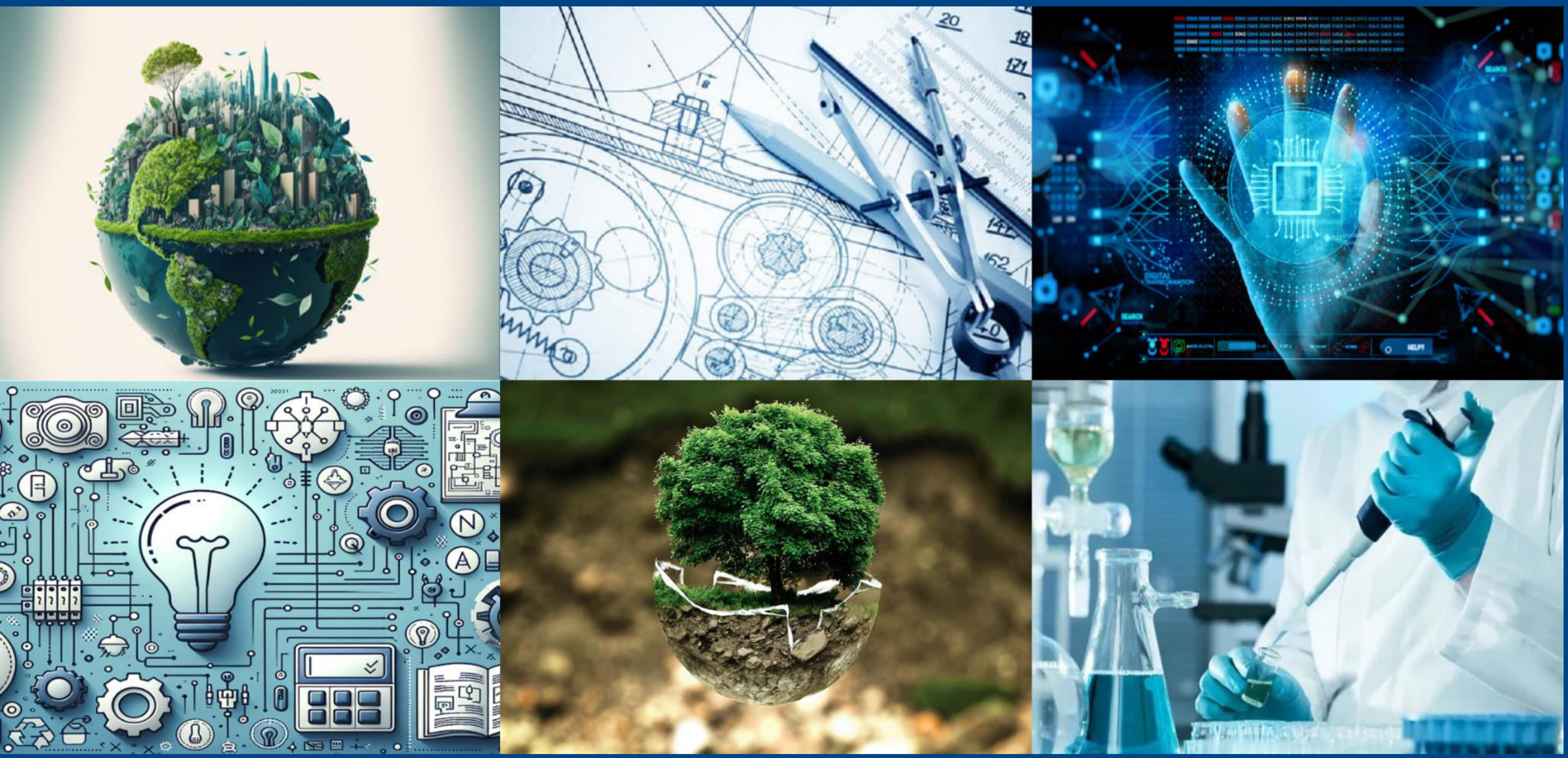
ISSN: 2582-7219

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 9, Issue 3, March 2026

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)

AI Multilingual Support System Developed for the Visually Impaired

Dr. Dipak D. Bage[1], Dr. Aniruddha S. Rumale[2], Krishnatanmay S. Bachhao[3], Piyush P. Bhadane[4], Sayali R. Dusane[5], Shruti K. Nathe[6], Vedant S. Thorat[7]

Assistant Professor, Dept. of Information Technology, Sandip Institute of Technology and Research Centre, Nashik, India[1]

Assistant Professor, Dept. of Information Technology, Sandip Institute of Technology and Research Centre, Nashik, India[2]

UG Student, Dept. of Information Technology, Sandip Institute of Technology and Research Centre, Nashik, India[34567]

ABSTRACT: Across the world, millions of people live without sight and face challenges that most others rarely notice in daily life [1]. Learning, traveling, moving around, and even simple conversations become constant struggles for visually impaired individuals [1]. Existing tools provide only partial assistance—some offer voice guidance, others directions—but most needs remain unmet [2], [3]. The work introduces a new form of intelligent assistant designed to move beyond these limited solutions. The proposed system is capable of sensing its surroundings, interpreting information, and responding in a natural and timely manner [4]. While supporting multiple languages to suit diverse users. Rather than functioning as a simple voice-based device, the system aims to unlock access to knowledge, learning opportunities, and personal independence. It is envisioned not merely as a product, but as an immersive experience that can significantly improve how visually impaired individuals learn, interact, and live. The broader impact points toward a future where accessibility, independence, and continuous learning coexist seamlessly.

KEYWORDS: Visual Impairment, Assistive Technology, Multilingual AI Assistant, Inclusive Design, Real-Time Assistance, Human-Centered AI, Barrier-Free Technology

I. INTRODUCTION

Sight is often taken for granted, but for those living with visual impairments, it becomes a daily challenge. According to the World Health Organization, more than 2.2 billion people worldwide suffer from vision loss or complete blindness [1]. Many of them face serious literacy issues due to limited or no access to proper assistive tools [2]. It creates barriers in education, mobility, adapting to different languages, and even basic independence [3]. Depending on others for reading or navigation not only slows down daily life but also reduces confidence and limits opportunities for learning [3].

