iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81211

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

InSight-AI Assistant for Visually Impaired

Shambhavi Srivastavv[1], Satyendra Singh[2], Kavya Singh[3]
Dept. of Computer Science & Engineering- Artificial Intelligence and Machine Learning, Shri Ramswaroop Memorial College of Engineering & Management, Lucknow, India

Abstract: ***Visual impairment affects millions worldwide and severely restricts independent mobility, obstacle avoidance, and access to written information. Traditional assistive tools offer limited support, while many advanced solutions remain costly, bulky, or fragmented, highlighting the need for an affordable and integrated approach. InSight is a holistic AI-powered mobile system designed to address this gap by leveraging the sensing and processing capabilities of a smartphone. Based on a three-layer architecture—Input, Processing, and Output—the system combines key assistive technologies into a unified platform. It integrates real-time object and face detection using CNN models, dual-mode navigation through GPS and AI-enhanced SLAM, and OCR with NLP for text reading, summarization, and translation. Information is conveyed through natural text-to-speech and spatial audio cues for intuitive interaction. With user-centric features such as stride calibration and dynamic obstacle alerts, InSight aims to enhance independence, safety, and accessibility for visually impaired individuals.***

Keywords: Visual Impairment; Assistive Technology; Computer Vision; SLAM; OCR; NLP; Mobile-Based Navigation; Inclusive Technology.

I. INTRODUCTION

Visual impairment is a serious global issue that affects millions of people and has a major impact on their daily life, independence, and social participation. According to global studies, more than 2.2 billion people suffer from some form of vision loss, and many of them do not have access to proper care or assistive technologies. Individuals with visual impairments often struggle with everyday tasks such as navigating unfamiliar places, avoiding obstacles, recognizing people, reading text, and understanding their surroundings. These challenges not only affect safety but also reduce opportunities in education, employment, and social interaction, increasing dependence on others. Therefore, there is a strong need for advanced and accessible technological solutions that can improve independence and quality of life for visually impaired individuals [2].

Traditional assistive tools such as white canes, Braille systems, tactile paths, and guide dogs have been widely used for a long time. While these tools are useful, they provide only limited support. They mainly help in detecting nearby obstacles but cannot provide detailed information about the environment. For example, they cannot identify objects, read written content, or assist in navigating complex indoor spaces like malls or metro stations. Recent research has introduced AI-based solutions, but many of these systems focus on only one feature, such as object detection or navigation. Applications like Seeing AI, Lookout, and Be My Eyes provide useful support but are still limited in providing a complete solution. Additionally, wearable devices such as OrCam MyEye and Envision AI Glasses are often expensive and may not be comfortable for long-term use. This highlights the need for a more affordable, integrated, and user-friendly system [11]–[15]. Smartphones have become an effective platform for assistive technology due to their wide availability and powerful features. Modern smartphones include high-quality cameras, GPS, motion sensors, microphones, and built-in AI capabilities. These features allow smartphones to perform complex tasks like object detection, navigation, and voice interaction without requiring additional hardware. This makes smartphone-based solutions more affordable and scalable, especially in developing countries [4]. In this context, the InSight system is proposed as a complete AI-based mobile application designed for visually impaired users. The main goal of InSight is to combine important features such as navigation, object detection, text recognition, and contextual understanding into a single platform. This eliminates the need to use multiple applications and provides a smoother and more efficient user experience [2].

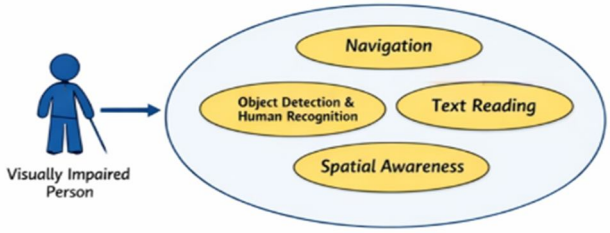


Figure 1. Bridging the gap between the user and environment

©IJRASET: All Rights are Reserved

The system is designed using a three-layer architecture: input, processing, and output. The input layer collects data from the smartphone's camera, GPS, and sensors such as accelerometers and gyroscopes. The processing layer uses advanced AI techniques to analyze this data. Convolutional Neural Networks (CNNs) are used for real-time object and human detection, enabling accurate identification of obstacles and surroundings [5], [9]. For navigation, the system uses a hybrid approach that combines GPS for outdoor navigation and Simultaneous Localization and Mapping (SLAM) for indoor environments where GPS is not reliable [6]. This ensures continuous navigation support in different environments. Additionally, Optical Character Recognition (OCR) is used to extract text from images such as signboards, labels, and documents, while Natural Language Processing (NLP) processes this text to provide meaningful and easy-to-understand output [7], [10].

The output layer focuses on delivering information in an easy and accessible way. Since visually impaired users rely more on audio, the system uses Text-to-Speech (TTS) to convert information into spoken form. Spatial audio is also used to provide directional cues, helping users understand where objects or obstacles are located. This improves awareness and helps users react quickly in real-time situations. Features such as obstacle alerts, personalized distance estimation, and face recognition further improve usability and independence [8].

The main objective of this work is to develop a system that integrates multiple technologies into a single, efficient assistive solution. By combining AI, computer vision, and language processing, InSight addresses the limitations of existing tools and provides a more complete and practical system. It also supports the goal of inclusive technology by making advanced solutions accessible and affordable. Overall, InSight has the potential to significantly improve mobility, independence, and quality of life for visually impaired individuals, contributing to a more inclusive and barrier-free society [1], [3].

II. LITERATURE SURVEY

Several research studies have explored different methods to improve assistive technologies for visually impaired individuals.

Joshi et al. (2024) developed an AI-based system called *AI-SenseVision*, which focuses on providing real-time assistance at a low cost.

