NUST
School of Electrical Engineering &
Computer Science

Project:

See 4 Me

Object Detection Aid for Visually Impaired

Internet of Things

(IT-863)

Submitted to: **Dr. Rafia Mumtaz**

Submitted by: **Safee Ullah Mughal (497945)**

Mohammad Noman (496862)

Mahnoor Fiaz (533243)

Laila Ikram (498215)
Zohaib Shahzad (496054)

Class: **MSIT-24 A**

Semester: **Spring 2025**

1. Abstract

*"**See For Me**" is a novel IoT-based assistive system developed to aid persons with visual impairment to see the objects in their immediate surroundings with distance through audio feedback provided in real time. It consists of an ESP32-CAM module used to capture photos and perform within the confines of an on-board model for efficient object detection. When an object is detected, a text-to-speech (TTS) mechanism converts the label into auditory form, so that the user can perceive it instantaneously. In a nutshell, this convenient, cost-effective, and wearable solution is an independent gadget with low power consumption and no constant dependency on the internet to operate. With the fusion of perception, processing, and communication, "See For Me" gives its users independence and awareness of their surroundings, thereby demonstrating how IoT positively impacts assistive technologies.*

2. Introduction

a. Background & Motivation

According to the World Health Organization, more than 250 million people worldwide have visual impairment, and many of these individuals face daily challenges in negotiating environments designed largely for sighted individuals. Simple tasks such as avoiding obstacles or identifying everyday objects may become very difficult, if not dangerous, for such people. Embedded systems and smart devices can now potentially address the aforementioned accessibility gap with the emergence of the Internet of Things (IoT) via assisted technologies that are connected, intelligent, and real-time.

b. Problem Statement

A white cane, looking forward toward a guide dog or any other means, will give a semblance of independence in the traditional sense. However, dynamic interpretation and communication of the contextual reality of the environment are beyond their realm. The present age demands lightweight visual devices such as object recognition tools requiring little or no use of expensive bulky equipment by visually impaired persons. Most of the existing solutions available today are either expensive, data-dependent, or hardware-dependent, thus not being specifically available for everyday use.

c. Objectives

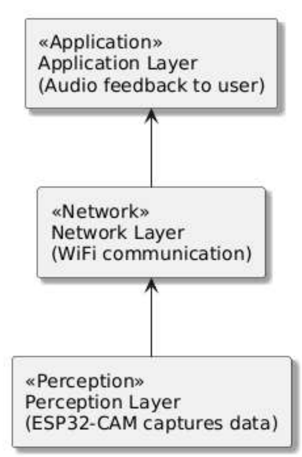
This project, entitled "See for Me", aims to develop a low-cost, portable, and stand-alone device under the banner of IoT-based device which should be able to detect and provide instant voice feedback of surrounding common objects around the user. For this purpose, the ESP32-CAM module-a compact microcontroller integrated with built-in Wi-Fi capabilities and camera-for images captured would be processed using a lightweight object detection algorithm and made voice-enabled by a Text-to-Speech (TTS) module-the identified object name would be converted to audio output to a speaker or headphone where a user would be able to listen-in real time-to understand the surroundings**.**

d. IoT Prospective

The project in question, from an IoT point of view, realizes a three-layer architecture.

1. **Perception Layer**: ESP32-CAM and camera sensor, which perceive the visual data from the real world.
2. **Network Layer**: Facilitating data transfer, and in extended versions, it is capable of communicating with cloud services via Wi-Fi.
3. **Application Layer**: This provides audio feedback and situational information to the end user.

