Grenze International Journal of Engineering and Technology, January Issue

A Smart Glass for Visually Impaired People

S. Pravinthraja[1], Asif Mohamed H.B[2], M.Sinthuja[3] and Amirtha Preeya V[4]
[1, 2, 4]Presidency University /CSE, Bangalore, India
[3]M.S.Ramaiah Institute of Technology/ISE, Bangalore, India
Email: {pravinthraja@gmail.com, asif.mohamed@presidencyuniversity.in, sinthujamuthu@gmail.com, Amirthavme@gmail.com}

***Abstract*—This research paper introduces a ground breaking solution aimed at improving the daily lives of visually impaired individuals. This research paper developed a Smart Glass system equipped with advanced technologies such as obstacle detection, object recognition, and transformation of visual information into audible form. This research paper is mainly focusing on people with visual impairments and more specifically their education life. It is presenting a concept of smart glasses to provide assistance in multiple tasks represented as modes to be chosen by the user. This innovative approach utilizes a combination of hardware and software to empower individuals with visual impairments, promoting independence and enhancing their overall quality of life.**

***Index Terms*— Smart Glass System, obstacle detection, object recognition, Visual-to-Audio transformation.**

I. INTRODUCTION

The "Smart Glass for Visually Impaired" project is a dedicated effort aimed at enhancing the independence, safety, andinclusivity of individuals with visual impairments. At its core, the project encompasses three essential components: Obstacle Detection, Object Detection, and Audio Form Conversion. These elements seamlessly integrate advanced sensor technologies and audio feedback to provide real-time information about the user's environment, addressing critical challenges faced by the visually impaired. The first component, Obstacle Detection, utilizes ultrasonic sensors emitting sound waves to detect obstacles. By measuring the time it takes for the waves to return, the system calculates distances and triggers alerts when obstacles are within a predefined range. This functionality ensures users receive timely warnings,
enhancing their safety and spatial awareness.
Object Detection, the second component, employs sophisticated technologies such asTensorFlow Lite. Using a camera, the system captures visual data, and a trained model analyzes it to recognize and classifyobjects. This real-time object recognition empowers users with a deeper understanding of their surroundings, contributing to improved navigation andenvironmental awareness. The third component, Audio Form Conversion, employs the pyttsx3 library forText-to-Speech conversion. Recognized obstacles and detected objects are transformed into audible feedback, providing users with spoken descriptions inreal-time. This audio feedback is deliveredthrough earphones, ensuring a discreet anduser-friendly experience that enhances accessibility. Beyond its technological sophistication, theproject's significance lies in its impact on inclusivity. The seamless integration of advanced sensor technologies and audio feedback not only addresses practical challenges but also fosters a sense of independence among visually impairedindividuals. By providing a tool that facilitates full engagement with the

Grenze ID: 01.GIJET.11.1.304_1

world, the project represents a transformative steptoward creating a society where everyone, regardless of their abilities, can activelyparticipate and contribute.

II. Literature Review

Samiur Rahman et al. (2018) presents a mobile camera system to enhance mobility and safety for the visually impaired. Using pre-stored floor images, their algorithm detects obstacles and floor type changes in real-time with a reported accuracy of 96%, offering potential benefits for improved accessibility in indoor environments[1].

Target detection in computer vision, a key research focus for two decades, rapidly identifies and locates predefined objects in images. Categorized into single-stage and two-stage detection algorithms based ontraining methods, this paper delves into representative models. It further examines commonly used datasets, analyzes various algorithms, and outlines potentialchallenges in target detection[2].

This research addresses the challenges of traditional blind canes by introducing a computer vision-based mobility aid. Utilizing Connected Component Labeling and Min-max threshold, the device achieves an 82.89% accuracy in obstacle detection. Unlike standard canes, this innovative approach aims to enhance blindindividuals' mobility without causingdisturbance to their surroundings [3].

