EDGE AI ENABLED SMART VISION ASSISTANT FOR VISUALLY CHALLENGED INDIVIDUAL

Raavi Satish[1], Peddinti Keerthana[2], Reddy Sowjanya [3], Mandala Hemanth[4], Paritala Satish[5], Vankala Atchut Rao[6]

[1] *Associate Professor, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Sciences, Andhra pradesh, India*
[2] *Student, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Science, Andhra pradesh, India*
[3] *Student, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Sciences, Andhra pradesh, India*
[4] *Student, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Sciences, Andhra pradesh, India*
[5] *Student, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Sciences, Andhra pradesh, India*
[6] *Student, Electrical and Electronics Engineering, Anil neerukonda Institute of Technology and Sciences, Andhra pradesh, India*

ABSTRACT

Assistive technologies play a crucial role in improving the quality of life for visually impaired individuals. This paper presents an Edge AI-enabled Smart Vision Assistant designed to provide real-time environmental awareness and navigation support. The proposed system utilizes edge computing to process visual data locally, thereby reducing latency, minimizing dependence on internet connectivity, and ensuring enhanced data privacy. A compact camera integrated with an embedded edge processor captures live video streams, which are analyzed using optimized deep learning models for object detection, text recognition, and scene interpretation.

The system employs lightweight convolutional neural networks tailored for resource-constrained devices to achieve efficient and accurate performance. Detected information, including obstacles, signage, and essential objects, is communicated to the user through an audio interface using text-to-speech technology. Additional functionalities such as obstacle avoidance, path guidance, and emergency alert mechanisms further enhance user safety and mobility.

The design emphasizes portability, low power consumption, and affordability, making it suitable for real-world deployment. Experimental evaluations demonstrate that the proposed system achieves reliable performance under varying environmental conditions while maintaining real-time responsiveness. By leveraging Edge AI, the Smart Vision Assistant overcomes limitations of cloud-based assistive solutions and ensures continuous operation in offline scenarios.

This work highlights the potential of integrating artificial intelligence with embedded systems to develop inclusive technologies. The proposed solution aims to empower visually challenged individuals by providing greater independence, situational awareness, and confidence in daily activities.

Keyword: - *Edge AI, Assistive Technology, Computer Vision, Object Detection, Text-to-Speech, Embedded Systems, Visual Impairment, Smart Navigation.*

1. INTRODUCTION

Visual impairment poses significant challenges to individuals in performing everyday activities such as navigation, object recognition, and reading environmental information. Conventional assistive tools, including white canes and guide dogs, provide basic mobility support but lack the capability to interpret complex visual data from the surroundings. As a result, visually challenged individuals often depend on external assistance, which limits their independence and confidence.

Recent advancements in Artificial Intelligence (AI) and embedded systems have opened new possibilities for developing intelligent assistive technologies. In particular, Edge AI enables data processing directly on local devices, reducing latency, ensuring faster response times, and minimizing dependence on cloud infrastructure. This approach also enhances data privacy and allows the system to function effectively in environments with limited or no internet connectivity. In this context, the proposed project presents an Edge AI-enabled Smart Vision Assistant designed to assist visually impaired individuals by providing real-time environmental awareness. The system utilizes a camera integrated with an embedded processor to capture visual input and apply computer vision techniques such as object detection, text recognition, and scene understanding. The extracted information is converted into audio output using text-to-speech technology, allowing users to perceive their surroundings through auditory feedback.

Furthermore, the system incorporates features such as obstacle detection and signage recognition to enhance safe navigation. The design focuses on portability, low power consumption, affordability, and ease of use, making it practical for daily applications.

Thus, the proposed system aims to bridge the gap between advanced AI technologies and accessibility needs, ultimately improving the quality of life for visually challenged individuals.

Assistive technologies for visually impaired individuals have gained significant attention in recent years due to advancements in artificial intelligence and computer vision. Many researchers have proposed systems that enhance navigation and environmental awareness.

It is reported that object detection techniques using deep learning significantly improve the accuracy of real-time assistive systems [1]. These systems enable identification of obstacles and important objects, thereby assisting users in navigation. Furthermore, wearable devices integrated with cameras have shown promising results in providing continuous environmental feedback [2].

