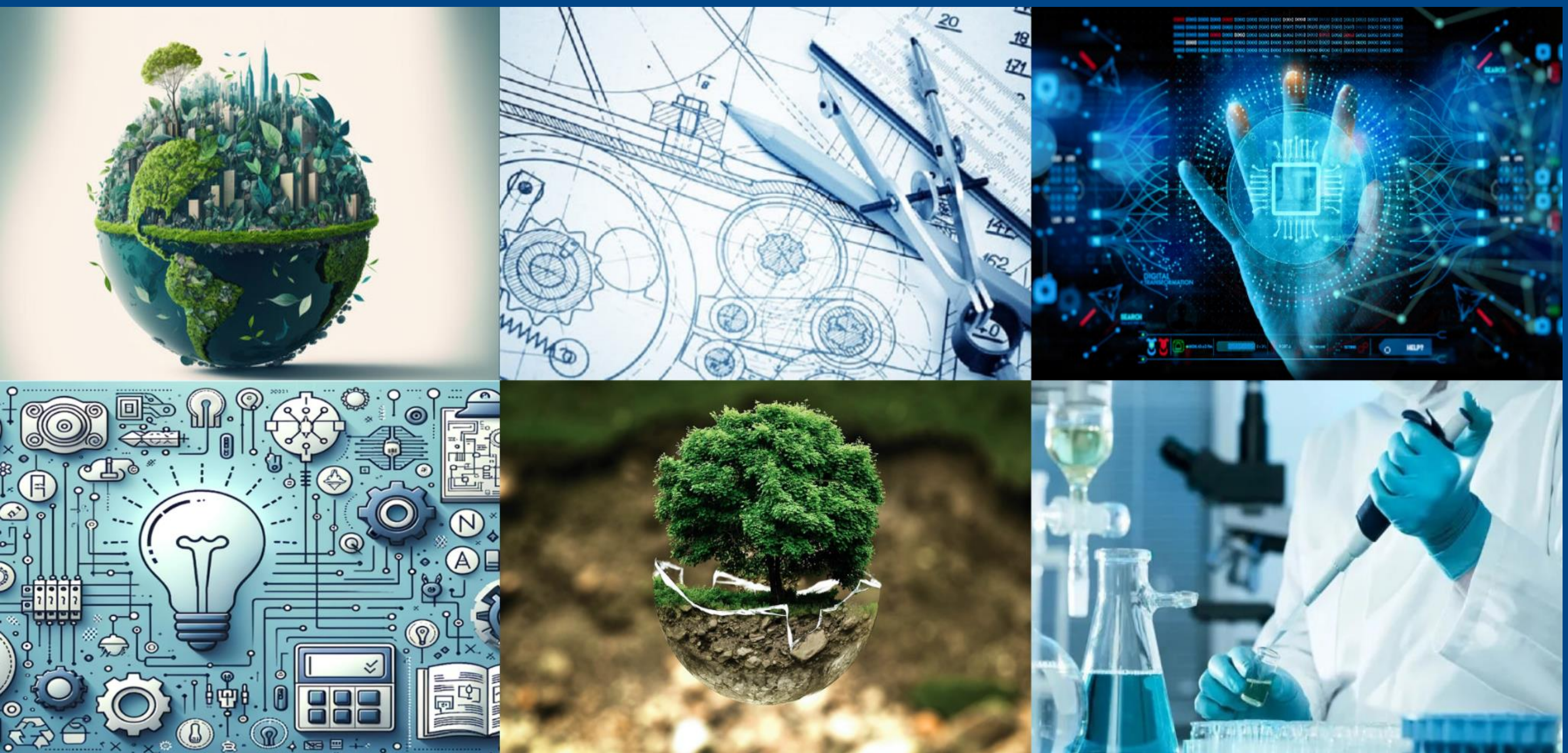
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

R.E.T.I.N.A. Real-Time Environmental Transcription and Interactive Navigation Assistant for Visually Impaired

Najiya Abdulrahiman, Mohammed Safwan K K , Brillons, Adarsh Narayanan , Fiza Rahmathullah

Assistant professor, Dept. of Artificial Intelligence and Data Science, Royal College of Engineering and Technology, Akkikavu, Thrissur, Kerala, India

PG Student , Dept. of Artificial Intelligence and Data Science, Royal College of Engineering and Technology, Akkikavu, Thrissur, Kerala, India

PG Student , Dept. of Artificial Intelligence and Data Science, Royal College of Engineering and Technology, Akkikavu, Thrissur, Kerala, India

PG Student, Dept. of Artificial Intelligence and Data Science, Royal College of Engineering and Technology, Akkikavu, Thrissur, Kerala, India

ABSTRACT: RETINA (Real-time Environmental Transcription and Interpretation for Navigation and Assistance for Visually Impaired) is an innovative project aimed at enhancing the navigation and environmental awareness of visually impaired individuals through advanced assistive technology. The system integrates a portable, headband-mounted camera with a Raspberry Pi, enabling real-time capture and interpretation of visual information from the user's surroundings. Utilizing cutting-edge computer vision and machine learning algorithms, along with the Google Coral Accelerator for rapid processing, RETINA facilitates swift object recognition and detailed auditory feedback. The Mobilenet tflite model plays a crucial role in generating accurate, contextually relevant descriptions of identified objects, enhancing the quality of information provided to users.