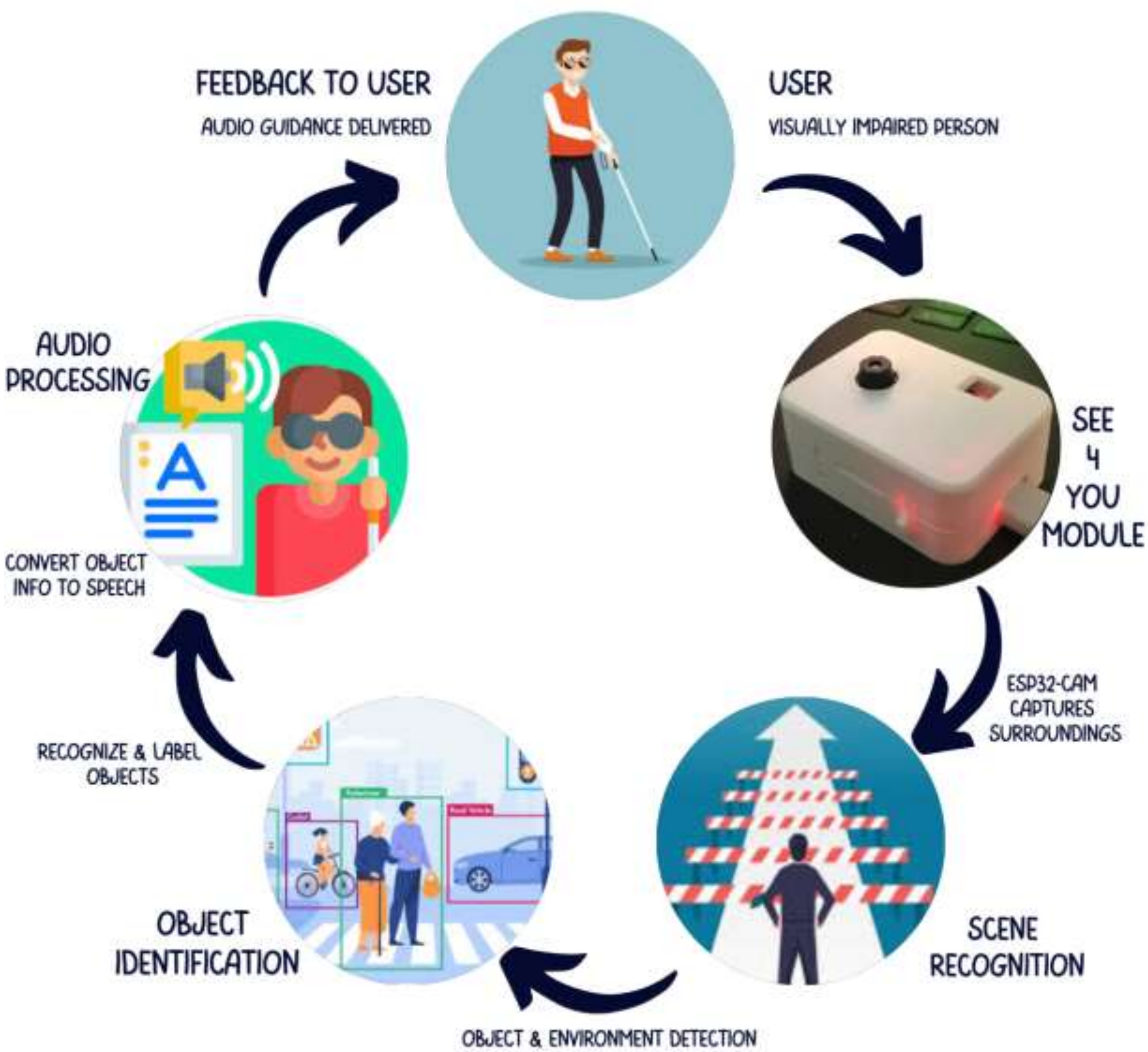
Through this solution, IoT is demonstrated to assist technological advancements with smarts lending to a real-time, worthwhile, and interactive experience using minimal resources. Being modular and scalable in design, the solution offers scope for future extensions such as obstacle detection, GPS integration for navigation, and cloud-based analytics for remote monitoring.



e. Scope of the Project

The first prototype is mainly concerned with the detection of certain predefined everyday objects and bringing it to the user's notice through sound. The device is envisaged to be:

- Wearable and battery-operated
- Not dependent on constant internet access
- Easily expandable for future IoT additions (i.e., voice commands, syncing with the cloud, emergency alerts)



"*See for Me*" is a pragmatic advanced step toward harnessing the potentials of IoT technologies to render humane and real-world solutions that enhance the possibilities for the easy and safe living of persons with disabilities.

3. Literature Review

The most important electronics for the visually impaired. Most electronic devices that assist people with vision loss rely on data collected from the environment and delivered to the user through tactile, auditory, or both formats. Different ideas are expressed and it remains to be determined whether the type of feedback is more appropriate. However, a system must have several basic components to function well, regardless of its services.An electronic device designed to provide navigation and orientation support for individuals with visual impairments should incorporate specific features that cater to their unique needs. These features serve as essential tools for blind individuals to evaluate the device's overall effectiveness and dependability.

Existing devices

a. Eye Replacement:

For blind or visually impaired persons, Bharambe et al. created an implanted device that serves as a replacement for their eyes and helps with navigation. The primary component of the embedded device is a TI MSP430G2553 microcontroller (Texas Instruments Incorporated, Dallas, TX, USA). The authors used an Android application to put the proposed algorithms into practice. The function of this application is to locate the user and provide better directions using GPS, upgraded GSM and GPRS.

b. Combining Artificial Vision and GPS(FAV&GPS)

To improve the mapping of the user's position and the distribution of surrounding objects, an assistive device for the blind using two functionalities based on the map-matching technique and artificial vision was invented. The first function helps in finding the needed item and allows the user to issue commands by pointing the head in the desired direction.

Automatic identification of visual targets is made available to others. This wearable gadget has two Bumblebee stereo cameras for video input located on the helmet, a GPS receiver, headphones, a microphone and an Xsens Mti tracking device for motion detection. It attaches to the user's head.

c. Banknote recognition (BanknoteRec):

He introduced a tool to help blind individuals classify different denominations of notes and coins. The input of the system was the OV6620 Omni vision CMOS camera, while the SX28 microcontroller was its process and output model (speaker). Red, green and blue average colors are used to describe the type of banknote using the RGB color model.

d. TED

To facilitate the transmission of information and navigation for blind people, a small dipole antenna was created and connected to an electro-tactile device (TED). The TED device and electrode array are designed to communicate wirelessly through this antenna. Front and rear antenna design.

e. CASBlip

Researchers unveiled CASBlip, a cutting-edge wearable assistance device designed specifically for those with visual impairments, in the study mentioned in reference [31]. With a primary focus on providing capabilities including object identification, orientation facilitation, and navigation aid, this system is specifically designed for people who are either completely or partially blind. The crucial sensor and acoustic modules are at the heart of the system's architecture. A set of glasses with sophisticated 1X64 3D CMOS image sensors and laser light beams precisely used for exact object detection are integrated into the sensor module.

f. RFIWS

The Radio Frequency Identification Walking Stick (RFIWS) was created to help blind people cross their sidewalks. This approach can be used to determine the approximate distance between a blind person and the edge of the sidewalk. Information is transmitted and received through a radio wave medium using radio frequency identification (RFID) devices [33]. The three primary elements of RFID technology are the RFID tag, the reader, and the hub.