In recent years, AI-driven assistive systems have emerged as a response to these challenges. Research by Prudhvi Naayini et al. [2] demonstrated that the integration of deep learning techniques with natural language processing can significantly enhance educational access and independence for visually impaired users [2]. However, their findings also highlighted unresolved challenges such as high system costs, complex usability, and insufficient multilingual support [4], [5], [6], which are critical for users from diverse linguistic backgrounds.

Building on, Baig et al. [3] developed a wearable device using Raspberry Pi with a small camera mounted on a hat, capable of detecting objects and providing contextual understanding. Yet, it struggled in low-light conditions and required internet connectivity, limiting its practical use. Similarly, Brilli et al. [4] introduced AIris, a multifunctional prototype that could read text, recognize faces, count money, and detect objects. While effective in tests, it only supported English, again exposing the weakness of current systems in multilingual adaptability.

Such advancements hold strong potential for visually impaired individuals who wish to pursue education, attend academic sessions, read printed materials, or independently manage daily tasks without constant assistance. Beyond accessibility, the broader objective is to restore independence, dignity, and confidence, enabling visually impaired users to actively participate in everyday life. This system serves as a companion to the visually impaired users

II. LITERATURE REVIEW

Establishing a solid research foundation requires a detailed review of existing work in the relevant domain. Analyzing prior studies helps identify effective methodologies, understand persistent challenges, and highlight unresolved issues. The process also aids in recognizing research gaps and determining how the proposed work can contribute meaningful improvements.

Relevance to current Research
The following section reviews key research contributions related to assistive technologies for visually impaired individuals.

Sr No.	Author	Title	Methodology	Algorithm	Conclusion	Remark
1	WHO Team (2023)	Increasing Eye Care Interventions to Address Vision Impairment	Technical brief reviewing global eye care strategies	Policy frameworks, statistical analysis	Eye care interventions can reduce global vision impairment burden	Focused on healthcare policy, not technical AI
2.	Prudhvi Naayini, Praveen Kumar Myakala, Chiranjeevi Bura, Anil Kumar Jonnalagadda, Srikanth Kamatala	AI-Powered Assistive Technologies for Visual Impairment	The authors reviewed AI-based tools like computer vision, NLP, and wearables for visually impaired assistance.	Deep learning models and real-time NLP frameworks were highlighted for object recognition and speech synthesis.	AI-driven assistive systems enhance independence, mobility, education, and social inclusion for visually impaired people.	The paper is comprehensive but notes challenges of affordability, usability, and multilingual adaptation.
3.	Mirza Samad Ahmed Baig, Syeda Anshrah Gillani, Shahid Munir Shah, Mahmoud Aljawarneh, Abdul Akbar Khan, Muhammad Hamzah Siddiqui	AI-based Wearable Vision Assistance System for the Visually Impaired	Developed a hat-mounted camera with Raspberry Pi 4 and LVLMs to provide real-time visual understanding.	Used MobileNet SSD for object detection, FaceNet for face recognition, and integrated GPT-based LVLM for contextual feedback.	The system achieved ~90% accuracy in good lighting and improved independence through personalized recognition and contextual descriptions.	Affordable and effective, but limited by hardware power, low-light conditions, and internet dependency.

4.	Dionysia Danai Brilli, Evangelos Georgaras, Stefania Tsilivaki, Nikos Melanitis, Konstantina Nikita	AIris: An AI-powered Wearable Assistive Device for the Visually Impaired	Built 3D-printed eyewear with Raspberry Pi and server-based ML for environmental analysis.	Applied ResNet-34 (face recognition), YOLO (object detection), Tesseract OCR (text), MAX Caption Generator	Prototype enabled scene description, text reading, money counting, and object/face recognition.	Strong multifunctional device but suffers from latency, hardware comfort issues, and only supports English.
5.	Issatay Tokmurziyev, Miguel Altamirano Cabrera, Muhammad Haris Khan, Yara Mahmoud, Luis Moreno, Dzmitry Tsetserukou	LLM-Glasses: GenAI-driven Glasses with Haptic Feedback for Navigation of Visually Impaired People	Designed lightweight glasses with ESP32 camera, YOLO-World detection, GPT-4o reasoning, and temple-mounted haptic actuators.	Used YOLO-World for object detection and GPT-4o for generating navigation instructions.	Achieved high accuracy in controlled tests (91.8% open spaces, 84.6% static obstacles, 81.5% dynamic obstacles) with intuitive tactile navigation.	Innovative use of haptics and GenAI, but early-stage; no multilingual communication, response time issues, and mostly tested indoors.
6.	Enci Wang, Xinhui Cai, Sixuan Chen	Research on a Computer Vision-Based Guide Glasses System for the Visually Impaired	Designed smart guide glasses integrating YOLOv5, ultrasonic sensors, GPS, and ESP32 for obstacle detection and navigation.	YOLOv5 (Nano version) with BlindVision-YOLO custom dataset for real-time target detection.	Achieved 74.1% mAP at 28.01 FPS with reliable obstacle alerts and voice feedback, improving safety in daily travel.	Strong in real-time detection and dataset design, but lacks multilingual communication and shows limitations in complex outdoor conditions.
7.	Fateme Zare, Mehdi Delrobaei, Paniz Sedighi	A Wearable RFID-Based Navigation System for the Visually Impaired	Low-cost, lightweight wearable system using passive RFID tags, tags communicate with transponders on the devices to guide the user via audio feedback.	RFID tags are nodes connected in a hexagonal pattern. Dijkstra's shortest-path algorithm, using tag scanning and orientation detection to guide the user.	The system proved simple, accurate, and reliable for indoor use, requiring no internet.	Navigation system with audio prompting that performs well in controlled indoor areas.In Future could include making it water-resistant, using Braille for input.