As per Kumar et al., Edge AI plays a crucial role in reducing latency and enabling real-time processing on embedded devices without relying on cloud infrastructure [3]. This is particularly important for assistive applications where immediate response is required.

In addition, text recognition technologies such as Optical Character Recognition (OCR) are widely used to read signage and printed text for visually impaired users [4].

It is also observed that combining multiple functionalities such as object detection, text recognition, and speech output improves the usability of assistive systems [5].

However, many existing solutions face challenges related to high power consumption, cost, and dependency on internet connectivity.

Therefore, the proposed Edge AI-enabled Smart Vision Assistant aims to overcome these limitations by providing a low-cost, portable, and efficient solution for real-time assistance.

1.1 Problem Statement

Visually impaired individuals face significant challenges in perceiving and understanding their surroundings, which affects their ability to navigate safely and perform daily activities independently. Traditional assistive tools such as white canes and guide dogs provide limited functionality, primarily focusing on obstacle detection, but they do not offer detailed information about the environment, such as object identification, text recognition, or scene understanding.

Existing smart assistive systems attempt to address these limitations using computer vision and cloud-based processing. However, many of these solutions suffer from high latency, dependency on continuous internet connectivity, and concerns related to data privacy. Additionally, such systems are often expensive, power-intensive, and not optimized for real-time performance on portable devices, making them less practical for everyday use.

Furthermore, the inability to read signage, recognize objects, and receive instant feedback about the surroundings limits the confidence and mobility of visually challenged individuals. There is a need for an efficient, low-cost, and portable solution that can provide real-time assistance without relying heavily on external infrastructure.

Therefore, the problem addressed in this project is the development of an Edge AI-enabled Smart Vision Assistant that can process visual information locally and provide immediate audio feedback. The system aims to enhance environmental awareness, improve navigation, and ensure greater independence for visually impaired individuals while maintaining low latency, affordability, and energy efficiency.

1.2 Objective

The main objective of this project is to design and develop an Edge AI-enabled Smart Vision Assistant that assists visually impaired individuals in understanding their surroundings and navigating safely.

Specific Objectives:

- To develop a system that captures real-time visual data using a camera and processes it using Edge AI techniques.
- To implement efficient computer vision algorithms for object detection, enabling identification of obstacles and important objects in the environment.
- To integrate text recognition for reading signage, labels, and printed text.
- To convert visual information into audio feedback using text-to-speech technology for easy user interaction.
- To design a**,** low-latency system by performing data processing locally on embedded devices without relying on cloud services.
- To ensure the system is portable energy-efficient, and user-friendly for daily use.
- To enhance the safety and independence of visually challenged individuals through real-time assistance.
- To develop a cost-effective solution that is accessible to a wider population.