g. Cognitive Guidance System (CG System)

A method for guiding blind people through an organized environment was proposed by Landa et al. using fuzzy Mandani-type decision rules and a vanishing point, this design uses the Kinect sensor and stereoscopic vision to determine the distance between the user and the obstacle and to guide the user along the path.

h. The UltraCane, an ultrasonic cane employed for navigation

She employed an ultrasound-based cane in addition to the C-5 laser stick created by Krishna Kumar et al. to help the blind. This project aims to replace ultrasonic sensors with laser sensors to eliminate laser risk. This team is capable of spotting impediments both above and below the ground.

i. Obs avoid employing throttling

Using the Kinect depth camera, Muhamad and Widyawan presented an obstacle avoidance system for the blind. prototype of the system under review. Using auto-adaptive Thresholding, the distance between the user and the obstruction is calculated. A laptop with a USB hub, a headset, and a Microsoft Kinect depth camera make up the system's main components.

4. System Architecture

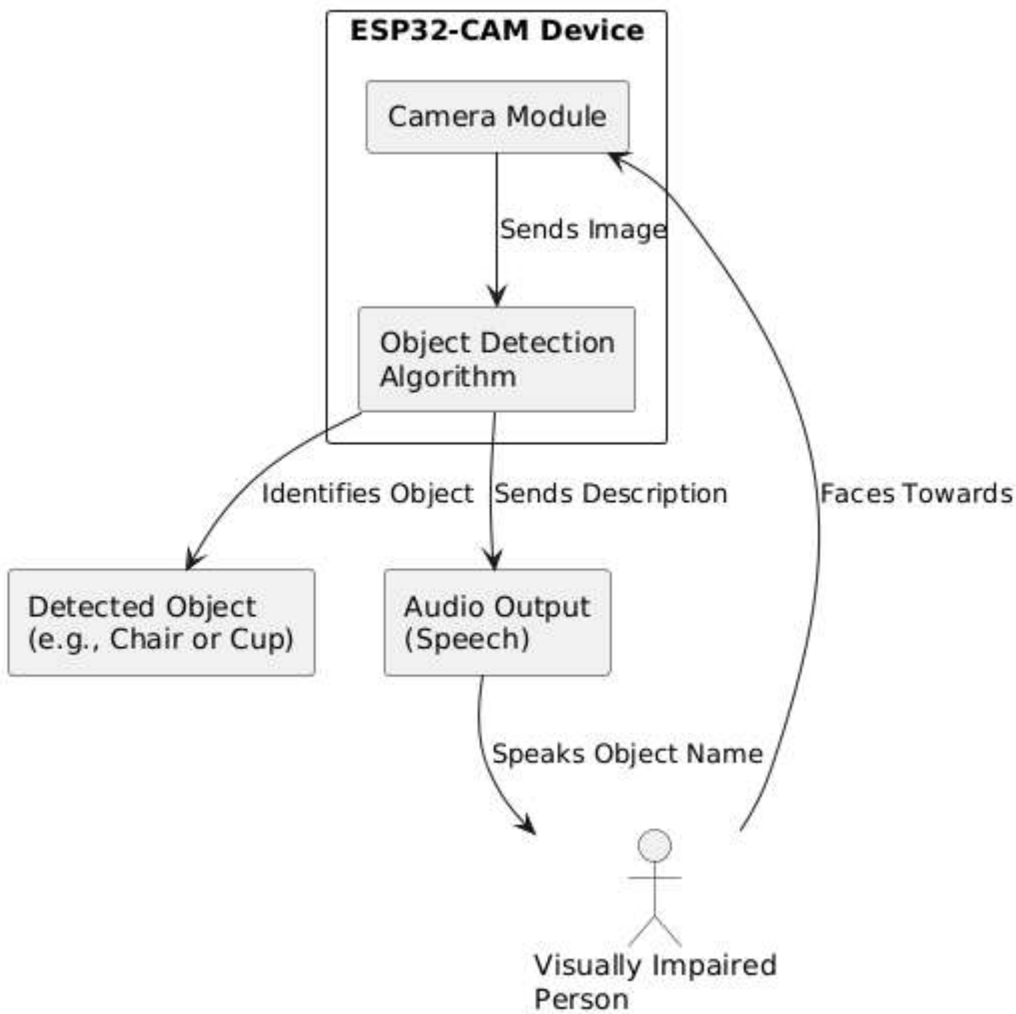
The architecture for the "See for Me" project is proposed in such a way to provide a real-time object identification and feedback system for blind or visually impaired people. The architecture has several layers which flow seamlessly together to create the project. Since the embedded system will be responsible for sensing, processing, and interacting with the user, we consider it the centre of the solution.

a. Overview of embedded system architecture

The embedded system is responsible for:

- Image capture using the ESP32-CAM camera module.
- Object detection by computer vision algorithm.
- Text to Speech (TTS), for audible feedback for the user.
- User input, to engage the system.