By continuously capturing video footage and processing it in real-time, RETINA delivers immediate audio cues about objects, obstacles, and environmental features through connected earphones. This innovative approach addresses key challenges faced by visually impaired individuals, such as limited environmental awareness and reliance on external assistance, ultimately promoting greater independence and safety. With a functional prototype already developed, RETINA stands at the forefront of assistive technology, potentially transforming the lives of visually impaired users by enabling more confident navigation and interaction with their surroundings.

KEYWORDS: Visually impaired, assistive technology, object recognition, Raspberry Pi, computer vision, Google Coral Accelerator, MobileNet V2, TensorFlow Lite, YOLO, Edge TPU.

.

I. INTRODUCTION

Visually impaired individuals in Kerala face significant challenges in navigation, with approximately 350,000 visually impaired people, including 35,000 who are blind. To address this, the RETINA (Real-time Environmental Transcription and Interpretation for Navigation and Assistance) system has been developed, integrating several hardware components to enhance environmental awareness. RETINA uses a headband-mounted camera, compatible with a Raspberry Pi 4 Model B, to capture live video from the user’s perspective. The Raspberry Pi processes the data, while the Google Coral USB Accelerator, connected via USB 3.0, offloads computationally intensive tasks using the Edge TPU for real-time object detection. MobileNet V2, optimized for edge devices, is deployed on the system via TensorFlow Lite to ensure efficient, low-latency detection of objects.

The system provides real-time feedback through bone conduction earphones, delivering auditory cues about detected objects and obstacles, without obstructing the user's surrounding sounds. A built-in microphone captures voice commands for interaction with the system. Powered by a portable power bank, RETINA enables continuous operation,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

ensuring safety and autonomy for visually impaired users. This project combines advanced assistive technology with hardware-specific optimizations, offering a more accessible, affordable, and efficient solution for visually impaired individuals, improving their mobility and quality of life in both urban and rural environments.

II.TECHNICAL SPECIFICATIONS

For the RETINA project, which enhances navigation and environmental awareness for visually impaired individuals using real-time object recognition and audio feedback, the following technical specifications are required:

Hardware:

- Raspberry Pi: A Raspberry Pi 4 Model B with 4GB RAM is recommended for its computational power to handle object recognition and real-time video processing tasks.
- Operating System: Install a compatible operating system like Raspbian or Ubuntu on the Raspberry Pi. Ensure the OS is updated to the latest version for security and stability.
- Camera: A lightweight mini camera mounted on a headband captures live video from the user's perspective. The camera should be compatible with the Raspberry Pi and capable of delivering high-quality video for processing.
- MicroSD Card: A microSD card with sufficient storage (at least 32GB) is required for the operating system, software libraries, and model data.
- USB Webcam: A USB webcam is used to capture live video footage from the user’s perspective. The webcam should be compatible with the Raspberry Pi and capable of providing high-quality video for real-time processing and object detection.
- USB Hub: A USB hub is needed to connect peripherals like the camera, microphone, and earphones to the Raspberry Pi.
- Power Supply: A reliable power supply, with a sufficient battery backup, ensures continuous operation for extended periods. Considerations for power efficiency are crucial for maximizing battery life.
- Earphones with built-in Microphone: A microphone captures the user’s voice commands, while bone conduction earphones provide audio output without obstructing the user's hearing, ensuring they receive auditory feedback about their environment.

Software Libraries:

- OpenCV: For real-time image processing and object detection.
- TensorFlow: To implement machine learning models like MobileNet V2 for efficient object classification.
- TensorFlow Lite: Optimized for edge devices, TensorFlow Lite enables faster inference of models such as MobileNet V2 when integrated with the Coral Accelerator.
- Text-to-Speech (TTS) Library: Converts recognized text into speech for auditory feedback.
- Speech Recognition Library: Enables the system to interpret user voice commands for interactivity.
- Threading Library: The threading library is used to manage multiple tasks concurrently, allowing real-time video capture, object detection, and audio feedback to run in parallel without blocking system performance.

Model Integration:

- YOLO (You Only Look Once): Used for real-time object detection, ensuring efficient classification of objects in the user's environment.
- MobileNet V2: A lightweight and efficient convolutional neural network (CNN) model used for real-time object detection. MobileNet V2 provides a balance between speed and accuracy, making it suitable for deployment on edge devices like the Raspberry Pi.

- TensorFlow Lite & Coral Accelerator: The TensorFlow Lite model of MobileNet V2, when integrated with the Google Coral USB Accelerator, utilizes the Edge TPU for accelerated inference, significantly improving processing speed and efficiency in object detection tasks.