2. METHODOLOGY

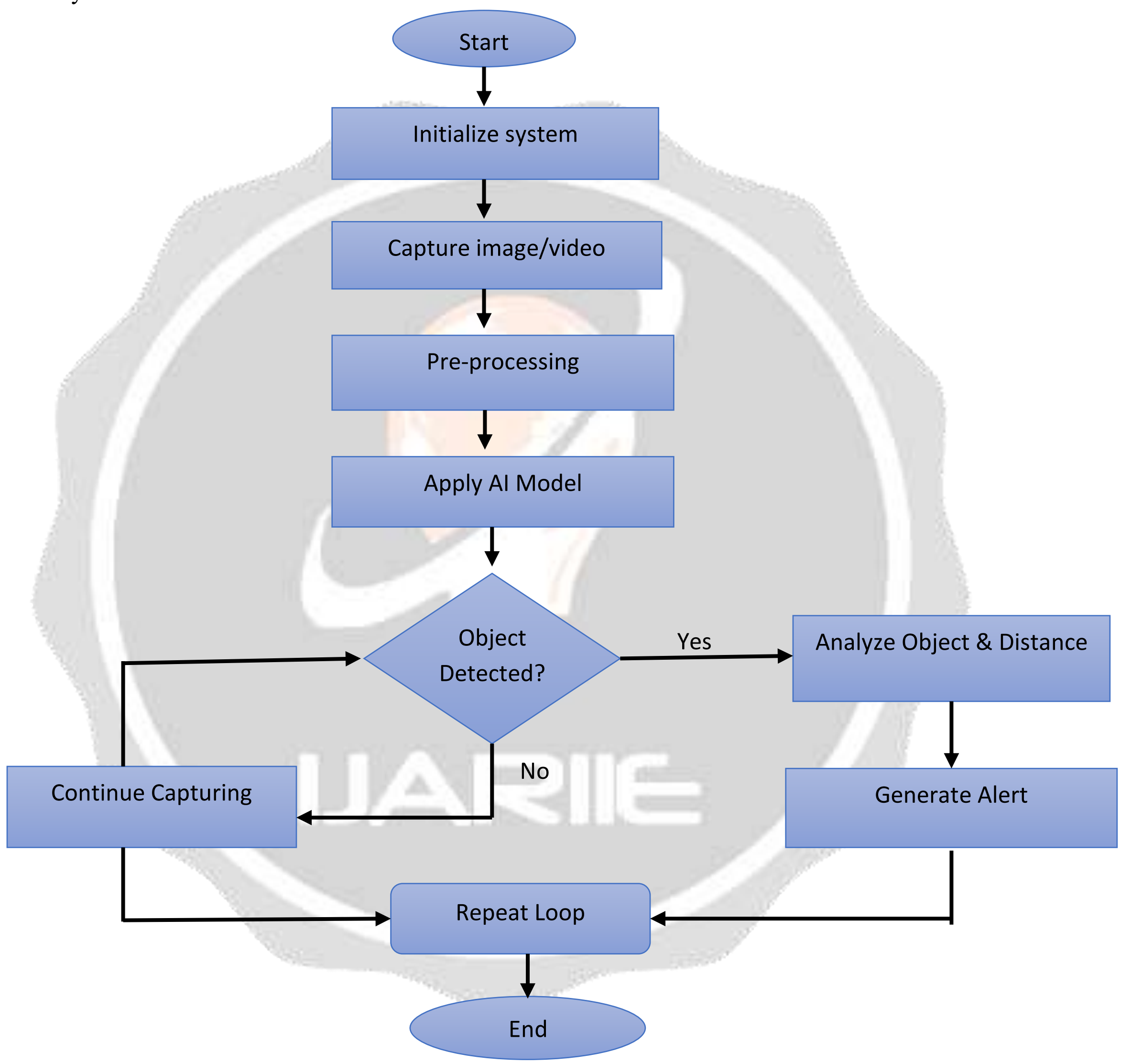
The methodology for the proposed Edge AI-enabled Smart Vision Assistant is structured to ensure real-time assistance for visually impaired users while maintaining portability, low latency, and energy efficiency. The system is divided into the following key phases:

1. Data Acquisition
    - A compact camera module is used to capture live video of the user's surroundings.
    - The captured data includes obstacles, objects, signage, and environmental features relevant to navigation.
2. Edge AI Processing
    - The visual data is processed locally on an embedded device using a high-performance edge processor.
    - Object detection algorithms identify obstacles, people, vehicles, and essential objects in real-time.
    - Text recognition is applied to detect and read signs, labels, and printed text.
    - Lightweight deep learning models, optimized for edge devices, are used to ensure fast processing and low power consumption.
3. Information Interpretation
    - The system interprets detected objects and text in a user-friendly format.
    - Contextual analysis may include distance estimation for obstacles and importance ranking of objects.
4. Audio Feedback
    - Processed information is converted to audio signals using Text-to-Speech technology.

- The user receives real-time spoken guidance about the environment, including object names, directions, and signage content.

5. Navigation Assistance
- Based on object detection and scene analysis, the system provides simple guidance for safe movement.
- Obstacle alerts and emergency notifications are delivered promptly through audio feedback.

6. Optimization and Testing
- The system is tested under different lighting and environmental conditions.
- Models are optimized to minimize latency and power consumption, ensuring portability and long-term usability.



Flow chart: Edge Ai Assistive System For Visually Challenged Individual

2.1 Hardware Components

A) Arduino Uno: The Arduino Uno acts as the central controller for handling all hardware-related operations. It continuously reads inputs from sensors and user controls, processes simple logic, and coordinates outputs like the display and communication modules. While the main vision processing is handled by an edge AI device, the Arduino ensures smooth interaction between physical components in real time.