This monograph provides a historical perspective on the impactful rise of deep learning in artificial intelligence, particularly in object recognition, detection,and segmentation within computer vision. Covering applications like image classification, face recognition, and semantic segmentation, it elucidates deep learning's superiority through concrete examples, making it essential for newcomers in signal processing and machine learning research [4].

This paper addresses challenges in autonomous driving object detection. It introduces an innovative hybrid system (LM-CNN-SVM) that combines Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for improved robustness, demonstrating superior performance in object recognition and pedestrian detection compared to existing methods.[5]. This article addresses the evolution of object detection, emphasizing the limitations of earlier techniques relying on hand-crafted features. It introduces an end-to-end solution using the Single Shot Detector (SSD) method to improve accuracy in deep learning-based object detection.[6]

This study addresses the challenge of objectdetection for visually impaired individuals,presenting a model using the Caffe model framework and Mobile net SSD method. The system converts objects to text and then speech, aiding real-time detection for visually impaired individuals with potentialapplications in portable devices[7].This paper surveys recent advancements in deep learning for object detection, covering frameworks, feature representation, proposal generation, context modeling, training strategies, and future research directions, comprising over300 contributions.[8] This paper addresses object recognition challenges using YOLO and YOLO_v3 algorithms, promoting real-world applications for visually impairedindividuals. Tested on MS-COCO Dataset,it employs gTTS and pygame for audio feedback [9][10][11].

This paper introduces "Third Eye," utilizingYOLOv5 for accurate object detection and gTTS, pyttsx3 for speech generation. Tailored for the visually impaired, it outperforms existing approaches, training on a custom dataset.[10]. This paper explores the progress in computer vision, highlighting a substantial increase in accuracy over the past decade. Focused on assisting the visually impaired, it utilizes Raspberry Pi for real-time object detection and recognition, providing voice feedback to enhance independence.[12][13].

III. Methodology

A. Object Recognition withTensorFlow

Data Capture: Visual data is collected through a camera, producing images or video frames.

Model Training with TensorFlow: A TensorFlow model, perhaps based on a pre- trained architecture like MobileNet orInception, is trained on a labeled dataset to recognize object classes.

Inference with TensorFlow: In real-time, the trained TensorFlow model analyzes newvisual data and predicts the probability distribution of object classes.

Textual Conversion: Information about detected objects, including class labels andpositions, is converted into textual form.

B. Object Detection withTensorFlow

Bounding Box Prediction: TensorFlow supports models for object detection.

Localization: The TensorFlow object detection model localizes each object, specifying its position through boundingbox coordinates.
Class Labeling: Object detection models using TensorFlow assign class labels, indicating the type of object present in eachbounding box generating spoken descriptions of the detected objects.
Headphone Output: The audio output is directed to headphones, providing a user- friendly way to listen to real-timeinformation about the environment.

C. Audio Output with pyttsx3

Text-to-Speech Conversion: The textual information about recognized and detected objects is passed to the pyttsx3 module.
Audio Feedback: pyttsx3 converts the textual information into audible speech,
Textual Conversion: The recognized object classes are converted into textualinformation.

1.Start:
The journey begins here.
2.1 *Camera (RecognizingCamera):*
Captures visual data using the connectedcamera module.
2.2 *Ultrasonic Sensor (MeasuringDistance):*
Begins by measuring the distance from thesensor.
The sensor detects obstacles in the physicalenvironment.
3. *TensorFlow (Detecting Object):*
Utilizes TensorFlow, a powerful tool forobject detection.
Analyzes the visual data to identify and locate objects in the camera's field of view.
4. *Fetching Classes of Object:*
Gathers information about the recognized objects.
Determines the class or category to which each detected object belongs.
5. *Text to Speech (TTS Module):*
Converts the identified object information into spoken words.
Utilizes a text-to-speech (TTS) module, possibly pyttsx3, to generate the corresponding audio.
6. Headphone:
Directs the audio output to the headphone.
7. Audio Output:
Listens to the spoken information about the detected object through the connected headphone.This flowchart in Fig.1 illustrates the sequential steps involved in capturing, recognizing, and conveying information about objects inthe environment, creating a seamless interaction between the visual and auditorysenses.