The embedded system operates with no externally wired dependencies, and can therefore be employed and augmented with communication and application layers. The major components of the embedded system architecture are shown in the diagram below:

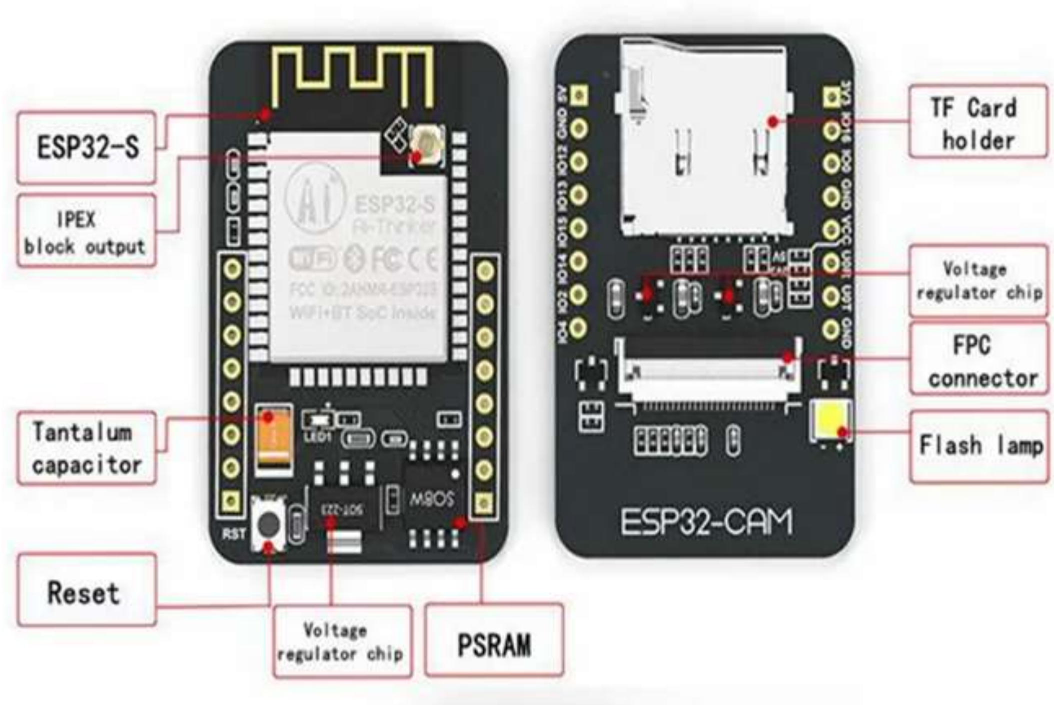


b. Components of the Embedded System

- **ESP32-CAM Module (Main microcontroller)**

  The ESP32 camera module will be the main device for object detection and image processing. This will cover steps for taking and processing images and the output of that image. It has a camera (OV2640) with Wi-Fi and Bluetooth.

  Resolution: **1600 × 1200 2MP**



- **ESP32 MB module**

  The ESP32-MB will be used as another microcontroller board to assist with overall control of the system as well as power control, and possibly handling some of the more complex processing or communications with external modules.

- **3D Enclosure**

To design the 3D enclosure, a custom enclosure was made to allow the ESP32-CAM, ESP32-MB board and the 5000mAH power bank to be securely enclosed. The dimensions were specific to allow a nice tight, light, and portable enclosure while allowing effective ventilation to move heat. Appropriate access points were provided for the camera lens, power inputs and required buttons for easy access to use and maintenance.

- **Battery Pack (5000 mAh)**

The power bank will provide the necessary energy to run the ESP32-CAM and ESP32-MB board. It allows the system to be portable and not restricted to a fixed power outlet, which is important for a wearable system such as this.

c. Power breakdown

Power consumption varies with it being in active mode versus light sleep or deep sleep mode.

ESP32-MB Board: Depending on the operations, it may have its own power consumption (around 100mA to 300mA when the ESP32-MB board is active).

Power Bank: A 5000mAh power bank can run the entire system for long periods of time

Total Power Consumption:

The power consumption of both components are:

ESP32-CAM active mode: **310mA**

ESP32-MB board: **200mA**

T**otal current consumption** will be:

$$Total\ Current\ Consumption = 310mA + 200mA = 510mA$$

Battery Capacity:

$$Effective\ Capacity = 5000\ mAh \times 0.9 = 4500\ mAh$$

We already know the power bank has **5000mAh** capacity, but after accounting for 90% efficiency, the available capacity is **4500mAh**.

Battery Life Calculation:

The **battery life** in hours can now be calculated as:

$$\text{Battery Life} = \frac{4500\ \text{mAh}}{510\ \text{mA}} \approx 8.82\ \text{hours}$$

The **battery life** of the system (ESP32-CAM and ESP32-MB board) with a **5000mAh power bank** will be approximately **8.82 hours** when both components are running in active mode with the flash on.

d. Communication

The communication module in the "See for Me" system is designed to efficiently manage the flow of information between its core components, ensuring real-time object detection and audio feedback without significant latency. This seamless communication is crucial for delivering an interactive and responsive user experience.

Intra-device Communication

- The ESP32-CAM module captures real-time images of the user's surroundings.
- These images are processed locally using a lightweight object detection algorithm. The detection process identifies objects, which are then labeled and prepared for audio output
- Communication between the camera, processing unit, and TTS module is handled internally, ensuring minimal delay in generating feedback.

Feedback Transmission

- Once an object is detected, the label is sent to the Text-to-Speech (TTS) module, where it is converted into spoken words.
- The audio message is immediately transmitted to the user through connected speakers or headphones.
- This real-time communication allows users to be instantly aware of objects in their path, enhancing their navigation capabilities.

Wireless Connectivity:

- The device supports Wi-Fi communication for additional functionalities, such as software updates or pairing with mobile applications.
- This optional connectivity allows for configuration changes, firmware updates, or even integration with cloud-based services if needed, although the primary operations remain fully offline for reliability.

Low Latency and Power Efficiency:

- The communication within the device is optimized for low latency, meaning the user receives feedback almost instantly after an object is detected.
- The use of local processing reduces power consumption and eliminates the dependency on constant internet access, making it suitable for long-term use in various environments.