III. GOOGLE CORAL USB ACCELERATOR

The Google Coral USB Accelerator plays a pivotal role in the RETINA system by significantly enhancing the real-time object detection capabilities through hardware-accelerated processing. Designed specifically for edge AI applications, the Coral Accelerator integrates seamlessly with the Raspberry Pi, enabling efficient, low-latency inference that is crucial for assisting visually impaired individuals in dynamic environments.

Fig 1: Google Coral Edge TPU

Key Features of the Google Coral USB Accelerator

Edge TPU (Tensor Processing Unit):
The Coral Accelerator is powered by Google's proprietary Edge TPU, which is optimized for running TensorFlow Lite models with minimal power consumption and high-speed inference. The Edge TPU delivers up to 4 trillion operations per second (TOPS) while consuming very little power, making it ideal for portable assistive devices like RETINA.

- *Plug-and-Play Compatibility:*
  The device connects via USB 3.0, ensuring high-speed data transfer and compatibility with various edge devices, including the Raspberry Pi 4. The plug-and-play nature of the Coral Accelerator allows RETINA to achieve enhanced performance without requiring extensive system modifications.
- *Real-time Processing Capabilities:*
  By offloading computationally intensive tasks such as object detection and classification to the Coral TPU, RETINA can achieve processing speeds of over 30 frames per second (FPS), which is a significant improvement compared to CPU-only processing. This ensures that users receive instant audio feedback about their surroundings without perceptible delays.
- *Optimized for TensorFlow Lite Models:*
  The Coral Accelerator is specifically designed to work with TensorFlow Lite models, such as MobileNet V2, which is used in RETINA for lightweight and efficient object detection. The combination of the Edge TPU and TensorFlow Lite results in a highly optimized system capable of handling real-time vision tasks without straining the Raspberry Pi's resources.
- *Energy Efficiency:*
  Given the wearable nature of the RETINA system, power efficiency is a crucial consideration. The Coral USB Accelerator operates with low power consumption while maintaining high processing efficiency, ensuring prolonged battery life and uninterrupted assistance for users.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. SYSTEM DESIGN AND METHODOLOGY

Research and Model Selection:

Extensive research was conducted to define system requirements and assess potential hardware configurations for optimal performance. Initial experiments using the YOLOv8 Nano model on the Raspberry Pi 4 revealed significant limitations in processing power, resulting in only 29% accuracy due to the Raspberry Pi's limited CPU capabilities. To overcome these performance bottlenecks, the MobileNet V2 model was integrated with TensorFlow Lite (TFLite) to improve inference efficiency. This configuration leveraged the Raspberry Pi's computational power more effectively, achieving a processing speed of 9 frames per second (FPS). Further hardware optimization was achieved by integrating the Google Coral USB Accelerator, which offloads computationally intensive tasks to the Edge TPU, enabling the system to process video at over 30 frames per second (FPS). This optimization significantly improved real-time object detection with minimal latency, ensuring quick and accurate feedback for the user.