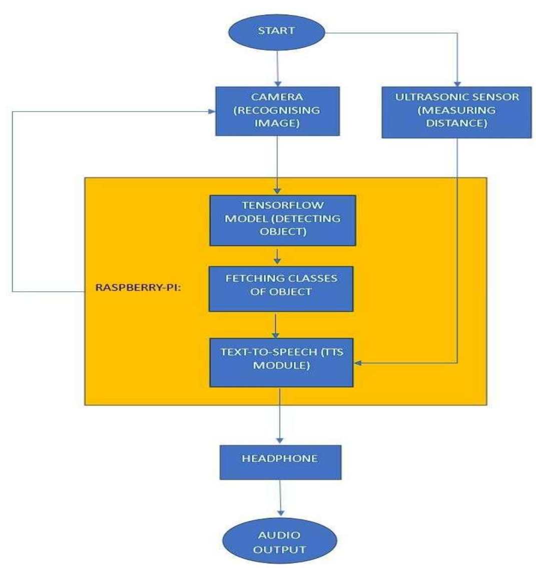

Fig 1: Flowchart of the model

1.Ultrasonic Sensor:
Connected to Raspberry Pi 3B+ at a specificpin.
Responsible for obstacle detection in the physical environment.
2. *Camera Module:*
Connected to Raspberry Pi 3B+.
Handles object recognition and objectdetection through visual data.
3. *Audio Output:*
Utilizes the pyttsx3 module.
Enables listening to detected objects, translating digital information into audible feedback.
4. *Power Supply:*
Provided by a 5V battery. Sustains the Raspberry Pi 3B+, empowering it to execute tasks related to obstacle detection, object recognition, and audio output. Fig 3. Depicts the hardware model of the proposed model.

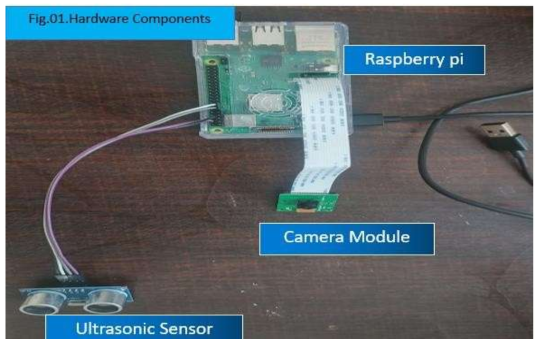

Fig 2: Hardware model

D. Design Procedure

Data Collection:

- Collect data from various environments to ensure the object detection model is robust in different scenarios.
- Include a diverse set of objects to enhance the model's ability to recognize a wide range of items.
- Ensure the dataset includes examples of obstacles to improve the obstacle detection system.
- Annotate the dataset meticulously, providing clear labels for object detection and obstacle scenarios.

Model Training:

- Use transfer learning techniques to leverage pre-trained models, optimizing the training process.
- Implement data augmentation strategies to increase the model's ability to generalize.
- Fine-tune the model based on initial training results and user feedback.
- Establish a continuous training schedule to keep the model updated with new data and improve its performance over time.

Sensor Integration:

- Implement a feedback mechanism to inform users when obstacles are detected.
- Test and calibrate ultrasonic sensors to ensure accurate and reliable obstacle detection.

Audio Transformation:

- Explore different Text-to-Speech (TTS) voices and speeds to enhance user experience.
- Implement a prioritization system for auditory alerts to convey critical information first.
- Fine-tune the audio transformation system based on user preferences and usability testing.
- Conduct user studies to ensure that audio cues are clear, understandable, and effectively convey information.

IV. IMPLEMENTATION DETAILS AND RESULTS

1. Ultrasonic Sensor for Obstacle Detection:

Hardware Setup

Sensor Selection: We have chosen and connected an ultrasonic sensor (HC-SR04) to hardware platform at respective pins.