5. Software Requirements (Mobile Application)

To complement the hardware capabilities of the “See For Me” device, a dedicated Android mobile application was developed using Java in Android Studio. This application serves as the core interface for processing real-time image data, performing object detection, and providing audio feedback to the user through Text-to-Speech (TTS).

a. Purpose of the App

- It receives compressed images in the form of binary data from the ESP32-CAM.

- It detects objects using OpenCV.
- It gives voice feedback using Text-to-Speech (TTS).
- It helps the user understand their surroundings through sound.

b. Main Features

- **Object Detection**: Finds objects like people, chairs, and doors.
- **Distance Estimation**: Tells how far an object is (e.g., "Person, 2 feet ahead").
- **Audio Output**: Speaks the object name through the phone speaker or Bluetooth headset.
- **Offline Functionality**: Works without internet after connecting to the ESP32-CAM.

c. Technologies Used

- **Programming Language**: Java
- **Platform**: Android Studio
- **Libraries**:
  - **OpenCV** for object detection
  - **Android TTS** for voice feedback
  - **WebSocket Client** for receiving video stream

d. Dependencies Used

- We used **pre-built OpenCV libraries** and integrated them directly into our app.
- GitHub repositories were used for:
  - **OpenCV Java wrappers** for Android
  - A simple **WebSocket library** to receive live stream data
- These were added manually and through Gradle, depending on the library.

e. Simple Working

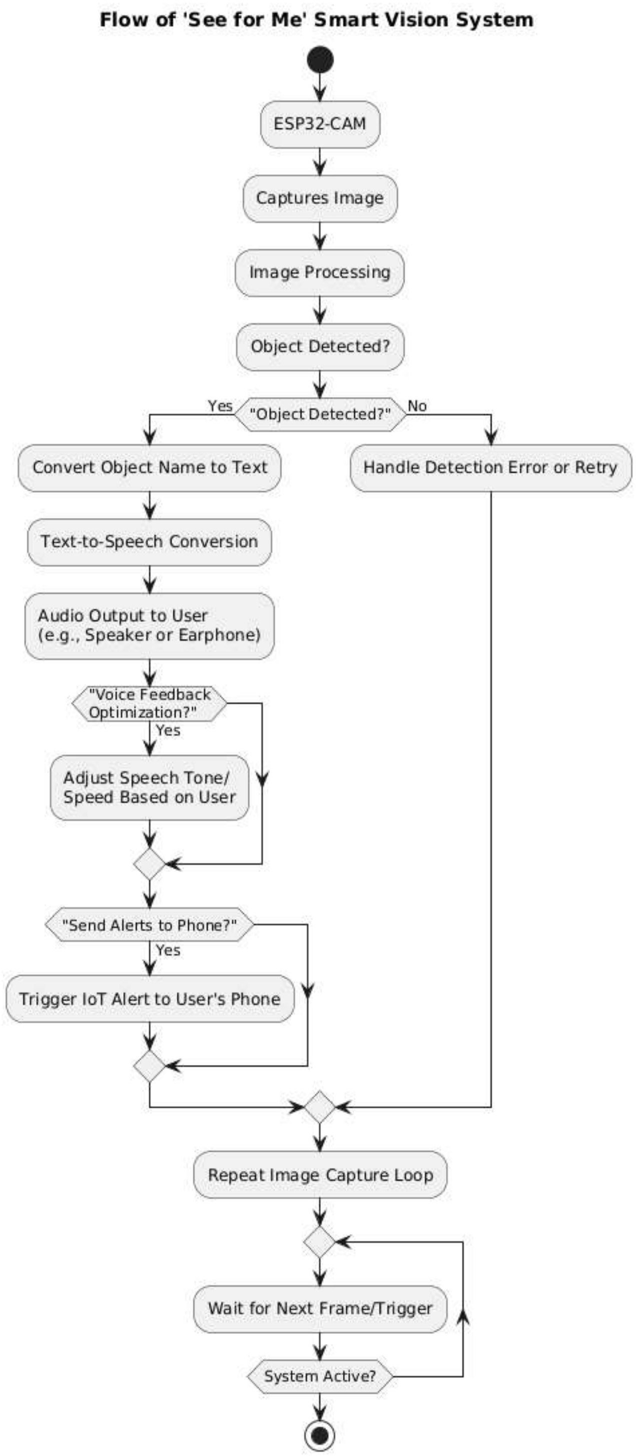
1. The app connects to the ESP32-CAM over Wi-Fi.
2. It receives and processes live images.
3. Objects are detected and labeled using OpenCV.
4. The label is converted into speech and played back to the user.

f. Why It’s Helpful

1. User friendly interface
2. Works in real time
3. Helps blind users “see” their surroundings with sound

6. Methodology

The "See for Me" system is a combination of image processing, object detection, and text-to-speech (TTS) technologies that support visually impaired users when navigating. The system works as follows:



a. **ESP32-CAM for Object Detection:**

The basis of the system is the ESP32-CAM module; the ESP32-CAM module contains an onboard camera (OV2640). The camera takes a snapshot of the surroundings wherever it points and performs image recognition on the image, but not in a traditional sense. Normally, image recognition performs complex processes, but our method works on a basic image recognition implementation. At this point, the ESP32-CAM is not always connected or dependent on the ESP32-MB board, but the ESP32-Board provides a means to communicate, perform additional processing, and control some of the limits of the ESP32-CAM.

b. **Object Recognition and Detection:**

Object detection is a computer vision technique for locating occurrences of objects in images or videos. Object detection algorithms use machine learning or deep learning to produce meaningful results. OpenCV is required for the device to run it natively. However, once those two obstacles are out of the way, doing things like edge detection, as this project demonstrates, is within the realm of possibility.