B) HC-SR04 Ultrasonic Sensor: The HC-SR04 Ultrasonic Sensor is responsible for detecting obstacles in front of the user. It works by emitting ultrasonic sound waves and measuring the time taken for the echo to return after hitting an object. Based on this, it calculates the distance between the user and the obstacle. This information is crucial for avoiding collisions and can be used to trigger alerts when objects are too close.

C) USB Web Camera: A USB Web Camera is used to capture real-time visual data from the surroundings. This camera acts as the vision input of the system, sending continuous video frames to the edge AI processor. The AI model analyse these images to detect objects, recognize scenes, or read text, thereby providing meaningful information to the user.

D) GSM Module: The GSM Module adds a safety feature to the system by enabling communication over a mobile network. When the user presses the emergency button, the GSM module can send an alert message or make a call to a predefined contact, such as a family member or caregiver. This ensures that help can be reached quickly in critical situations.

E) Push Button: The Push Button serves as a simple interface for the user to interact with the system. It can be used to activate or deactivate certain features, or more importantly, to trigger an emergency alert through the GSM module. Its simplicity makes it easy to use, especially for visually impaired individuals.

F) LCD Display: The LCD Display Module provides a visual output of the system's status and detected information. Although the primary feedback for the user is usually audio, the LCD is useful for debugging, monitoring, and assisting caregivers or developers. It can display messages such as detected objects, distance measurements, or system status updates.