8.	Wiktor Mucha, Florin Cuconasu, Naome A. Etori, Valia Kalokyri, Giovanni Trappolini	TEXT2TASTE: A Versatile Egocentric Vision System for Intelligent Reading Assistance Using LLM	Smart glasses capture video, detect text, and use GPT-4 plus RAG for personalized responses.	DETIC for detection, EasyOCR, GPT-4 with RAG.	96.77% accuracy in retrieving multilingual menu text; user satisfaction 4.87/5.	Real-world testing; multilingual and personalized smart reading aid.
9.	Abhinav Pratap, Sushant Kumar, Suchinton Chakravarty	Adaptive Object Detection for Indoor Navigation Assistance: A Performance Evaluation of Real-Time Algorithms	Using a special indoor dataset of objects like doors and furniture to see how well each worked in real-world indoor settings.	YOLOv5 (looks at the whole image at once), SSD (uses preset boxes to find objects), Faster R-CNN (Adds region proposals first, then detects objects), Mask R-CNN (segmentation)	YOLOv5: 90.1% accuracy, 0.357 s; SSD: 79.8%, 0.72 s; others low speed and accuracy	YOLOv5 best for indoor real-time help; SSD OK; others too slow
10.	Alexander Mehta and Ritik Jalisatgi	Help the Blind See: Assistance for the Visually Impaired through Augmented Acoustic Simulation	Fast object detection and 3D mapping, then uses musical notes to help visually impaired users understand their surroundings through sound.	Custom segmentation reduces time complexity from $O(n^2)$ to $O(n)$	R-value 0.866 and 87.5% accuracy in noisy, outdoor nighttime tests	Musical notes make navigation quick even in noisy, low-light settings
11.	Goyal (2024)	Comparative Study of Microcontrollers	Compared Arduino, Raspberry Pi, ESP32	Benchmarking hardware	Raspberry Pi 4 offers best balance of cost and AI capability	Useful for hardware selection in assistive systems.
12.	Sharma et al. (2022)	Real-time Object Detection	Implemented YOLOv5n for embedded detection	YOLOv5n	Achieved efficient detection on Pi 4	Confirms Pi's suitability for low-cost AI
13.	Patel et al. (2020)	OCR with EasyOCR on Raspberry Pi 4	Applied EasyOCR for assistive reading	EasyOCR	Effective for printed text recognition	Struggles with handwritten or small fonts

14.	Kumar et al. (2021)	Face Recognition using DeepFace on Raspberry Pi	Implemented DeepFace for identification	DeepFace	Provided fast, contactless recognition	Accuracy drops in poor lighting or occlusion
15.	Setiawan & Yusuf (2022)	IoT Device Control with Offline Speech Recognition	Offline speech recognition for IoT control	Vosk engine	Reliable hands-free control in quiet settings	Accuracy reduced in noisy environments
16.	P. Singh, R. Kaur, M. Sharma (2022)	Text-to-Speech for Assistive Systems on Raspberry Pi 4	Implemented TTS module on Raspberry Pi for assistive reading	pyttsx3 / gTTS	Generated clear speech output for visually impaired users	Affordable and effective, but voice quality needs refinement
17.	L. Chen, A. Singh, B. Kumar, C. Zhao, D. Johnson (2020)	Scene Summarization using NLP on RaspberryP4	Combined camera input with NLP-based summarization	NLP parsing + rule-based models	Produced contextual scene descriptions for navigation	Useful for quick overviews, but summaries can oversimplify

The review highlights that while progress has been made, significant gaps remain. Current solutions often lack multilingual adaptability, affordability, and practical usability. The research aims to build on past work while addressing these gaps by designing a system that is practical, cost-effective, and reliable for real-world use.

III. SYSTEM ARCHITECTURE

The proposed system is an intelligent assistive platform which is developed to enhance environmental awareness for visually impaired people with the help of continuous real-time auditory feedback. Raspberry Pi 4 is the main and central component of the system architecture, which acts as an embedded computing unit responsible for integrating multiple hardware and software components. This is done by combining multimodal sensing with on-device intelligence, the platform transforms raw environmental signals into structured information that can be interpreted through audio output.

Environmental data are acquired and extracted using a vision sensor, a microphone, and ultrasonic distance sensors, supporting multi-input acceptance. Still images and video streams are captured and processed by the camera module which supports computer vision tasks like text recognition, object detection, and scene-level understanding. Voice commands captured through the microphone enable hands-free interaction and system control, improving accessibility and usability. Ultrasonic sensors continuously detect surrounding obstacles to provide safe movement for the user, enabling spatial awareness that helps in safe navigation. The fusion of visual, auditory, and distance-based sensory inputs allows the system to operate effectively in dynamic environmentsumber equations consecutively.

The Raspberry Pi 4 serves as the central processing unit, aggregating sensory data and executing task-specific intelligent modules. Each module is designed to operate independently while maintaining synchronization with the core control logic to ensure stable real-time operation. The platform leverages the quad-core ARM Cortex-A72 processor and 4 GB of RAM available on the Raspberry Pi 4 to perform local AI and machine learning inference directly on the device. This edge-based processing approach offers a cost-effective alternative to high-end embedded platforms such as the NVIDIA Jetson Nano, while maintaining sufficient computational performance for assistive tasks.

System outputs are primarily delivered through auditory interfaces, including speakers or earphones, enabling continuous and intuitive feedback without reliance on visual displays. The modular system design supports low-latency operation, improved portability, and reduced power consumption. Additionally, by performing computation at the edge, the system minimizes dependence on cloud services, thereby enhancing data privacy and operational reliability. Overall, the proposed architecture demonstrates an efficient, scalable, and affordable approach for real-time assistive systems targeting visually impaired users.

A. Input Module

The main role of the system is to gather information from the nearby environment of the user. This raw input data is then passed further for computation. Generally, the inputs are taken from the camera and various sensing devices [7]. By offering non-visual inputs, the sensing unit improves environmental awareness. The principle of operation for the Sensing Unit is to determine distance by calculating the time difference between sending and receiving sound waves. For a number of modules, including face recognition, object detection, and scene summarization, this serves as the raw dataset. This raw dataset is then cleaned, transformed and is further parsed for processing**.**

B. Central Processing Unit

The system's main processing brain is the Raspberry Pi 4 [11]. It gathers raw data from the sensing unit and camera, processes it using intelligent modules, and then sends the results to the user interface. Several modules within this unit function independently while maintaining connectivity to guarantee seamless real-time operation**.**

