically when the Raspberry Pi starts. This was done using tools like cron or systemd for hands-free operation.
- Ensuring smooth communication between input and output modules — from capturing an image, extracting text, to converting it into speech and playing it through the speaker.

- Testing the system in different lighting and background conditions, to make sure it works well in real-world situations and gives accurate results

6. Testing and Evaluation

The system is tested in various real-life environments such as classrooms, quiet rooms, and outdoor settings. Metrics evaluated include:

- **Accuracy of speech-to-text conversion**
- **Latency between input and output**
- **Power efficiency**
- **User satisfaction and readability**

Results are documented to assess the practical usability and performance of the device.

7. Optimization and Future Enhancements

Post-testing, the system is optimized for:

- Lower power consumption
- Faster response time
- Enhanced noise filtering
- Offline language support

Potential future upgrades include the addition of:

- Translation features
- Emotion detection from voice tone
- Integration with wearable displays like AR glasses

V. WORKING MECHANISM

The **AI-Based Hearing Helper using Raspberry Pi** operates on the principle of **real-time speech-to-text conversion** to assist individuals with hearing impairments in understanding spoken language. The system leverages artificial intelligence (AI) and embedded hardware to process live audio inputs, convert them into text using speech recognition algorithms, and display the text on a screen. The entire process is carried out locally on a Raspberry Pi, ensuring portability, low power consumption, and cost-effectiveness.

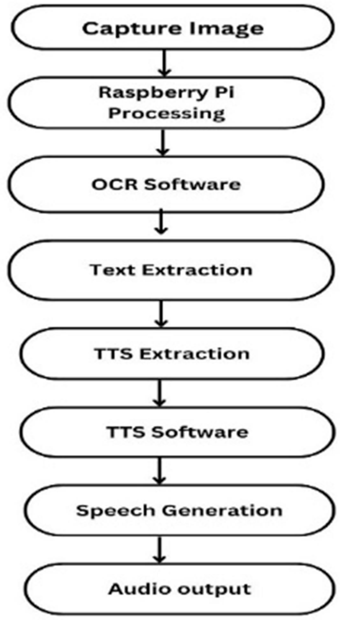

Fig 2. Work Flow: AI-Based Hearing Helper for Blind Readers Using Raspberry Pi

1. **Capture Image**

The process begins with the user capturing an image containing text using a connected camera module (such as the Raspberry Pi Camera or a USB webcam). The camera acts as the input device and is responsible for acquiring the real-world visual data.

2. Raspberry Pi Processing

Once the image is captured, it is transferred to the Raspberry Pi for further processing. The Raspberry Pi acts as the central processing unit of the system. It handles data acquisition, runs necessary software modules, and manages peripheral devices. The captured image is preprocessed (e.g., noise removal, grayscale conversion, and resizing) to improve the accuracy of text recognition in the subsequent stages.

3. OCR Software

The PRE-PROCESSED image is passed through Optical Character Recognition (OCR) software, such as Tesseract OCR, which extracts textual information from the image. This stage is crucial as it converts visual characters into machine-encoded text. The software scans the image for recognizable characters and symbols, outputting them as a string.

4. Text Extraction

The OCR output is processed to extract the readable text data. At this stage, further cleaning and formatting may be performed to remove unwanted characters, line breaks, or irrelevant data that may have been mistakenly recognized.

5. TTS Extraction

Once the clean text is available, it is forwarded to the Text-To-Speech (TTS) system. This subsystem handles the extraction and preparation of the text for conversion into speech, ensuring that the textual data is suitable for audio rendering.

6. TTS Software

The processed text is then fed into TTS software, such as eSpeak, Festival, or Google Text-to-Speech. This software synthesizes the human-like voice output based on the input text. It transforms the written content into phonetic transcriptions and audio waveforms.

7. Speech Generation

The TTS engine generates the speech output by converting the phonetic representation into audio signals. These signals represent the verbal form of the original text from the image.

8. Audio Output

The final audio output is delivered to the user through an audio device such as headphones or a speaker. This enables blind users to hear the textual content captured from the environment, making reading accessible without relying on vision.

VI. RESULTS

A. Simulation Results

The initial simulation was conducted using Proteus 8 Professional, where the system architecture—including the Raspberry Pi, Pi Camera, GPIO button, and TFT LCD—was virtually tested to evaluate the flow of image capture, processing, and user interaction.