Object identification

Object recognition is a computer vision technique for identifying objects in images or videos. Object recognition is a key output of deep learning and machine learning algorithms.

OpenCV(open source computer vision library)

A machine learning and computer vision library is accessible for free under the name OpenCV. To facilitate the use of machine perception in commercial goods and to provide a common foundation for computer vision applications, OpenCV was developed.

Aspects of the OpenCV Library

- Visual reading and writing.
- Perform feature detection, process images (filter, transform), and capture and save videos.

- Identify particular objects in movies or pictures, like faces, eyes, and cars.
- Analyze video, that is, determine its motion, eliminate the backdrop, and locate objects.

Object Detection in "See For Me" Using OpenCV

In the “See For Me” project, object detection is one of the core components enabling real-time environmental awareness for visually impaired users. We achieve this using OpenCV integrated into a customized Android application.

The detection process works as follows:

- **Live Video Capture:**

  The ESP32-CAM module captures live video at a resolution of 1600x1200 pixels. This high resolution ensures clarity and sharpness, which are critical for accurate detection—especially under good lighting conditions.

- **Data Streaming to Mobile App:**

  The video feed is streamed wirelessly to a smartphone via Bluetooth or Wi-Fi, where the image processing takes place in real time.

- **Processing with OpenCV:**

  Once the video reaches the app, OpenCV begins analyzing the individual frames. Object detection is performed by applying a combination of:

  - Edge detection
  - Haar cascades or pre-trained DNNs for specific object categories (like humans, vehicles, obstacles)
  - Contour analysis to localize the object
- **Labeling and Feedback:**

Detected objects are enclosed in bounding boxes, labeled with their names (e.g., “Person,” “Chair”), and then the label is converted into text and speech using Android’s text-to-speech engine.

- **User Feedback:**

The spoken object names are delivered to the user through Bluetooth earphones, creating a seamless and non-intrusive audio guide.

This real-time visual interpretation allows “See For Me” to act as a digital extension of vision, providing constant updates about the user’s surroundings.

Distance Estimation for Safer Navigation

To make object detection meaningful for practical navigation, “See For Me” incorporates a distance estimation mechanism, helping users assess how far away an obstacle is before taking action.

Key Implementation Details:

Effective Range:

The current setup allows detection and relative distance estimation of objects up to 6 meters (approximately 18 feet) under good lighting conditions. This range was optimized based on real-world testing, balancing camera capability, user safety, and system speed.

Estimation Method:

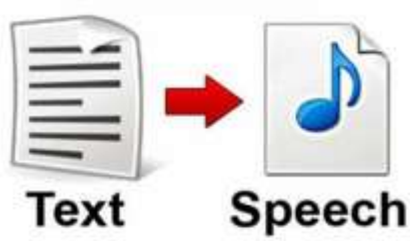
i. We approximate distance using the apparent size of the object in the video feed.
ii. Objects detected closer to the camera appear larger, and OpenCV's bounding box dimensions help us infer this distance.
iii. For known object classes (like human figures), we use reference height ratios to calibrate distance.

Resolution Advantage:

The 1600x1200 resolution significantly enhances distance accuracy by allowing fine-grain object scaling and depth estimation.

c. Text-to-Speech (TTS):

Text-to-Speech (TTS) is a form of assistive technology that converts written text into spoken audio output. In the **"See For Me"** project, the TTS module plays a crucial role in providing real-time auditory feedback to visually impaired users by vocalizing the names of detected objects along with their estimated distances.



Role of TTS in the System:

The TTS module bridges the gap between object detection (performed via OpenCV on the mobile app) and user comprehension by:

- Receiving text labels (e.g., "Chair," "Door," "Person") from the object detection algorithm.
- Converting these labels into clear, natural-sounding speech.
- Delivering audio output through the mobile device's speaker or connected headphones.

This allows users to **understand their surroundings without visual cues**, enhancing independence and safety.

Key Features of the TTS Module:

a) **Real-Time Conversion**
   - The TTS engine processes text **instantly** (with minimal latency) to ensure timely feedback.
   - Works in sync with the object detection system to avoid delays.

b) **Natural-Sounding Speech**
   - Uses pre-trained neural networks for human-like intonation and pronunciation.
   - Adjustable speech speed, pitch, and volume to suit user preferences.

c) **Offline Functionality**
   - Unlike cloud-based TTS services (e.g., Google TTS), this system uses an embedded or mobile-based TTS engine to function without internet dependency.
   - Ensures reliability in areas with poor or no connectivity.

d) **Multilingual Support (Optional Future Enhancement)**
   - Can be extended to support multiple languages for broader accessibility.

Implementation of TTS in the "See For Me" System

The TTS module operates as follows:

1. **Input from Object Detection:**
   - OpenCV identifies an object (e.g., "Cup") and estimates distance (e.g., "1.5 meters").
   - The mobile app formats this into a text string: **"Cup, 1.5 meters ahead."**
2. **Text Processing:**
   - The TTS engine parses the text, applying linguistic rules for proper pronunciation.
3. **Audio Generation:**
   - The processed text is converted into a digital speech signal.

- Output is played through the mobile device’s audio system.

4. **User Feedback Optimization:**
   - Users can adjust speech **speed, volume, and tone** via the mobile app.
   - Critical alerts (e.g., "Obstacle close!") can be prioritized with different voice tones.