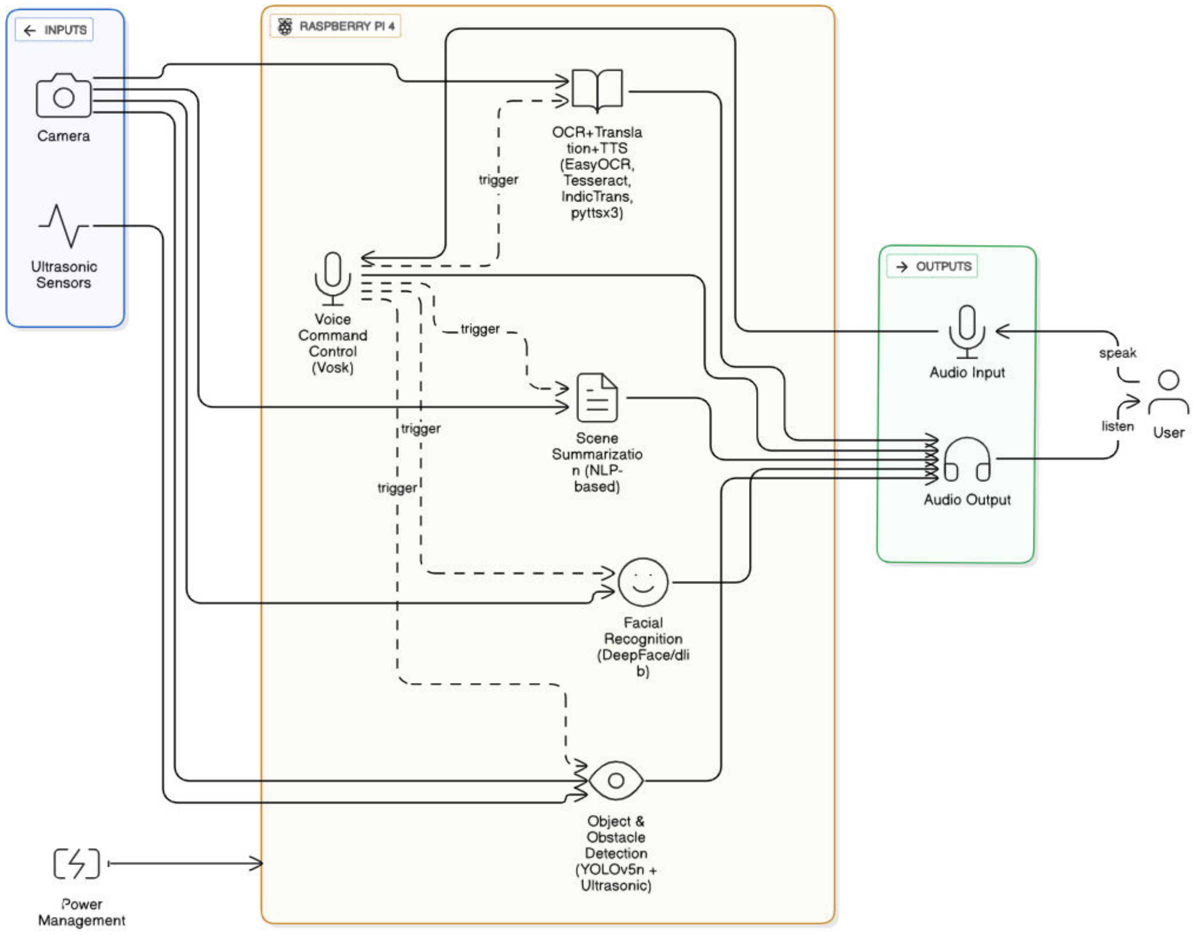


FIGURE 3.1: SYSTEM DESIGN

1. Object Detection Module

The objects in the visual field are recognized and categorized by this module[12]. It obtains the object's type and details. The basic idea is to divide the image into smaller portions, search for patterns in the edges, colors, and shapes, and then compare them to object models that have been stored. The function in the system is to identify significant objects, warn users of impediments, and give contextual information about the surroundings. Applications include real-time navigation support, recognizing potentially dangerous objects to issue safety alerts[9], and inventory tracking in intelligent environments. The benefits are that it works with several objects in real time and improves situational awareness. The restrictions are that camera quality and lighting affect performance and it may incorrectly classify partially visible objects.

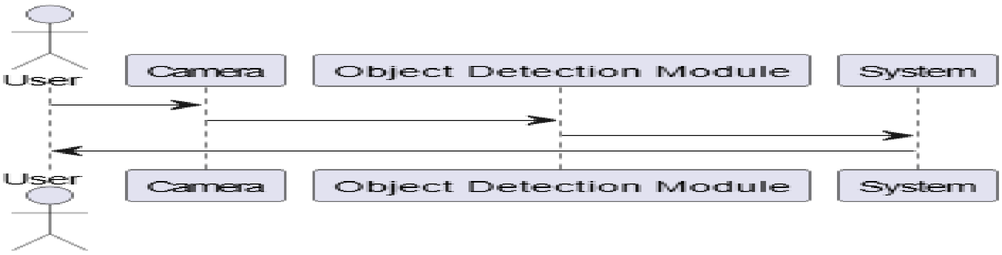


FIGURE 3.2: OBJECT DETECTION

2. Face Recognition Module

The detection and identification of human faces is the area of expertise for this module[14]. It searches facial features and compares them to profiles that have already been saved. The working principle is that the module looks for facial patterns like eyes, nose, mouth, and facial outline in the captured image. Important feature points are taken out and compared with stored facial data to confirm or identify the individual. The function in the system is to enable security features and tailored responses by offering user-specific recognition. The uses are authentication and security, systems for tracking attendance, and systems for Title individualized help such as greeting known users. The benefits are that it provides contactless and non-intrusive identification and is able to identify people fast. The restrictions are that poor lighting can cause performance to suffer and accuracy can be affected by angles, partial occlusions, and facial expressions.