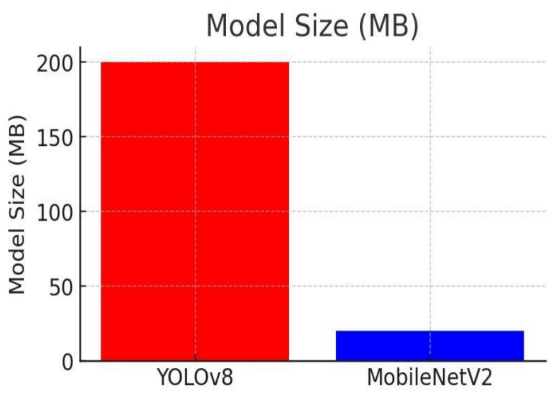


Fig 2: YOLOv8 vs MobileNetv2 Model size differences

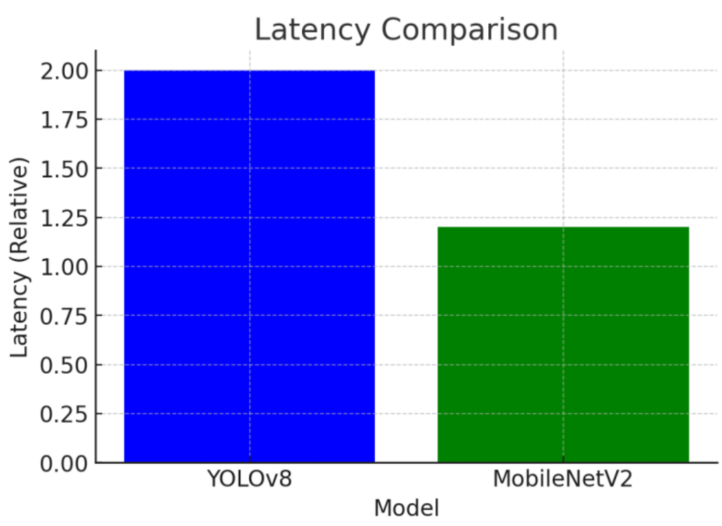


Fig 3: YOLOv8 vs MobileNetv2 Latency Comparison

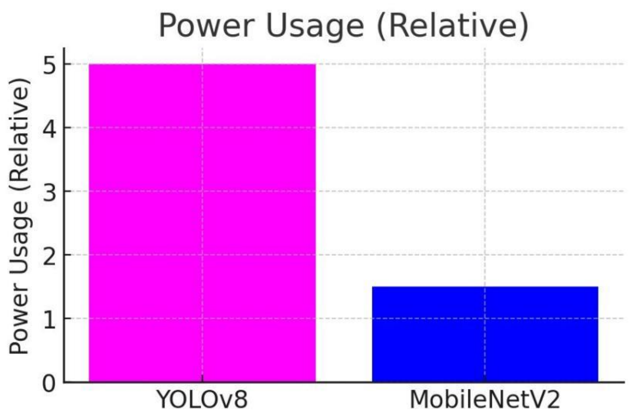


Fig 4: YOLOv8 vs MobileNetv2 Power Usage

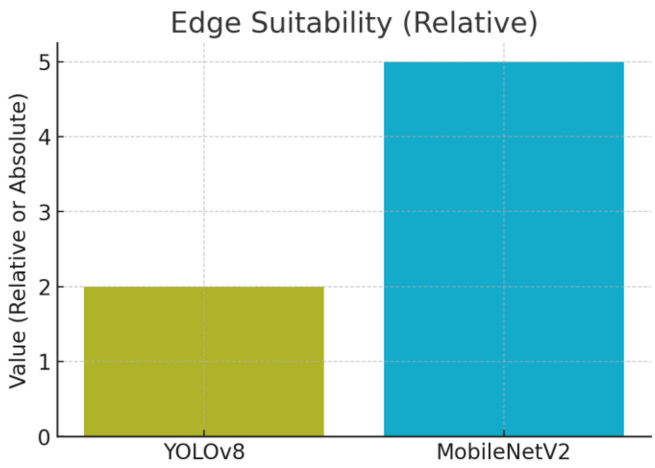


Fig 5 : YOLOv8 vs MobileNetv2Edge Suitability

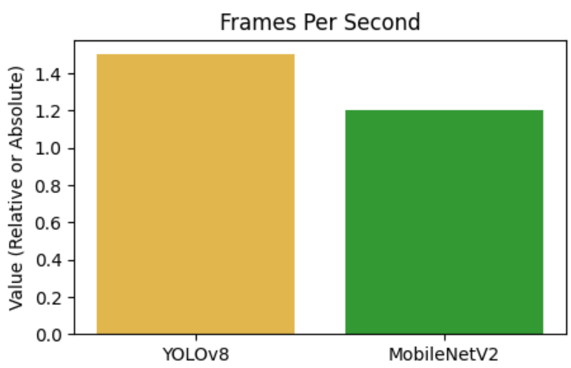


Fig 6: YOLOv8 vs MobileNetv2 Performance chart

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Software Integration and Optimization

The system's hardware integration involved configuring the Raspberry Pi 4 Model B with essential dependencies and deploying the MobileNet V2 object detection model, along with the necessary voice output capabilities through the TTS library.

Algorithm 1 Capture and Process Video Frame

Input: Webcam video stream
Output: Preprocessed video frame ready for inference

Initialisation:
1: Open the webcam feed using **cv2.VideoCapture(0)**
2: Start a continuous loop for video frame capture

LOOP Process
3: Capture a frame from the webcam
4: Check for the ESC key press to exit
5: Convert the captured frame from BGR to RGB using **cv2.cvtColor**
6: Resize the frame to match the model's expected input size
7: Convert the resized frame to bytes format
8: Set the input tensor for the model
9: Run inference using **interpreter.invoke()**

10: Return: Preprocessed video frame

Algorithm 2 Object Detection and

Input: Preprocessed video frame
Output: Annotated frame with detected objects

Algorithm 3 Text-to-Speech Announcement

Input: Detected objects from the model
Output: Spoken announcement of detected objects
LOOP Process
1: For each detected object, check if it was detected within the last 10 seconds using the last_spoken dictionary
2: If the object hasn't been announced recently, announce it using the text-to-speech engine (pyttsx3)
3: Update the last_spoken timestamp for that object to the current time
4: Return: Updated last_spoken dictionary

Initial testing revealed significant processing delays when running both object detection and audio feedback simultaneously, due to the limited processing power of the Raspberry Pi's CPU. To address this, multithreading techniques were implemented, allowing the Raspberry Pi to manage both the object detection process via the Coral USB Accelerator and the audio output through the wireless earphones concurrently, ensuring smoother real-time performance

Algorithm 4 Multi-threaded Execution

Input: None
Output: Concurrent execution of two functions

Initialisation:
1: Define **function_1** to print "thread_1"
2: Define **function_2** to print "thread_2"
3: Create two threads:
4: **t1** for f**unction_1**
5: **t2** for **function_2**

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

LOOP Process
4: Start both threads using t1.start() and t2.start()
5: Wait for both threads to complete using t1.join() and t2.join()
4: Return: Execution of both threads completed