Calibration: The sensor is calibrated to provide accurate distance measurements within your required range.
Integration: Sensor readings are continuously captured to detect obstacles inreal-time.

2. ***TensorFlow for ObjectDetection:***

Model Integration

Pre-trained Model: A pre-trained TensorFlow model for object detection is integrated into the system.
Object Detection Algorithm: The model isconfigured to identify objects within the camera's field of view.
Real-time Processing: Object detection is performed continuously on each frame captured by the camera.

3. ***OpenCV for Computer Vision:***

Computer Vision Integration

Camera Capture: OpenCV is used to capture video frames from the camera, providing input for both obstacle detectionand object detection.
Pre-processing: Frames may undergo pre-processing steps such as resizing or normalization before being fed into the object detection model.
Visualization: OpenCV is utilized to visualize the results of both obstacle and object detection, drawing bounding boxes and labels.

4. ***Pyttsx3 for Audible Feedback:***

Text-to-Speech Integration

Textual Information Generation: Objectand obstacle information is co
Audible Alerts: Pyttsx3 is employed to convert the textual information into audiblespeech.
Alerts Differentiation: Different audible alerts are generated based on the severity ortype of detected objects and obstacles.

5. ***Real-time Implementation:***

Continuous Processing

Loop Structure: The entire systemoperates within a continuous loop, ensuringreal-time processing of sensor readings, object detection, and audible feedback.
Feedback Mechanism: Audible alerts are continuously provided based on changing environments, adjusting the alerts as needed.

6. ***Safety Considerations:***

Alert Prioritization: Audible alerts are prioritized based on the proximity of obstacles detected by the ultrasonic sensor.
User Interaction: The system is designed to provide clear and understandable audiblealerts for different levels of obstacle and object proximity.

7. ***Testing and Optimization:***

Testing

Integration Testing: The systemundergoes thorough integration testing to ensure accurate object detection, obstacle detection, and reliable audible alerts.
Field Testing: Real-world testing is conducted in various scenarios to validate the system's effectiveness.

Optimization

Performance Tuning: Code optimization is performed for enhanced performance, considering factors like frame rate, latency,and resource utilization.
Calibration Adjustment: Periodic checks and adjustments are made to the ultrasonic sensor calibration to account for changes inthe environment.

TABLE 1. FOR DETECTED OBJECT WITH ACCURACY

Object	Detected Accuracy	False Rate	Existing Accuracy	Improvement
Keyboard	72%	3%	63%	+9%
Table	71%	4%	62%	+9%
Chair	73%	2%	61%	+12%
Laptop	74%	2%	60%	+14%
Person	70%	3%	64%	+6%
Bag	75%	4%	62%	+13%
Bottle	72%	2%	65%	+7%
Bed	71%	3%	63%	+8%