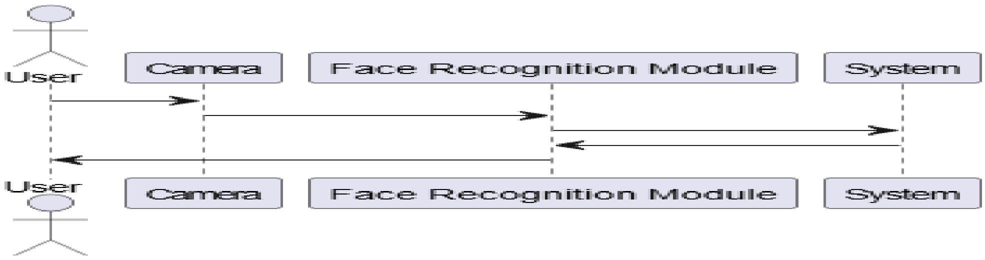


FIGURE 3.3: FACE RECOGNITION

3. Audio Command Module

This module permits hands-free interaction by translating user voice commands into system instructions[15]. The working principle is that a microphone records the user's speech, and then the audio is analyzed to identify words and turn them into commands. These commands are linked to particular system functions such as "Summarize scene," "Detect objects," and "Change language." The function in the system is to offer an instinctive, natural method of controlling the device without requiring manual input. The uses are voice-activated assistive technology, smart wearables or smart homes, and navigation assistance for visually impaired people. The benefits are that it provides a quick and easy way to interact and reduces reliance on touchscreens or buttons. The restrictions are that background noise can lower accuracy and unclear speech or different accents may impact recognition.

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)

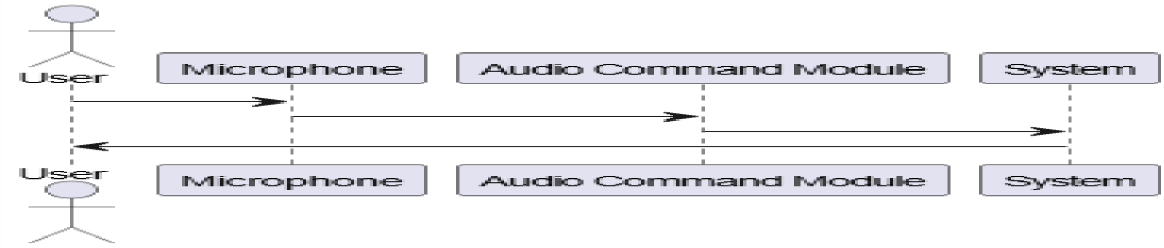


FIGURE 3.4: AUDIO COMMAND

4. Text Vocalization Module

This module transforms text outputs produced by the system, such as object names or scene summaries, into spoken words for the user[16]. The working principle is that textual data is processed into phonetic components, and then human-like speech is generated and played through a speaker. The function in the system is to provide real-time audible feedback so that the system can function even in the absence of a display interface. The uses are reading descriptions aloud to people who are blind or visually impaired, supplying warnings with audio alerts, and providing reading aids in multiple languages. The benefits are that it transforms data into a format that is simple to comprehend and enables interaction without using hands. The restrictions are that speech may sound robotic if not optimized, and clarity may require adjusting the tone and speed.

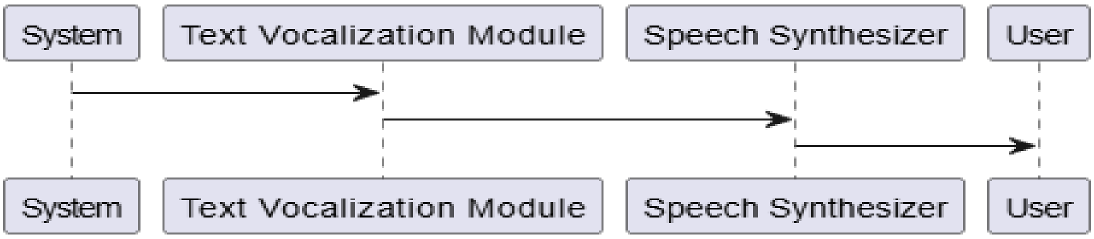


FIGURE 3.5: TEXT VOCALIZATION

5. Scene Summarization Module

Scene summarization provides a comprehensive description of the entire environment[17], [21], in contrast to object detection, which lists specific objects. It creates a useful synopsis by combining data from the camera, sensing unit, and detected objects. The working principle is that after determining the objects' locations and connections, the system creates a sentence or synopsis that sums up the scene as a whole. For example, "A person is sitting on a bench with a dog nearby." The function in the system is to give users a quick overview of their surroundings without being overly detailed. The uses are assisting visually impaired people in navigating and comprehending their surroundings, analyzing and summarizing surveillance footage, and producing automated scene reports. The benefits are that it summarizes several items at once, saving time, and provides context rather than just facts. The restrictions are that summaries could be imprecise, lacking, or oversimplify complex scenes.

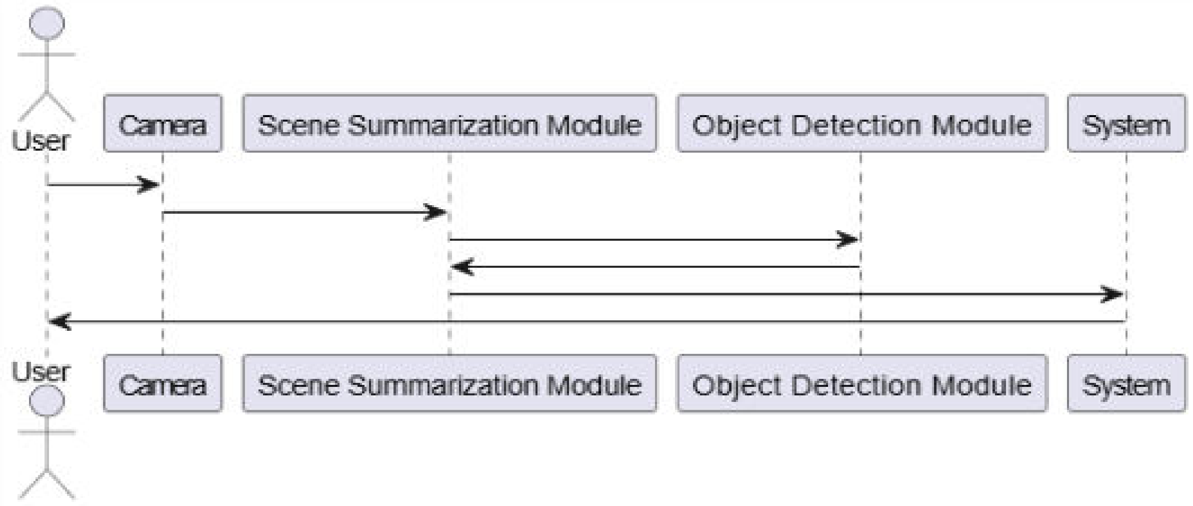


FIGURE 3.6: SCENE SUMMARIZATION

6. Multilingual Module

This module guarantees that the system's outputs can be provided in multiple languages based on user preference[8], [21]. The working principle is that speech commands or text outputs are translated into the chosen language and then output as either audio or text. The function in the system is to increase accessibility and enable efficient use of the device by users with various linguistic backgrounds. The uses are assisting users who speak multiple languages, providing real-time translation support, and enabling the use of assistive technology in various geographical areas. The benefits are that it improves the system's usability worldwide and fills in the gaps in communication. The restrictions are that contextual accuracy may not always be guaranteed in translations and performance depends on the accuracy of language models.