Why TTS is Essential for Visually Impaired Users:

- **Instant Environmental Awareness:**
  Unlike Braille or tactile feedback, speech is immediate and requires no additional learning.
- **Hands-Free Operation:**
  Users can navigate without needing to interpret vibrations or physical cues.
- **Adaptability:**
  Can be customized for different languages, accents, and speech preferences.

d. Power Management:

The system's 5000mAh charger allows the device to remain portable and be used for long periods without being tethered to a power source. The ESP32-CAM and ESP32-MB are both low-power devices to enable continuous functionality in a portable device.

7. IoT Communication and Protocols

The "See For Me" assistive system leverages local area communication protocols to ensure real-time, reliable, and internet-independent functionality. The design prioritizes both efficiency and accessibility, particularly in environments where continuous internet access may not be available. The system uses two primary communication protocols: WebSocket and UDP (User Datagram Protocol), operating over a Local Area Network (LAN) where both the ESP32-CAM device and the user's mobile device are connected to the same network.

a. WEB SOCKET (real time image streaming)

The ESP32-CAM module acts as a WebSocket server, enabling the transmission of live image frames captured by the onboard camera to the mobile device. WebSocket is chosen over traditional HTTP communication due to its persistent, full-duplex connection, which allows real-time, low-latency data transfer between the server (ESP32-CAM) and the client (mobile application). This continuous connection ensures a smooth and fast video stream, which is critical for detecting objects in the environment without delay.

Key Features of WebSocket in this system:

- Maintains a persistent connection over TCP.
- Enables real-time transmission of image data with minimal overhead.
- Reduces latency compared to traditional HTTP polling.
- Operates fully within the LAN, eliminating the need for internet access.

This setup ensures that the mobile device can continuously receive image streams from the ESP32-CAM, perform object detection (if offloaded to the phone), and respond quickly.

b. UDP (light weight text-based communication)

Alongside image streaming, the system also uses the UDP (User Datagram Protocol) for sending text-based information such as object labels or command/status messages from the mobile device to the ESP32-CAM (or vice versa). UDP is a connectionless protocol known for its minimal transmission overhead and fast delivery, making it ideal for short, real-time messages where speed is more critical than guaranteed delivery.

In this context, UDP is used to:

- Transmit detected object labels to the ESP32 for TTS conversion.
- Communicate status signals or control messages between the ESP32 and mobile device.
- Maintain fast, lightweight communication without requiring an acknowledgment handshake.

Advantages of using UDP in this system are

- Low latency and reduced processing time.
- Ideal for short messages that require fast delivery.
- Works efficiently within a local network.

8. Implementation and Results

Some of the overall objectives of this phase are as follows.

a. Cost effective

Cost effectiveness is saving money by replacing a product or process to make it work better. This is done to improve the economic results of the organization by reducing the cost of purchase and increasing efficiency in all areas.

Cost analysis

Esp32 cam module	$8
3d printing	$3
Battery	$3
Connections	$3
Total cost	$17

b. Efficiency

Performance: all necessary features listed are supported.

1. Live broadcast
2. Object detection
3. Object identification
4. Distance estimation
5. Acoustic feedback

- **Cordless connectivity**: The assistive device must be wireless and wirelessly connected to the database to ensure the exchange of information.
- **Reliability**: The hardware and software requirements for the device are met.Simple: An easy-to-use interface and user-friendly controls can make the device simple for the user to utilize.
- **Wearable:** According to our research and analysis, wearing a device offers the user greater flexibility and comfort than holding it in their hands.
- **Economically feasible**: Users can afford the technology if they want to improve their quality of life; otherwise, few can.

In order to address the aforementioned flaws, we intend to continue this research by examining each function separately and building an intelligent framework that provides all the aforementioned capabilities while being more scalable and reasonably priced.

AUTHOR	DEVICE	YEAR	SENSOR
Shantanu gangway	Smart stick	2010	Photo diodes
S.Chew	Smart cane	2011	IR
Shantanu gangwar	C5 laser cane	2012	GPS technology
Lislie kay OBE	Sonic Torch	2013	Ultrasonic sensor
Vignesh.N	Smart shoe	2015	Ultrasonic sensor + Arduino UNO

Our aim is to design a device which should have all the required features but should be cost effective as all people can not afford high cost technologies. So keeping in view this point we have designed a device which just cost us $17. On checking the device it was clear that the device is so efficient in all the objectives which were set at the beginning including live streaming, object detection, identification and giving an acoustic feedback.. Our device is user friendly, economically feasible and reliable. The future works in our design includes refining each function to its best level and having the ability for the customized design provision.

9. Future Work

a. Future Enhancements for Object detection

- While the current version uses monocular cues (single-camera logic), future versions of “See For Me” may integrate stereo cameras or depth sensors for more precise spatial awareness.
- Integration with sensor fusion (e.g., using ultrasonic modules or LIDAR) is also considered to provide cross-validated distance readings.

b. Future Enhancements for the TTS Module

- **Emotional Tone Adjustment:**
  Different urgency levels (e.g., calm voice for general objects, urgent tone for obstacles).
- **Voice Customization:**
  Allow users to select preferred voices (male/female/neutral).
- **Contextual Feedback:**
  Instead of just "Chair," provide "Chair on your left, 2 steps away."

The Text-to-Speech (TTS) module is a critical component of the "See For Me" system, enabling visually impaired users to interact with their environment naturally and efficiently. By converting detected objects into real-time spoken feedback, the system enhances mobility, safety, and independence. Future improvements in voice personalization and multilingual support could further enhance its usability.