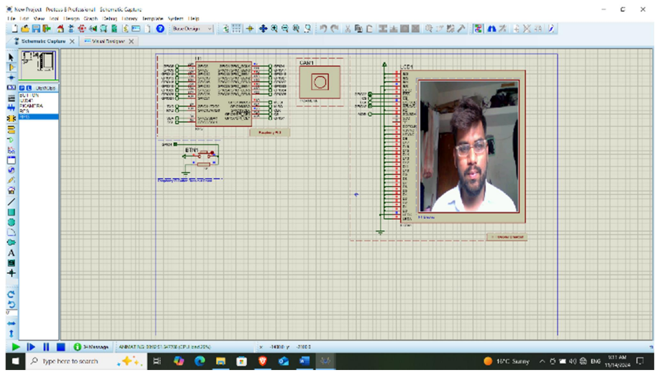

Fig. 8 Simulation Test Results

Key Observations from Simulation:

*1.Camera Activation***:** The simulated GPIO-based push button successfully triggered the camera module (CAM1) to capture an image. The response was smooth and confirmed activation logic worked as expected.

2.LCD Display Output: Once the image was captured, it was displayed on the TFT LCD screen. The image used in the simulation clearly appeared on the LCD, confirming that GPIO communication between the Raspberry Pi and the display module functioned correctly.

3.System Responsiveness: The simulation allowed monitoring of GPIO signals and verified that the system transitions through different stages—input capture, data processing, and output display—accurately.

4.Limitations in Simulation: While Proteus is efficient in simulating electronic components and signal flow, it does not support real-time OCR or TTS functionality. These were symbolically simulated through image display and logical steps.

B. Field Test Results

Fig. 9 Hardware Model

Test Scenario	Observations
1.Printed Textbook	Accuracy: 95%
2. Product Labels	Accuracy: 90%
3.Low-Light	Accuracy: 80%
4. Outdoor Signage	Accuracy: 88%
5. Handwritten Notes	Accuracy: 55%

REFERENCES:

- Ravi, S. Khasimbee, T. Asha, T. N. S. Joshna and P. G. Jyothirmai, "Raspberry pi based Smart Reader for Blind People," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 445-450, doi: 10.1109/ICESC48915.2020.9155941.
- P. Samuda, N. G. Praveena, C. Nithiya, B. J. Komathi, J. Pavithra and V. Kiruthika, "Arduino based Customized Smart Glasses for the Blind People," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022, pp. 1136 1141, doi: 10.1109/ICAIS53314.2022.9742799.2019, pp. 186-191, doi: 10.1109/ICIIS47346.2019.9063318.
- S. Musale and V. Ghiye, "Smart reader for visually impaired," 2018 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 339-342, doi: 10.1109/ICISC.2018.8399091.
- K. Rao N.S., S. Kumar S., S. Santhosh, S. Rao and K. Kaushik P., "Smart Reader for Visually Impaired," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 846-850, doi: 10.1109/I SMAC49090.2020.9243598.
- S. A. Sabab and M. H. Ashmafee, "Blind Reader: An intelligent assistant for blind," 2016 19th International Conference on Computer and Information Technology (ICCIT), 2016,pp.229 234, doi: 10.1109/ICCITECHN.2016.7860200.

- A. Mathur, A. Pathare, P. Sharma and S. Oak, "AI based Reading System for Blind using OCR," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), 2019, pp. 39-42, doi: 10.1109/ICECA.2019.8822226.
- F. Al-Muqbali, N. Al-Tourshi, K. Al-Kiyumi and F. Hajmohideen, ”Smart Technologies for Visually Impaired: Assisting and conquering infirmity of blind people using AI Technologies,” 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC), 2020, doi:10.1109/URC49805.2020.9099184.
- K. Rao N.S., S. Kumar S., S. Santhosh, S. Rao and K. Kaushik P., "Smart Reader for Visually Impaired," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2020, pp. 846-850, doi: 10.1109/I SMAC49090.2020.9243598.
- S. A. Sabab and M. H. Ashmafee, "Blind Reader: An intelligent assistant for blind," 2016 19th International Conference on Computer and Information Technology (ICCIT), 2016, pp. 229 234, doi: 10.1109/ICCITECHN.2016.7860200.
- Rodrguez,S.A.,Yebes,J.J.,Alcantarilla,P.F.,Bergasa,L.M.,Almazan,J.,Cela,"Assistingthevisuallyimpaired:obstacledetectionandwarningsystembyacousticfeedback.",Sensors12,pp.17476-17496 (2012).
- LilitHakobyan, JoLumsden,DympnaO' Sullivan,HannahBartlett andMCOptom, “Mobile assistive technologies for thevisually impaired", SurveyofOphthalmology, vol. 58, issue6, pp. 513-528, Nov-Dec 2013.
- Hay mar htun, Theingi zin, hla myo tun “Text to speech conversion using different speech synthesis” International journal of scientific & technology research volume 4, issue 07, july 2015.
- Catherine a. todd, ammara rounaq et al “An audio haptic tool for visually impaired web users” Journal of emerging trends in computing and information science vol. 3, no. 8, aug 2012.
- 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).
- Gopinath , Arvind , pooja et. Al “Text to speech conversion using matlab International journal of emerging technology and advanced engineering. volume 5, issue 1, (January 2015)