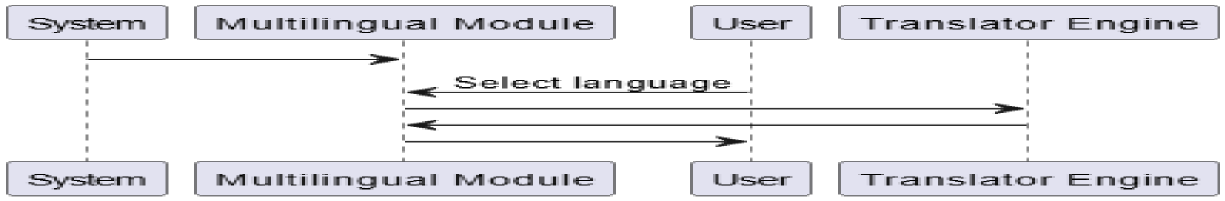


FIGURE 3.7: MULTILINGUAL MODULE

C. User Interface

The user interface is the part where interaction between the user and the system takes place. It ensures that data is displayed in voice in an understandable and practical way. The working principle is that it provides auditory cues to output information after receiving commands from the user via voice. Applications include presenting results through audio and acting as a medium for real-time communication.

The proposed AI-based Multilingual Assistive System was implemented and tested using Raspberry Pi 4 (4GB RAM) with YOLO-based object detection, EasyOCR for text reading, Vosk for speech recognition, pyttsx3/gTTS for voice output, and ultrasonic sensors for obstacle detection[11], [12]. The system was evaluated across multiple parameters — accuracy, latency, and usability — to determine real-time performance.

A. Object Detection Accuracy

The system was tested in three environmental conditions: indoor (good lighting), outdoor (daylight), and low-light. The YOLO-nano model achieved high accuracy in well-lit environments, while performance decreased slightly in darker conditions.

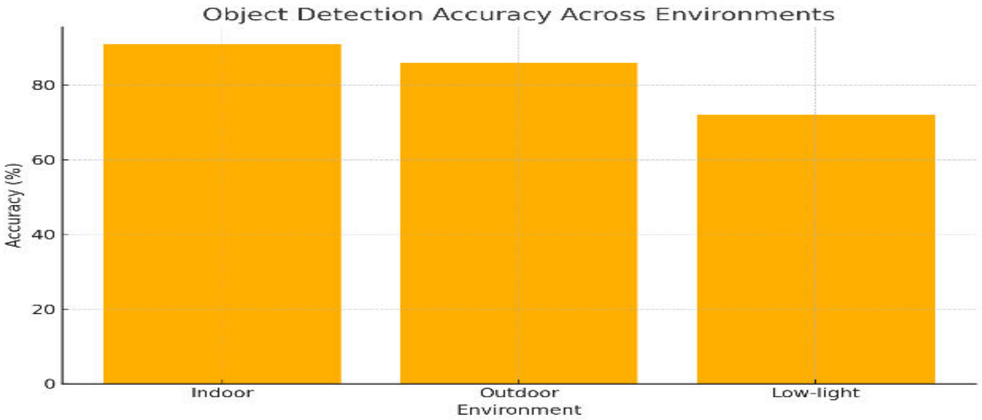


FIGURE 3.8: OBJECT DETECTION ACCURACY

The model's performance demonstrates reliability and adaptability in dynamic lighting conditions. The real-time frame rate (~8 FPS) on Raspberry Pi confirms its ability to process visual input efficiently for navigation and object awareness.

B. OCR Text Recognition

The text recognition module, implemented using EasyOCR, showed strong results for printed and medium-font text, while smaller fonts and handwritten inputs caused accuracy drops due to image blur and text complexity.

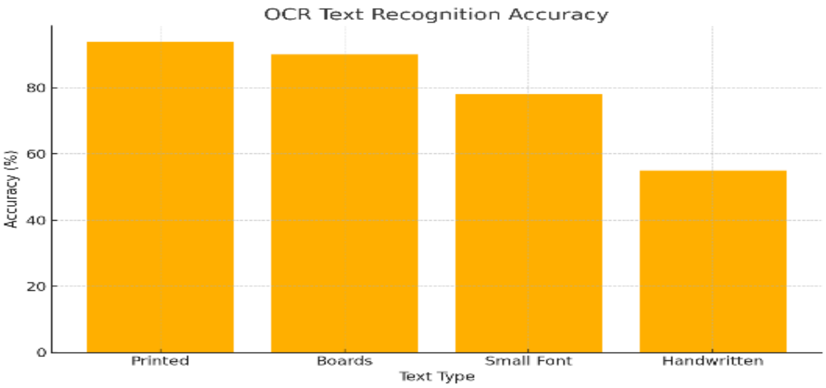


FIGURE 3.9: OCR TEXT RECOGNITION

OCR accuracy reached nearly 94% for printed materials, making it effective for reading newspapers, labels, and educational material. This demonstrates the system's ability to assist visually impaired users in academic and daily contexts. Further improvement could be achieved through adaptive font-size detection.

C. Speech Command Recognition

The Vosk offline speech recognition engine was tested in different environments ranging from quiet indoor rooms to noisy outdoor spaces. The recognition rate remained above 90% in quiet conditions[15] but dropped to around 74% in noisy environments.