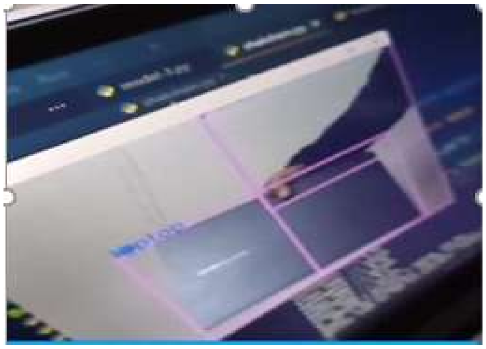

Fig 3: Detection of Laptop

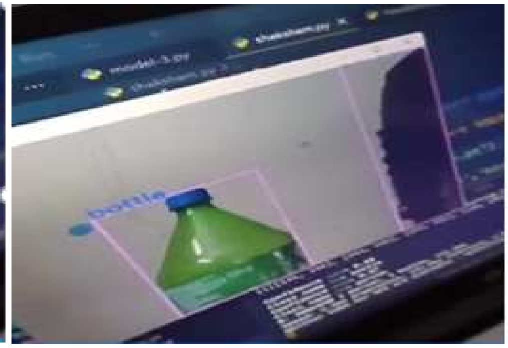

Fig 4: Detection of Bottle

V. CONCLUSION

Smart Glass for Visually Impaired project isa significant step forward in improving the lives of people with visual impairments. This advanced solution integrates technologies like obstacle detection, object recognition, and auditory feedback to enhance user independence and safety. The obstacle detection feature provides real-time alerts about potential hazards, giving users more confidence in navigating their surroundings, especially in dynamic environments. The object recognitioncapability goes beyond just identifying obstacles; it allows the Smart Glass to describe various objects, enabling users to interact more effectively with their surroundings. The emphasis on auditory feedback makes the Smart Glass user-friendly by converting visual information into audible feedback. This not only helps with immediate decision-making but also promotes independence by reducing the need for external assistance. Looking ahead, technology's rapid evolution holds the promise of even more sophisticated smart assistive devices. Anticipated advancements in sensors, machine learning, and integrationcapabilities could lead to a future where individuals with visual impairments can navigate and engage with the world on equal terms.

REFERENCES

[1] Samiur Rahman et al 2018 J. Phys.:Conf. Ser. 960 012046. Jun Deng et al 2020 J. Phys.: Conf. Ser. 1684 012028

[2] Shahani S and Gupta N. (2023). The Methods of Visually Impaired Navigating and Obstacle Avoidance 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC). 10.1109/ICAISC58445.2023.10200915.

[3] Xiaogang Wang (2016), "Deep Learning in Object Recognition, Detection, and Segmentation", Foundations and Trends® in Signal Processing: Vol. 8: No. 4, pp 217- 382. http://dx.doi.org/10.1561/2000000071

[4] Uçar A, Demir Y, Güzeliş C. Object recognition and detection with deep learning for autonomous driving applications. SIMULATION. 2017;93(9):759-769.

[5] Vaishnavi, K., Reddy, G. P., Reddy, T. B., Iyengar , N. C. S., & Shaik , S. (2023). Real-time Object Detection Using Deep Learning. Journal of Advances in Mathematics and Computer Science, 38(8), 24–32.

[6] M. Sinthuja, Dr. Pravinthraja, Dr. S.R Manisekar, Karthik V (2023) “Currency Recognition App for Blind people" 2nd International Conference on Human-Centric Smart Computing (ICHCSC 2023) New Delhi, India on July 05th-06th. eBook ISBN 978-981-99-77116, February 24, Series ISSN2190-3018,

[7] C. Sagana et al., "Object Recognition System for Visually Impaired People," 2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Nitte, India, 2021, pp. 318-321, doi: 10.1109/DISCOVER52564.2021.9663608.

[8] K. Sheela sobana Rani, S Pravinth Raja, M. Sinthuja, B Vidhya Banu, R. Sapna, Kenenisa Dekeba, "Classification of EEG Signals Using Neural Network for Predicting Consumer Choices", Computational Intelligence and Neuroscience, vol. 2022, Article ID 5872401, 7 pages, 2022.

[9] Koppala Guravaiah, Yarlagadda Sai Bhavadeesh, Peddi Shwejan, Allu Harsha Vardhan, S Lavanya,Third Eye: Object Recognition and Speech Generation for Visually Impaired,Procedia Computer Science,Volume 218,2023,Pages 1144-1155.

[10] M. Sinthuja, Pravinthraja, S.Sheeba Rachel, Suganthi.P, “Vision Enhancement using Image”, Hinweis Second International Conference on Networking, Communication and Computing- NCCT-2023, Ooty, India on August 25 -26.

[11] Charlin Pushpalatha, M. Sinthuja, Pravinthraja, Amutha Raj, Sridhar, Sheeba Rechal, T. Anitha, T. Rajendran, Yosef Asrat Waji. “Segmentation and Classification of Clustering using U-Net with Deep Learning Model", Journal of Healthcare Engineering, vol. 2022, Article ID 1601354,pp 1-10, Feb 2022.