Additionally, the autorun feature was configured to automatically launch the camera detection application as soon as the Raspberry Pi is powered on, using a configuration file. This ensures a seamless user experience, eliminating the need for manual intervention during startup. The hardware setup is optimized for continuous operation, with the Raspberry Pi, Coral Accelerator, and audio output peripherals working together without requiring additional manual setup after the initial boot.:

```
[Desktop Entry]
Type=Application
Name=Camera Detection
Exec=sh -c "export DISPLAY=:0 && cd /home/rasp/Desktop/coral && . /home/rasp/Test_coral/bin/activate && python cam.py"
Path=/home/rasp/Desktop/coral
Terminal=true
X-GNOME-Autostart-Delay=5
```

Model Training and Deployment

The MobileNet model was fine-tuned using a dataset representing real-world environments commonly encountered by visually impaired users. To ensure compatibility with the Coral USB Accelerator, the MobileNet model was converted to TensorFlow Lite (TFLite) format, required by the Edge TPU for efficient inference. The trained TFLite model was deployed on the Raspberry Pi 4, with the Coral TPU accelerating the object detection process. Initially, the system processed over 9 frames per second (FPS) using just the Raspberry Pi, but after integrating the Coral TPU, performance improved to 20 FPS, ensuring smooth real-time detection with better power efficiency.

The system's portability was enhanced with the autorun feature, ensuring the Raspberry Pi automatically starts the detection application upon power-on, eliminating the need for manual setup. A portable external power bank powers the Raspberry Pi, Coral Accelerator, and other peripherals, enabling continuous, on-the-go operation. Audio feedback is delivered through wired earphones, allowing users to hear real-time descriptions of detected objects, improving navigation and usability in diverse environments.

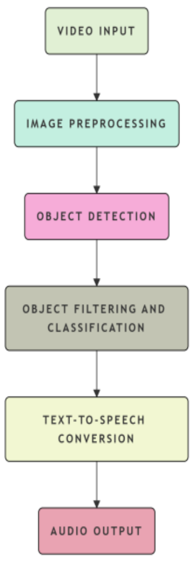


Fig 7: Flowchart of Methodology

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

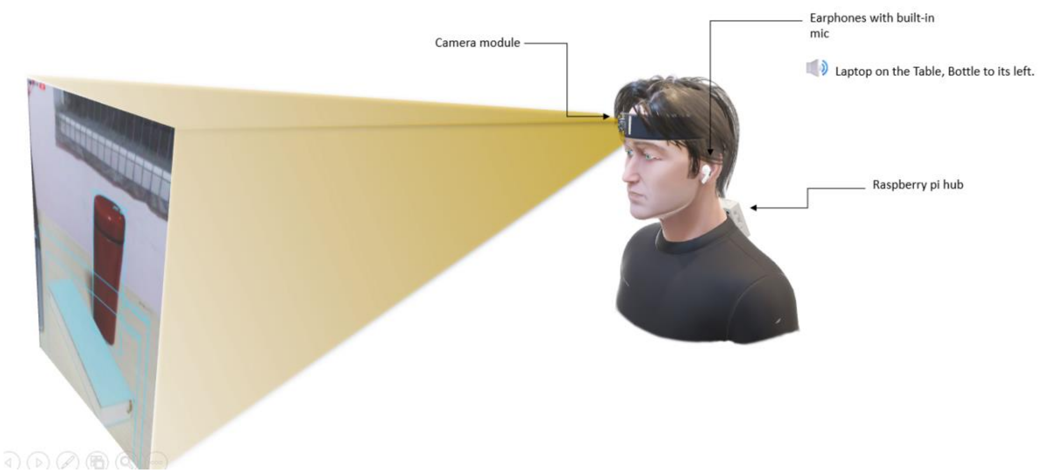


Fig 8 : Working Principle of the Wearable Assistive Vision System

V. APPLICATIONS

THE RETINA ASSISTIVE TECHNOLOGY SYSTEM HAS A WIDE RANGE OF POTENTIAL APPLICATIONS AIMED AT IMPROVING THE QUALITY OF LIFE FOR VISUALLY IMPAIRED INDIVIDUALS BY PROVIDING REAL-TIME OBJECT DETECTION AND DESCRIPTIVE AUDIO FEEDBACK. ITS PORTABILITY, EASE OF USE, AND REAL-TIME PROCESSING CAPABILITIES MAKE IT A VALUABLE TOOL IN VARIOUS ENVIRONMENTS AND SCENARIOS.:

- ***AI Assistant:*** It enables the user to easily integrate common language voice commands viaGoogle Assistant SDK as well as Google's Cloud Speech API.
- ***Personal Mobility Assistance:*** RETINA can significantly enhance the independence of visually impaired users by providing real-time object recognition and spatial awareness in both indoor and outdoor environments. The system helps users navigate safely through crowded public spaces, avoid obstacles, and recognize common objects such as doors, stairs, vehicles, and street signs. This feature can be particularly beneficial for daily commuting, shopping, and independent living.
- ***Workplace Accessibility:*** In professional settings, RETINA can assist visually impaired individuals by identifying essential objects and workplace hazards, allowing them to operate more independently in offices, manufacturing plants, and retail environments. The system's real-time audio descriptions can provide information about work tools, document locations, and even interactive elements such as computer screens or devices with tactile interfaces.
- ***Educational Support :*** RETINA has the potential to support visually impaired students by assisting with classroom navigation, recognizing objects such as books, desks, and digital devices, and providing real-time descriptive feedback during lessons. It can also be used in libraries to help identify bookshelves, categorize objects, and provide an audio description of educational materials, improving the overall learning experience.
- ***Home Assistance*** **:** The system can improve day-to-day life at home by recognizing household objects such as furniture, appliances, and personal items. Users can receive audio cues to locate items such as keys, remote controls, or kitchen utensils, reducing their dependence on others and fostering greater self-sufficiency. Additionally, RETINA can provide safety alerts for hazards such as open cabinets, sharp objects, or spills.
- ***Public Transportation Support :*** RETINA can aid visually impaired individuals in navigating public transportation systems by recognizing bus stops, train platforms, and station signs. It can provide guidance on locating ticket counters, identifying approaching vehicles, and detecting available seating areas, making travel more accessible and stress-free.
- ***Shopping and Retail Assistance*** : In shopping environments, RETINA can help users identify products, read price tags, and navigate store aisles. The system's object detection and audio feedback allow users to recognize different product categories and locate essential items without the need for external assistance. Integration with barcode

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

scanning and product databases can further enhance the shopping experience by providing detailed product descriptions and pricing information.

- ***Healthcare and Medical Assistance*** **:** RETINA can support visually impaired individuals in healthcare settings by assisting them in identifying medications, locating medical equipment, and navigating hospital corridors. It can also provide reminders and audio alerts for medication schedules, promoting better health management and reducing the risk of errors in self-care routines.
- ***Recreational Activities :*** Beyond practical applications, RETINA can enhance the recreational experiences of visually impaired individuals by helping them engage with hobbies such as gardening, cooking, and fitness activities. The system can identify objects used in these activities and provide guidance on their usage, fostering greater participation and enjoyment in leisure pursuits.
- ***Emergency Situations and Safety :*** RETINA can provide critical support in emergency scenarios by identifying exit routes, emergency equipment such as fire extinguishers, and potential hazards like smoke or obstructions. In case of a fire or other emergencies, the system can guide users towards the nearest safe exit, enhancing personal safety andresponse capabilities.

VI. EXPERIMENTAL RESULTS

A. Performance Testing and Evaluation:

A comprehensive testing strategy was employed to validate the system's functionality, stability, and performance. Unit and integration tests were conducted to ensure seamless interaction between hardware and software components. Object detection accuracy, inference speed, and power consumption were evaluated across different environments. The final system, integrated with the Google Coral, achieved real-time object detection at up to 20 frames per second (FPS), providing reliable audio feedback. This enhanced environmental awareness and situational understanding for users, ensuring the system met performance expectations.

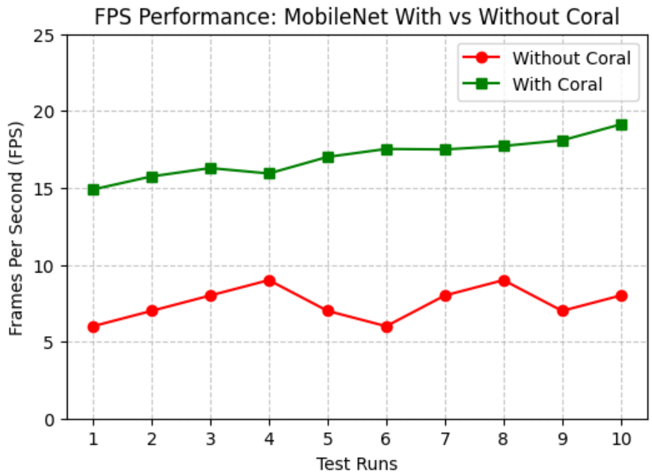


Fig 9: Mobilenet model performance with and without Coral TPU

B. User Trials and Feedback*:*

User trials were conducted with visually impaired participants to assess the usability and effectiveness of the RETINA system in real-world scenarios. Feedback from participants indicated improvements in navigation confidence and ease of use. Insights gathered from the trials were used to make refinements to the system, enhancing comfort, usability, and overall user experience.

The successful prototype, with real-time object detection at 20 FPS and auditory feedback, significantly enhances safety and independence for visually impaired individuals.