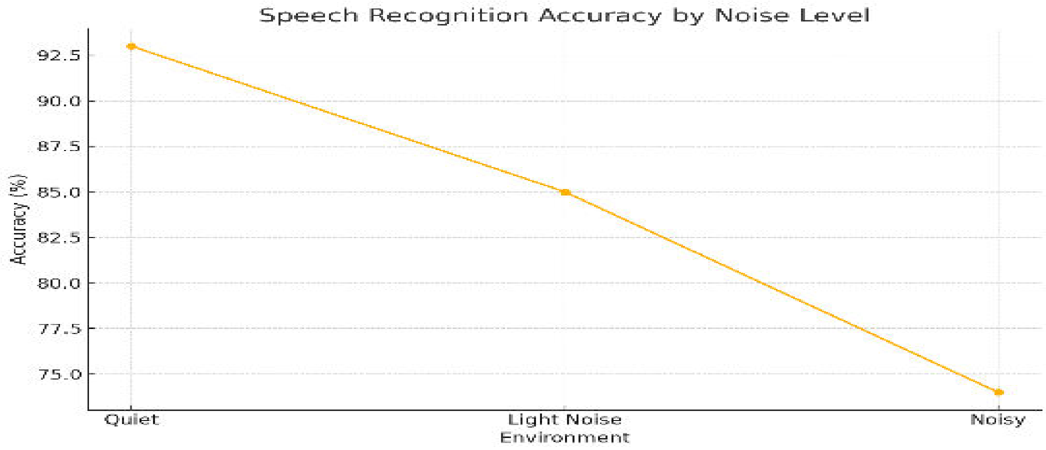


FIGURE 3.10: SPEECH COMMAND RECOGNITION

Speech-based control ensured hands-free operation and independence for users. Though accuracy decreased in heavy noise, the integration of noise suppression and directional microphones could enhance future versions. The system successfully executed all major commands like "detect object," "read text," and "change language."

D. User Evaluation and Usability Testing

A group of ten participants, including visually impaired users, evaluated the system's usability and interaction. Feedback indicated high satisfaction in areas like ease of use, response time, and clarity of voice output.

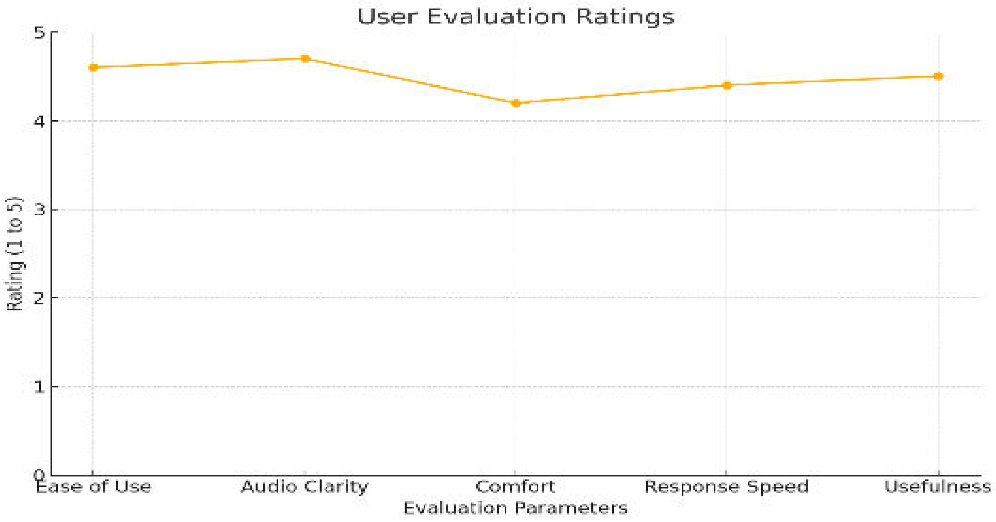


FIG 3.11: USER EVALUATION RATINGS

Participants rated the system's ease of use and multilingual feedback highly, averaging above 4.5/5. They particularly valued its real-time response and clear voice alerts, which enhanced their confidence in navigation and independent reading. The hardware comfort was rated slightly lower, indicating the need for ergonomic design improvements.

E. Overall System Performance Summary

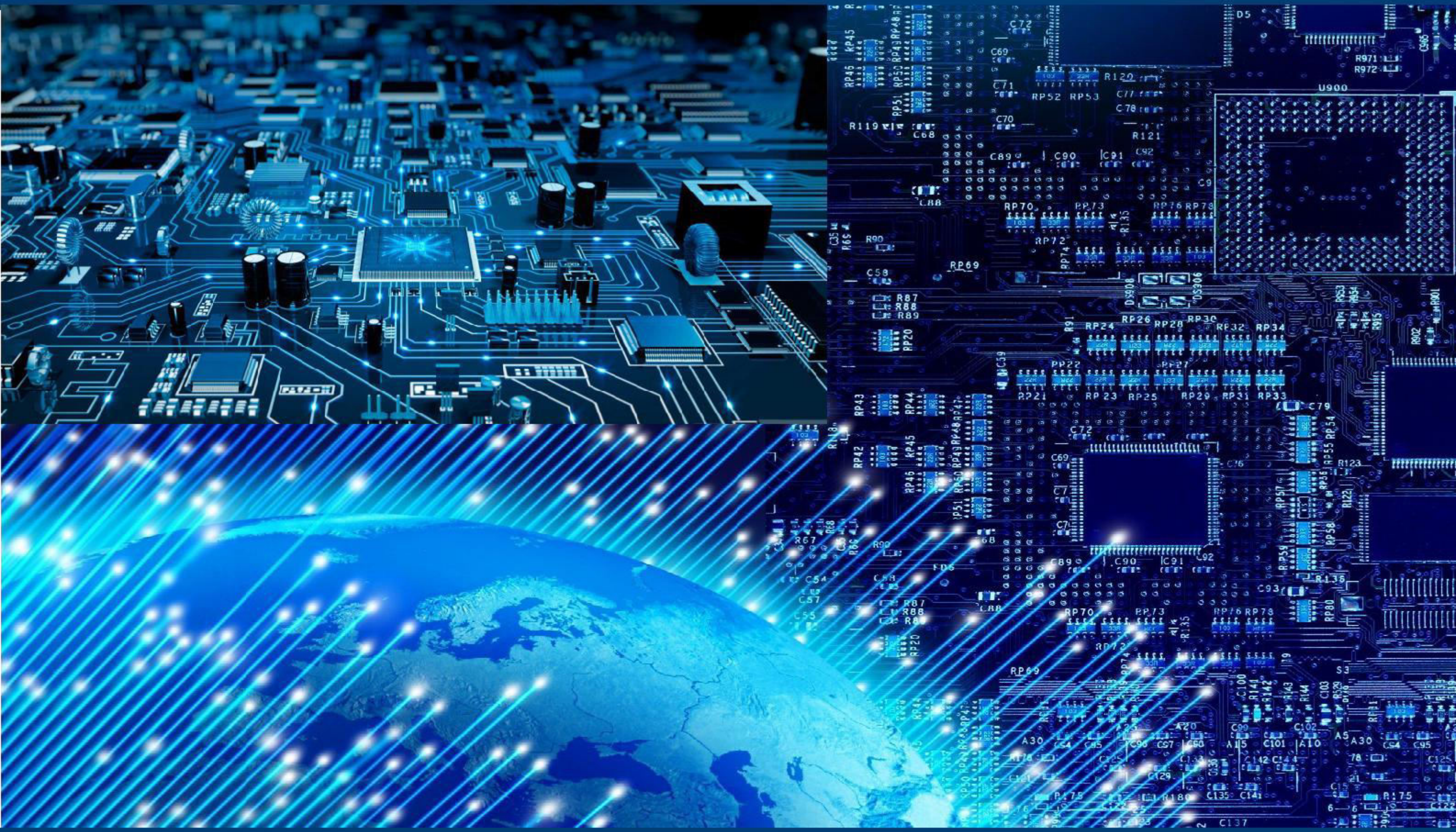
The integrated system achieved reliable real-time performance on low-cost embedded hardware, delivering accurate detection, text reading, and multilingual voice output. It maintained consistent latency under one second for most operations, ensuring seamless user experience. The findings confirm that the system effectively enhances accessibility and independence for visually impaired individuals through a unified, AI-driven platform.