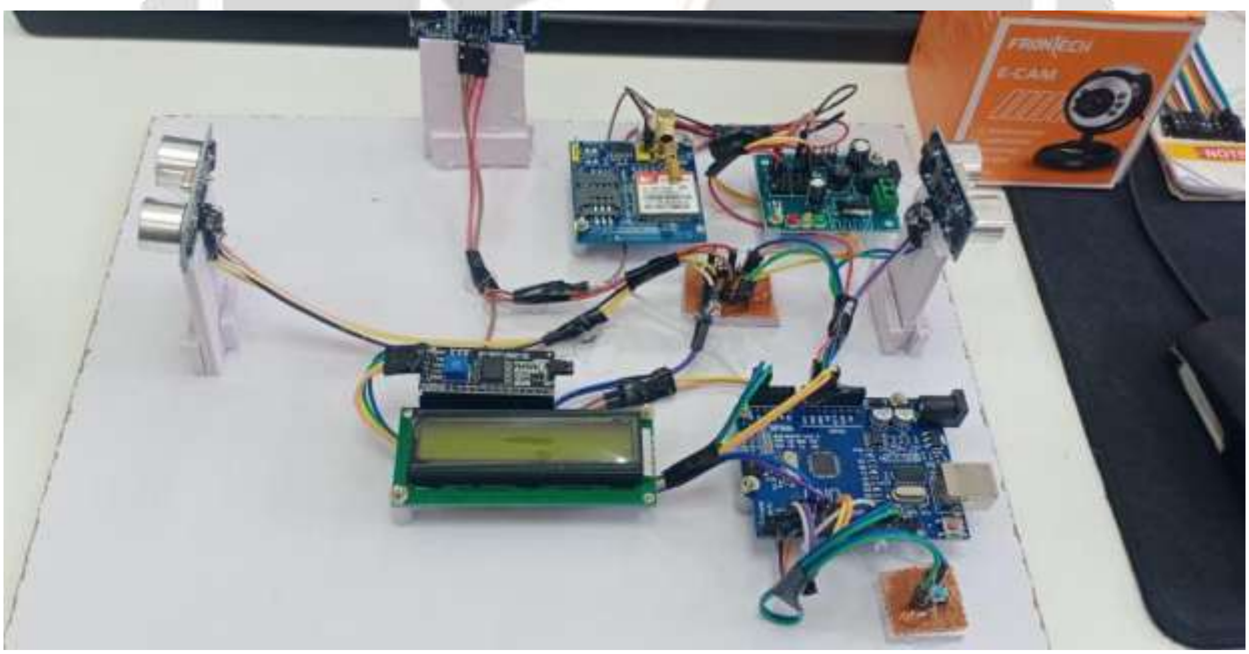


Fig -1: Hardware setup of the proposed model

2.2 System Architecture

The system architecture of Edge AI–enabled smart vision system is designed to process environmental information in real time and convert it into meaningful assistance for visually challenged users. It follows a structured flow where data is captured, processed locally, and then delivered as feedback without relying heavily on the internet.
At the input level, the system begins with the USB Web Camera and the HC-SR04 Ultrasonic Sensor. The camera continuously captures visual data from the surroundings, which serves as the primary input for object detection and scene understanding. At the same time, the ultrasonic sensor measures the distance to nearby obstacles, ensuring immediate awareness of objects that may not yet be identified by the vision model.
The captured visual data is then sent to an edge computing unit, typically a device like a Raspberry Pi 4 or NVIDIA Jetson Nano. This unit forms the core processing layer of the architecture. It runs trained artificial intelligence models such as object detection algorithms, which analyse each frame from the camera to identify objects, people, text, or hazards. Because this processing happens on the device itself (edge AI), the system achieves low latency and can function even without internet connectivity.
Parallel to this, the Arduino Uno manages the sensor and control subsystem. It reads distance values from the ultrasonic sensor, monitors the Push Button for user interaction, and controls output peripherals. The Arduino communicates with the edge device when necessary, ensuring that both sensing and intelligent processing work together smoothly.
Once the data is processed, the system moves to the output layer. The results from the AI model and sensor readings are converted into user-understandable feedback. Audio output is typically the primary mode, where detected objects or warnings are conveyed through speech. Additionally, the LCD Display Module shows text information such as object names, distance values, or system status, which is helpful for developers or assisting individuals nearby.
For safety and emergency communication, the GSM Module is integrated into the architecture. When triggered, usually through the push button, it sends alert messages or calls to predefined contacts, ensuring that the user can quickly reach help when needed.
Overall, the architecture can be understood as a layered system: the input layer collects environmental data, the processing layer (edge AI + microcontroller) interprets it, and the output layer communicates the results to the user. This integration of vision, sensing, and communication technologies creates a responsive and reliable assistive system that enhances navigation and safety for visually challenged individuals.

3. DEEP LEARNING BASED OBJECT DETECTION USING YOLO

The use of YOLO plays a central role in enabling real-time smart vision for visually challenged individuals. YOLO, which stands for *"You Only Look Once,"* is a deep learning–based object detection algorithm designed to identify multiple objects in an image quickly and efficiently. This makes it highly suitable for assistive systems where fast response is critical. The USB Web Camera continuously captures live video from the user's surroundings. These video frames are sent to an external processing unit (such as a laptop), where the YOLO model is deployed. Unlike traditional methods that process images in multiple steps, YOLO analyse the entire image in a single pass, dividing it into grids and predicting bounding boxes along with class probabilities. This allows the system to detect objects like people, vehicles, obstacles, and everyday items almost instantly.
Once YOLO processes the image, it generates outputs such as object labels (for example, "person," "chair," or "car") and their positions in the frame. These results are then transmitted to the Arduino Uno, which integrates them with sensor data. At the same time, the HC-SR04 Ultrasonic Sensor provides precise distance measurements for nearby obstacles. This combination ensures that the system not only recognizes objects but also understands how close they are to the user.
The detected information is then converted into meaningful feedback. Audio output informs the user about detected objects in real time, such as "person ahead" or "vehicle approaching." Additionally, the LCD Display Module can show the detected object names and distances for monitoring purposes. If a dangerous situation is detected or the user presses the Push Button, the GSM Module can send an alert message to a caregiver. The main advantage of using YOLO in your project is its speed and accuracy, which are essential for real-time assistance. Since the system

is designed for visually impaired users, even a small delay could affect safety. YOLO's ability to process frames quickly ensures immediate feedback, making navigation safer and more reliable.

Overall, YOLO acts as the "intelligent vision engine" of your system, transforming raw visual data into actionable information. When combined with the Arduino-based hardware setup and sensors, it creates a powerful assistive solution that enhances environmental awareness and independence for visually challenged individuals.