10.Conclusion

Many methods have been created with the goal of assisting visually impaired people and improving their quality of life generally. Unfortunately, the vast majority of these devices show functional limits. In order to better serve the community of people who are blind or visually impaired, The goal is to demonstrate how assistive technology has advanced. In doing so, this literature review seeks to meticulously analyze the key strategies described in earlier works that serve this group of readers. It explores their improvements, benefits, drawbacks, and accuracy.This project's main goal is to provide insight into the trajectory of development by illuminating the advancements made in the field of assistive technology. The goal is to rigorously analyze the complex features of these systems, taking into account both their advantages and disadvantages. This in-depth investigation aims to unearth subtleties that may open doors for next scientists and inventors. This study aims to give a springboard for the conception and design of technologies that not only assure the safety but also stimulate independent mobility for people with visual impairments by recognizing and addressing common problems.

11.References

[1]. Kassim AM, Shukor AZ, Chan XZ, Yasuno T (2018) Exploratory study on navigation system for visually impaired Person. Aust J Basic Appl Sci 7(14):211–217

[2].Kassim AM, Shukor AZ, Chan XZ, Yasuno T (2019) Performance study of developed SMART EYE for visually impaired person. Aust J Basic Appl Sci 7(14):633–639

[3].Kassim AM, Yasuno T, Suzuki H, Jaafar HI, Aras MSM, Jafar FA, Sivarao S (2016) Conceptual design and implementation of electronic spectacle type based on obstacle detection system for visually impaired person. J Adv Mech Des Syst Manufact (JAMDSM) 10(7):JAMDSM0094

[4]. Kassim AM, Yasuno T, Suzuki, AMSM, Shukor AZ, Jaafar HI, Jafar FA (2018) Vision-based tactile paving detection method in navigation systems for visually impaired persons. In: Advances in human and machine navigation systems, pp 29–46Google Scholar

[5]. Felix M, Action N (2021) Ultra-low power techniques in energy harvesting wireless sensor networks: recent advances and issues. Sci Afr 11:e00720Google Scholar

[6]. Ziyu Z et al (2021) A review of technologies and applications on versatile energy storage systems. Renew Sustain Energy Rev 148:111263Google Scholar

[7]. Mohit A, Manju B, Sukhvinder SB (2021) Wireless battery recharging through UAV in wireless sensor networks. Egypt Inform JGoogle Scholar

[8]. Qichao Z et al (2021) State-of-health estimation of batteries in an energy storage system based on the actual operating parameters. J Power Sour. 506:230162Google Scholar

[9].Ahmed NA (2021) Integration of energy storage system and renewable energy sources based on artificial intelligence: an overview. J Energy Storage 40:102811Google Scholar

[10]. Tianmei C, Yi J, Hanyu L, Antao Y, Meiyi L, Bing C, Ying X, Qiang C (2020) Applications of lithiumion batteries in grid-scale energy storage systems. Trans Tianjin Univ 26:208–217CrossRef Google School

[11].Shah, C.; Bouzit, M.; Youssef, M.; Vasquez, L. Evaluation of RUNetra tactile feedback navigation system for the visually-impaired. In Proceedings of the International Workshop on Virtual Rehabilitation, New York, NY, USA, 29–30 August 2006; pp. 72–77.

[12].Hersh, M.A. The Design and Evaluation of Assistive Technology Products and Devices Part 1: Design. In International Encyclopedia of Rehabilitation; CIRRIE: Buffalo, NY, USA, 2010. [Google Scholar]

[13].Marion, A.H.; Michael, A.J. Assistive technology for Visually-impaired and Blind People; Springer: London, UK, 2008. [Google Scholar]

[14].Tiponut, V.; Ianchis, D.; Bash, M.; Haraszy, Z. Work Directions and New Results in Electronic Travel Aids for Blind and Visually Impaired People. Latest Trends Syst. 2011, 2, 347–353. [Google Scholar]

[15].Tiponut, V.; Popescu, S.; Bogdanov, I.; Caleanu, C. Obstacles Detection System for Visually-impaired Guidance. New Aspects of system. In Proceedings of the 12th WSEAS International Conference on SYSTEMS, Heraklion, Greece, 14–17 July 2008; pp. 350–356.

[16].Dakopoulos, D.; Bourbakis, N.G. Wearable obstacle avoidance electronic travel aids for blind: A survey. IEEE Trans. Syst. Man Cybern. Part C 2010, 40, 25–35. [Google Scholar] [CrossRef]

[17].Renier, L.; De Volder, A.G. Vision substitution and depth perception: Early blind subjects experience visual perspective through their ears. Disabil. Rehabil. Assist. Technol. 2010, 5, 175–183. [Google Scholar] [CrossRef]

12. Appendices

a. ESP32 CAM datasheet:

- Product Name: ESP32-CAM.
- WiFi+Bluetooth module: ESP-32S.
- Camera Module: OV2640 2MP.
- Flash Light: LED Built-in on Board.
- Operating Voltage: 3.3/5 Vdc.
- Onboard TF card slot, supports up to 4G TF card for data storage .
- RAM: Internal 512KB + External 4MB PSRAM.
- Power consumption:
  - Flash off: 180mA@5V.
  - Flash on and brightness max: 310mA@5V.
  - Deep-Sleep: as low as 6mA@5V.
  - Modern-Sleep: as low as 20mA@5V.
  - Light-Sleep: as low as 6.7mA@5V
- Dimensions: 40.5mm x 27mm x 4.5mm