IV. CONCLUSION

A comprehensive review of multiple papers and articles related to this field has been conducted, leading to the conclusion that there has been great successive progression but still lacks major gap in terms of adaptive learning in consideration of visually impaired people. While analyzing previous studies, it was inferred that the existing approaches had not achieved significant accuracy along with ensuring reliability and efficiency in a cost-effective manner. Current assistive technologies help visually impaired people in some areas, like object detection or navigation[2],[3],[4], but they are often expensive, not user-friendly and work in only one language [2]-[10]. The System will be able to fulfill the language barrier gap by providing multilingual support[8], [21] and brings together all the features in real time leveraging cost-effective hardware. These early findings confirm that the work is progressing in the right direction and should be continued to reach the final objectives. Observations suggest that the system could make a real difference in assistive technology for visually impaired users.

REFERENCES

[1] WHO Team UHC/Life Course EXDGO (UHA), World Health Organization, Increasing eye care interventions to address vision impairment, Technical Brief, 8 March 2023.
[2] P. Naayini, P. K. Myakala, C. Bura, A. K. Jonnalagadda, S. Kamatala, AI-Powered Assistive Technologies for Visual Impairment, June 2022.
[3] M. S. A. Baig, S. A. Gillani, S. M. Shah, M. Aljawarneh, A. A. Khan, M. H. Siddiqui, AI-based Wearable Vision Assistance System for the Visually Impaired, Sept. 2023.
[4] D. D. Brilli, E. Georgaras, S. Tsilivaki, N. Melanitis, K. Nikita, AIris: An AI-powered Wearable Assistive Device for the Visually Impaired, Nov. 2021.
[5] I. Tokmurziyev, M. A. Cabrera, M. H. Khan, Y. Mahmoud, L. Moreno, D. Tsetserukou, LLM-Glasses: GenAI-driven Glasses with Haptic Feedback for Navigation of Visually Impaired People, May 2024.
[6] E. Wang, X. Cai, S. Chen, Research on a Computer Vision Based Guide Glasses System for the Visually Impaired, Aug. 2023.
[7] F. Zare, M. Delrobaei, P. Sedighi, A Wearable RFID-Based Navigation System for the Visually Impaired, Dec. 2019.
[8] W. Mucha, F. Cuconasu, N. A. Etori, V. Kalokyri, G. Trappolini, TEXT2TASTE: A Versatile Egocentric Vision System for Intelligent Reading Assistance Using Large Language Models, Feb. 2024.
[9] A. Pratap, S. Kumar, S. Chakravarty, Adaptive Object Detection for Indoor Navigation Assistance: A Performance Evaluation of Real-Time Algorithms, July 2022.
[10] A. Mehta, R. Jalisatgi, Help the Blind See: Assistance for the Visually Impaired through Augmented Acoustic Simulation, Oct. 2018.
[11] Y. Goyal, Comparative Study of Microcontroller: Arduino Uno, Raspberry Pi 4, ESP32, International Journal for Research in Applied Science and Engineering Technology (IJRASET), Vol. 12, Issue VII, July 2024.
[12] S. Sharma, R. K. Gupta, A. Verma, P. Singh, Real-time Object Detection on Raspberry Pi 4 using YOLOv5n, 2022.
[13] M. Patel, K. Shah, R. Mehta, OCR with EasyOCR on Raspberry Pi 4 for Assistive Reading, 2020.
[14] A. Kumar, R. Yadav, S. Verma, Face Recognition using DeepFace on Raspberry Pi 4, 2021.
[15] P. Setiawan, R. Yusuf, IoT Device Control with Offline Automatic Speech Recognition on Edge Device, 12th International Conference on System Engineering and Technology (ICSET), 2022.
[16] P. Singh, R. Kaur, M. Sharma, Text-to-Speech for Assistive Systems on Raspberry Pi 4, 2022.
[17] L. Chen, A. Singh, B. Kumar, C. Zhao, D. Johnson, Scene Summarization using NLP on Raspberry Pi 4, 2020.
[18] D. Jurafsky, J. H. Martin, Speech and Language Processing, 2022.
[19] G. Jocher, YOLOv5: Real-Time Object Detection, 2020.
[20] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: Closing the Gap to Human-Level Performance, 2014.



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

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

www.ijmrset.com