Fig -2: Real-Time Object Detection Using YOLO in Assistive System

3.1 Result & Discussions

The proposed smart vision system was successfully implemented using the YOLO algorithm integrated with hardware components such as the Arduino Uno, HC-SR04 Ultrasonic Sensor, GSM module, and a USB web camera. The system was tested in real-time environments to evaluate its performance in assisting visually challenged individuals.

The experimental results demonstrate that the YOLO-based object detection module is capable of identifying multiple objects such as humans, vehicles, and common obstacles with high accuracy and minimal latency. The system was able to process live video frames efficiently and provide real-time audio feedback, enabling users to understand their surroundings effectively. The use of YOLO significantly improved detection speed compared to traditional methods, making it suitable for assistive applications.

The integration of the ultrasonic sensor enhanced the reliability of the system by providing accurate distance measurements for nearby obstacles. This ensured immediate alerts when objects were within a critical range, thereby improving user safety. The Arduino Uno effectively coordinated the sensor inputs and controlled output devices such as the LCD display and GSM module. The GSM module was tested for emergency communication, and it successfully transmitted alert messages to predefined contacts when triggered using the push button. This feature adds an additional layer of safety, particularly in critical situations where the user may require assistance.

Despite the successful implementation, certain limitations were observed. The performance of the YOLO model depends on lighting conditions and camera quality, which may affect detection accuracy in low-light environments. Additionally, since the processing is performed on an external system, there may be minor delays due to data transmission between the processing unit and the Arduino.

Overall, the proposed system demonstrates an effective combination of deep learning and embedded systems to assist visually impaired individuals. The integration of real-time object detection, obstacle sensing, and emergency communication provides a reliable and practical solution for enhancing independent navigation. Future improvements can focus on optimizing the model for low-light conditions, integrating edge AI hardware for standalone operation, and enhancing system portability.

4. CONCLUSIONS

This paper presents the design and implementation of an Edge AI–enabled smart vision system for visually challenged individuals using the YOLO algorithm. The proposed system integrates both software and hardware components, including the Arduino Uno and HC-SR04 Ultrasonic Sensor, to provide real-time environmental awareness and assistance.

The experimental results indicate that the system is capable of performing efficient object detection with minimal latency, thereby enabling timely audio feedback to the user. The incorporation of ultrasonic sensing enhances obstacle detection accuracy, while the GSM module ensures reliable emergency communication. The Arduino Uno effectively coordinates the operation of all peripheral components, contributing to the overall system stability.

The proposed model offers a low-cost, portable, and efficient solution for assisting visually impaired individuals in navigating their surroundings independently. However, the system performance may vary under low-light conditions and depends on the efficiency of the external processing unit used for implementing the YOLO algorithm. Future work can focus on improving system robustness by integrating dedicated edge AI hardware, optimizing the detection model, and incorporating advanced features such as voice interaction and GPS-based navigation. Overall, the system demonstrates the potential of combining deep learning techniques with embedded systems to develop practical assistive technologies.

5. ACKNOWLEDGEMENT

We are greatly indebted to Anil Neerukonda Institute Of technology and Sciences for providing a conducive environment that has motivated me to pursue and achieve my academic and project goals. I would like to express my sincere thanks to my project guide,.Dr.R.Satish, for his invaluable guidance, intellectual insights, constant encouragement, and support throughout the project. His novel approach to ideas and dedication helped me successfully navigate through the challenges of this project. We also express our sincere gratitude to our coordinator, Dr.V. Sandeep Reddy, for her leadership and continuous motivation, which was essential in completing our academic semester successfully. His guidance ensured that we stayed
focused and on track throughout the duration of the project.

6. REFERENCES

[1]. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi,"*You Only Look Once: Unified, Real-Time Object Detection*,"Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[2]. Anind K. Dey et al "*Context-Aware Computing Applications in Assistive Technologies*,"IEEE Pervasive Computing, vol. 10, no. 3, pp. 32–39,2011*(Used for wearable and assistive system concept support)*

[3]. Kumar et al "*Edge AI: Enabling Technologies and Applications*,"IEEE Internet of Things Journal, vol. 7, no. 5, pp. 1–10, 2020.

[4]. Tesseract OCR,"*An Open SourceOCR Engine*,"

[5]. World Health Organization,"*Assistive Technology for Persons with Disabilities*,"WHO Report, 2018.