© 2025 IJMRSET | Volume 8, Issue 4, April 2025 | DOI:10.15680/IJMRSET.2025.0804392

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

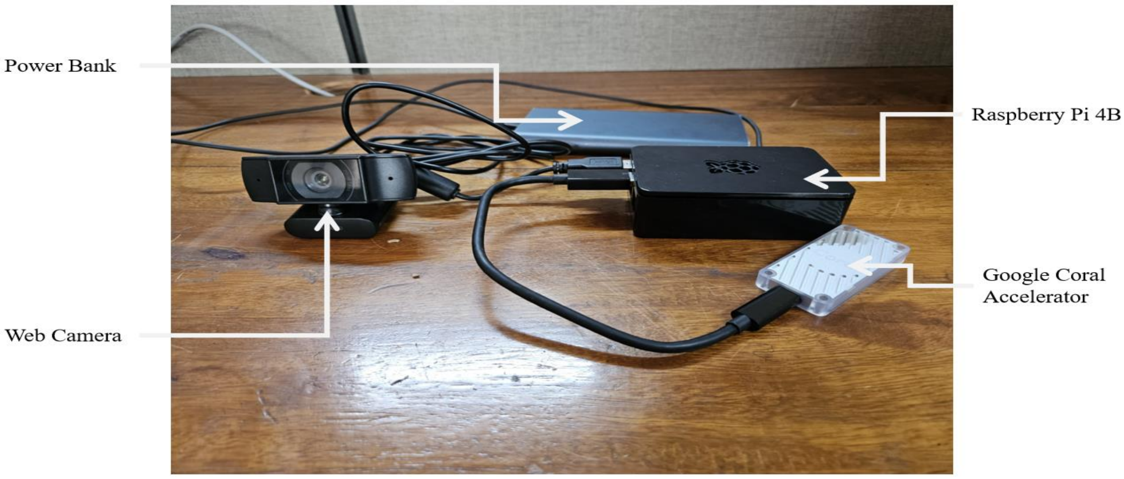


Fig 10: Prototype of the device

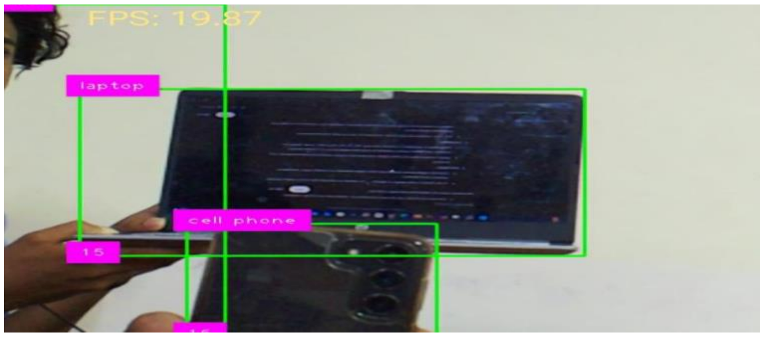


Fig 11 a: Object Detection Results

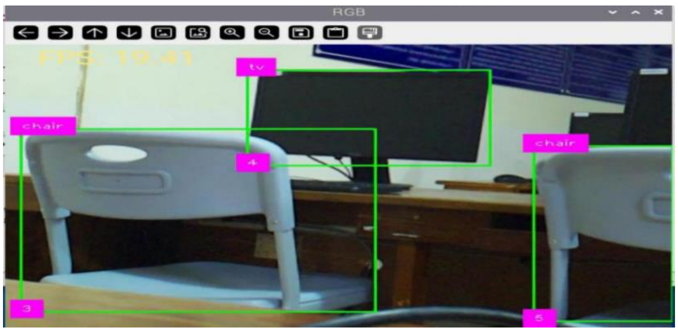


Fig 11 b: Object Detection Results

VII. CONCLUSION

The RETINA assistive technology project has made significant advancements in enhancing the mobility and independence of visually impaired individuals, with a strong focus on leveraging the power of the Coral TPU. Through a multi-phase approach, we have tackled key challenges and optimized object detection models for real-world applications, placing emphasis on hardware acceleration.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The transition from YOLOv8 Nano to MobileNet with TensorFlow Lite has highlighted the need to balance accuracy with processing efficiency. The Coral TPU has played a pivotal role in this, offering substantial improvements in processing speed and energy efficiency compared to traditional hardware. By utilizing the Coral TPU for hardware acceleration and combining it with techniques like multithreading, we have achieved real-time performance, enabling faster and more reliable object detection and audio feedback.

As the project progresses, RETINA will continue to refine its use of the Coral TPU, ensuring seamless and responsive assistance in practical environments. With its robust hardware capabilities, the Coral TPU is key to providing visually impaired users with greater autonomy, confidence, and inclusivity in their daily lives, setting the stage for the future of assistive technology.

REFERENCES

[1] Nick A. Bild MS ,“VoiceGPT”, Git hub, Jan 2023,

[2] Kaushal Trivedi, “Building a Multilingual Translation Tool with OpenAI ChatGPT API”,Medium, March 2023

[3] Hirak Ghael, “A Review Paper on Raspberry Pi and its Applications”, Research gate, Jan 2020,

[4] Sai Cheong Siu, “ChatGPT and GPT-4 for Professional Translators: Exploring the Potential of Large Language Models in Translation”, Research gate, May 2023.

[5] Konaite, M., Owolawi, P. A., et al. "Smart Hat for the Blind using Raspberry Pi and TensorFlow Lite." Developed a real-time detection system with SSD MobileNetV2 and audio feedback.

