Object detection and Audio Feedback system for Visually impared

Tresa Nancy Thareparambil Paul

***Abstract*—In this research, a smartphone-based obstacle detection system for those with visual impairments is shown. It records video in real time, uses a deep learning model to process it, and then uses text-to-speech to deliver audible feedback. The system does not require new hardware because it uses OpenCV for video processing and a pre-trained object identification model for recognition. Although it improves accessibility, issues with processing speed, model correctness, and internet dependence exist. Future developments will concentrate on enhancing offline capabilities, cutting latency, and improving detection accuracy. This project showcases an affordable and useful assistive technology that helps visually impaired individuals navigate more easily.**

***Index Terms*—Obstacle detection, visually impaired assistance, real-time object recognition, deep learning, text-to-speech, OpenCV, smartphone-based system, accessibility technology, assistive AI, computer vision.**

I. Introduction

Technological advancements have significantly contributed to improving accessibility for visually impaired individuals. This project introduces a real-time obstacle detection system that utilizes a smartphone camera and deep learning-based object recognition to assist users in navigating their surroundings. By employing OpenCV and a pre-trained neural network, the system processes video input, identifies objects, and provides auditory feedback through text-to-speech conversion. Designed to be cost-effective, it requires only a smartphone and a laptop, eliminating the need for additional hardware. The system transmits video from the smartphone to the laptop, where image processing and object detection take place. The detected objects are then converted into speech and conveyed to the user. This AI-powered solution enhances mobility and safety for individuals with visual impairments.

II. Related Work

Various studies have explored the use of technology to assist visually impaired individuals in navigating their environment. Traditional methods often rely on ultrasonic sensors, LiDAR, or specialized wearables to detect obstacles and provide alerts. However, these systems can be costly and require complex hardware setups. With recent advancements in deep learning, researchers have investigated the potential of smartphone cameras for real-time object detection. Models such as YOLO, SSD, and ResNet have been widely used for object recognition, with some approaches utilizing cloud computing for better accuracy and others focusing on edge computing to reduce processing delays. This project differs by offering a smartphone-based, cost-effective solution that eliminates the need for additional hardware while integrating text-to-speech technology for improved accessibility.

III. Problem Statement

It can be extremely difficult for those with visual impairments to independently and securely navigate their environment. Conventional assistive technology, such wearable sensors and canes, can be costly, need extra gear, and might not accurately identify obstacles in real time. Current computer vision-based solutions frequently rely on expensive hardware or cloud computing, which may cause latency problems. A real-time, affordable, and easily accessible solution that makes use of widely accessible technology is required. In order to help visually impaired people navigate their surroundings more skillfully without the need for specialist technology, this project intends to create a smartphone-based obstacle detection system that uses computer vision and text-to-speech conversion to deliver real-time audio feedback.

IV. Analysis/Proposed Solution/Proposed Approach

This study suggests a smartphone-based obstacle detection system that is both affordable and easily accessible for people with visual impairments. The system uses a laptop to process real-time video from a smartphone camera using deep learning-based object detection, and it converts text to speech to deliver aural feedback. The system guarantees precise obstacle identification by using OpenCV for image processing and a pre-trained neural network for object recognition. After being identified, the objects are translated into voice and sent to the user through audio streaming. This strategy makes use of generally accessible technology, which makes it inexpensive and simple to adopt in contrast to traditional solutions that call for extra hardware. In order to improve reliability, future developments will concentrate on raising detection accuracy, cutting latency, and enabling offline capability.

A. System Architecture

The proposed system follows a structured architecture to enable real-time obstacle detection and audio feedback for visually impaired individuals. The architecture consists of the following key components:

Smartphone Camera (Input Source) – Captures real-time video and streams it to a laptop using IP Webcam. Video Processing (Laptop) – The laptop receives the video stream and processes it using OpenCV for frame extraction. Object Detection (Deep Learning Model) – A pre-trained neural

network (ResNet50) processes the frames to identify objects. Text-to-Speech Conversion – The detected objects are converted into speech using Pyttsx3 for audio feedback. Audio Feedback to User – The processed speech output is streamed back to the smartphone using Audio Relay. This architecture ensures a real-time, cost-effective solution without requiring additional hardware.

V. Implementation

system's frontend is primarily the smartphone, which serves as the user interface and video input source. The user activates the IP Webcam application, which streams real-time video to the backend, a laptop running the object detection algorithm. The backend processes the video frames using OpenCV and a pre-trained ResNet50 model in PyTorch for object recognition. Once an object is identified, the system uses Pyttsx3 to convert the detected object's name into speech. The processed audio feedback is then sent back to the smartphone using Audio Relay, allowing the user to hear real-time updates about obstacles. This seamless connection between the frontend and backend ensures a smooth and efficient user experience without requiring additional hardware.

VI. Results and Discussion

The developed obstacle detection system effectively identifies objects in real-time and provides auditory feedback to support visually impaired users. Testing in different environments, both indoors and outdoors, demonstrated that the ResNet50 model accurately detects obstacles, particularly under well-lit conditions. However, performance declines in low-light settings or when objects are partially hidden.

The text-to-speech functionality ensures prompt audio feedback, enhancing the user's awareness of their surroundings. However, minor delays were noticed due to network latency in video streaming and audio transmission. Enhancing the processing efficiency and incorporating offline capabilities could further improve responsiveness. Overall, the system offers a practical and affordable solution, highlighting the potential of smartphone-based object detection to assist visually impaired individuals in navigation.

VII. Conclusion

This project presents a cost-effective obstacle detection system for visually impaired individuals using a smartphone and deep learning. By integrating OpenCV, ResNet50, and Pyttsx3, it detects objects in real-time and provides audio feedback, enhancing navigation without additional hardware. While the system performs well in good lighting, challenges like low-light detection and minor latency remain. Future improvements will focus on accuracy, efficiency, and offline functionality, making AI-driven assistive technology more reliable and accessible.

VIII. REFRENCES

1] A smartphone-based application for scale pest detection using multiple-object detection methods JW Chen, WJ Lin, HJ Cheng, CL Hung, CY Lin… - Electronics, 2021 [2] ML] Automatic acne object detection and acne severity grading using smartphone images and artificial intelligence QT Huynh, PH Nguyen, HX Le, LT Ngo, NT Trinh… - Diagnostics, 2022 - mdpi.com[3] Object Detection System for Visually Impaired Persons Using Smartphone D. Ravi Kumar, Hiren Kumar Thakkar, Suresh Merugu, Vinit Kumar Gunjan Suneet K. Gupta [4]Object detection system for visually impaired persons using smartphone DR Kumar, HK Thakkar, S Merugu, VK Gunjan… - … 2020: Proceedings of …, 2022 - Springer[5]Deep learning smartphone application for real-time detection of defects in buildings H Perez, JHM Tah - Structural Control and Health Monitoring, 2021 - 6]Recent trends in smartphone-based detection for biomedical applications: a review S Banik, SK Melanthota, Arbaaz, JM Vaz… - Analytical and …, 2021