[6] Al-Buraihy, E., & Wang, D. "Object Detection System for Blind People Using Deep Learning." Improved cross-lingual image captions with high accuracy.

[7] Wang, L., & Wong, A. "Assistive Devices via MobileNetV2-SSD Optimization." Balanced accuracy, size, and speed for visually impaired users.

[8] Ghosh, A., Al Mahmud, S. A., et al. "Assistive Technology Using TensorFlow on Raspberry Pi." Integrated Raspberry Pi 4, Coral USB, and TensorFlow for real-time detection.

[9] Shehada, D., Turky, A., Khan, W., Khan, B., & Hussain, A. Lightweight facial emotion recognition system for visually impaired users.

[10] Yousif, A. J., & Al-Jammas, M. H. Comprehensive review of deep learning approaches for video captioning challenges.

[11] Terven, J., Córdova-Esparza, D.-M., & Romero-González, J.-A. (Year). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS.

[12] Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. (2011). Facial expressions in the wild (SFEW/AFEW). In Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV) (pp. 1-8). IEEE.

[13] Goodfellow, I. J., et al. (2013). Challenges in representation learning: A report on three machine learning contests. In Proceedings of the International Conference on Neural Information Processing (NeurIPS) (pp. 117–124). Springer.

[14] Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 111(15), E1454–E1462.

[15] Aafaq, N., Mian, A., Liu, W., Gilani, S. Z., & Shah, M. (2019). Video description: Survey of methods, datasets, and evaluation metrics. ACM Computing Surveys (CSUR), 52(6), 1-36.

[16] Baraldi, L., Grana, C., & Cucchiara, R. (2017). Hierarchical boundary-aware neural encoder for video captioning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1-9). IEEE.

[17] Xiao, H., & Shi, J. (2019). Video captioning with text-based dynamic attention and step-by-step learning. School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China.

[18] Li, H., Song, D., Liao, L., & Peng, C. (2019). RevNet: Bring reviewing into video captioning for a better description. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE.

[19] Assistive Technology for Visually Impaired using Tensor Flow Object Detection in RaspberryPi and Coral USB Accelerator by Ghosh, Amit; Al Mahmud, Shamsul Arefeen et. al.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[20] A Lightweight Facial Emotion Recognition SystemUsing Partial Transfer Learning for Visually Impaired People by Dina Shehada, Ayad Turky, Wasiq Khan, Bilal Khan, and Abir Hussain et al.

[21] Exploring deep learning approaches for video captioning: A comprehensive review by Adel Jalal Yousif , Mohammed H. Al-Jammas.

[22] H.N. Alkalouti, M.A. AL Masre, Encoder-decoder model for automatic video captioning using YOLO algorithm, in: 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2021

[23] M. Nabati, A. Behrad, Multi-sentence video captioning using content-oriented beam searching and multi-stage refining algorithm, Inf. Process. Manag. 57 (6) (2020), 102302.

[24] N. Aafaq, A. Mian, W. Liu, S.Z. Gilani, M. Shah, Video description: survey of methods, datasets, and evaluation metrics, ACM Comput. Surv. (CSUR) 52 (6) (2019) 1–37.

[25] Huanhou Xiao, Jinglun Shi, Video captioning with text-based dynamic attention and step-by-step learning, Pattern Recognit. Lett. (133) (2020) 305–312. Pages.

[26] Mancini, A., et al. (Year). An insight into assistive technology for the visually impaired and blind people: State-of-the-art and future trends. Journal on Multimodal User Interfaces, Springer.

[27] Bhowmick, A., et al. (2015). Embedded multisensor system for safe point-to-point navigation of impaired users. IEEE Transactions on Intelligent Transportation Systems, 16(11), 1-10

[28] Ahmad, N. S., et al. (2020). Assistive object recognition system for visually impaired. International Journal of Engineering Research & Technology (IJERT), 9(9), 1-8. Shaikh, S., et al. (2018). Multi-sensor obstacle detection system via model-based state-feedback control in smart cane design for the visually challenged. IEEE Access, 6, 54017-54027.

[29] Al-Malki, R. S., & Al-Aama, A. Y. (2023). Arabic captioning for images of clothing using deep learning. Sensors, 23(8), 3783.

[30] Mashetty, Harish, et al. "Deep Fake Detection with Hybrid Activation Function Enabled Adaptive Milvus Optimization-Based Deep Convolutional Neural Network." 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI). IEEE, 2025.

[30] Lasheen, M. T., & Barakat, N. H. (2022). Arabic image captioning: The effect of text pre-processing on the attention weights and the BLEU-N scores. International Journal of Advanced Computer Science and Applications (IJACSA), 13(7).

[31] Al-Muzaini, H. A., Al-Yahya, T. N., & Benhidour, H. (2018). Automatic Arabic image captioning using RNN-LSTM-based language model and CNN. International Journal of Advanced Computer Science and Applications (IJACSA), 9(6).

[32] Feng, Y., & Lapata, M. (2013). Automatic generation of captions for images using probabilistic models. IEEE Transactions on Pattern Analysis and Machine Intelligence.

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com