The system uses camera input to detect surroundings and deliver immediate feedback to the user. This study highlights the importance of making assistive technology affordable while maintaining reliability.

Walle et al. (2022) presented a comprehensive survey of AI and vision-based systems designed for visually impaired users. The study explains how deep learning techniques have improved the accuracy of object detection. Deep learning enables systems to learn from large datasets, allowing better recognition of objects in complex environments.

Caraiman et al. (2017) introduced the concept of sensory substitution through their system, which converts visual data into sound signals. The system uses a camera to capture the surroundings and transform the information into audio cues. This helps users understand object positions and avoid obstacles using hearing instead of vision.

Budirions et al. (2022) focused on smartphone-based assistive solutions. Their research shows that smartphones are widely available and can serve as cost-effective tools for navigation and object detection. However, they emphasized the need for fast processing and user-friendly interfaces, especially since visually impaired users cannot rely on screen-based interaction.

Jafri et al. (2014) studied object recognition in indoor environments. The research highlights challenges such as constantly changing object positions, which make detection difficult. The study discusses different computer vision techniques to improve accuracy in such dynamic settings.

Croce et al. (2019) proposed a hybrid navigation system that works both indoors and outdoors. While GPS is effective outdoors, it fails indoors. Their system combines multiple technologies to ensure continuous navigation support, helping users move smoothly between different environments.

Ramadhan (2018) explored wearable assistive systems that use sensors to detect obstacles and provide alerts. These systems are convenient because they allow hands-free operation and continuous assistance.

Sethuraman et al. (2023) developed a wearable device called *MagicEye*, which integrates object detection and navigation features. The device supports visually impaired individuals in performing everyday tasks more independently.

Lin et al. (2019) demonstrated how deep learning improves wearable assistive systems. Their work shows that neural networks can enhance object recognition accuracy, especially in complex and crowded environments.

Chang et al. (2019) designed a system for recognizing medicines using deep learning. This application helps visually impaired individuals identify correct medicines, reduce the risk of errors, and improve safety.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The review of existing studies shows that significant progress has been made in assistive technologies for visually impaired individuals. Various solutions such as AI-based systems, wearable devices, and smartphone applications have improved mobility and independence.

However, several challenges still exist. Issues such as system accuracy, real-time performance, affordability, and user-friendliness continue to affect the effectiveness of these technologies. Many existing systems focus on specific tasks rather than providing a complete solution.

The proposed system, *InSight: AI Assistant Application for Visually Impaired*, aims to address these gaps by integrating multiple features such as object detection, text recognition (OCR), navigation, and voice interaction into a single platform. By combining these functionalities into one mobile application, the system seeks to provide a more efficient, accessible, and user-friendly solution.

III. PROPOSED METHODOLOGY

The proposed system, *InSight: AI Assistant Application for Visually Impaired*, is designed as a smart mobile application that uses artificial intelligence to assist visually impaired users in understanding their surroundings and moving safely. The system combines technologies like computer vision, natural language processing, and audio feedback to provide real-time support. The overall design follows a modular approach, which means the system is divided into different parts, and each part performs a specific task while working together smoothly.

The application mainly uses the smartphone camera to continuously capture images of the user's surroundings. These images are then analyzed using advanced AI models such as convolutional neural networks (CNNs) and vision transformers, which help in identifying objects like chairs, tables, obstacles, signboards, and other common items. The system not only detects objects but also tries to estimate their position and distance from the user.

This is done using techniques like depth estimation or combining data from available sensors. Based on this analysis, the system gives audio instructions to the user, such as "Object on your right" or "Move slightly left," helping them navigate safely without needing vision.

Apart from object detection, the system also includes a text-reading feature using Optical Character Recognition (OCR). This allows the application to extract text from images captured through the camera, such as medicine labels, signs, or printed documents. The extracted text is then converted into speech using a text-to-speech (TTS) system so that users can hear the information easily. This feature is especially useful in daily activities where reading is required. To make the application more user-friendly, a voice command system is also included. This feature uses natural language processing (NLP) to understand user commands. For example, the user can say things like "Read this for me," "What is ahead?" or "Take me to the door." The system first converts the speech into text and then processes it to understand the user's request and perform the correct action. This makes the application easy to use without needing complex controls or visual interaction.

For navigation, the system uses a combination of different methods to work both indoors and outdoors. GPS is used for outdoor navigation to provide directions and location-based guidance. However, since GPS does not work well inside buildings, the system switches to vision-based techniques and sensor-based tracking for indoor environments. It may also use technologies like SLAM (Simultaneous Localization and Mapping) to better understand the surroundings and guide the user step by step. At the same time, the system continuously checks for obstacles and provides real-time updates so that the user can move safely at all times.

The application is developed as an Android app using tools like Flutter or native Android development frameworks, making it easy to use on smartphones. The AI models can either run directly on the device for faster response and offline use or connect to cloud services such as Google ML Kit, Hugging Face, or Gemini when more advanced processing is needed. The system is designed to be fast, accurate, and simple to use, ensuring that users get immediate and reliable assistance in real-life situations. Since the system follows a modular structure, each feature like object detection, OCR, navigation, and voice interaction can be improved separately without affecting the whole application.

To check how well the system works, it is tested in different environments, both indoors and outdoors. Important factors like accuracy, response time, and user experience are evaluated during testing. The interface is designed in a way that minimizes the need for visual interaction, focusing mainly on audio instructions and simple controls. The aim of this methodology is to bring multiple assistive features together into one single application, so users do not need to depend on multiple tools. By combining AI, vision technology, and voice interaction, the InSight system provides a complete solution that helps visually impaired individuals to live more independently, safely, and confidently.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

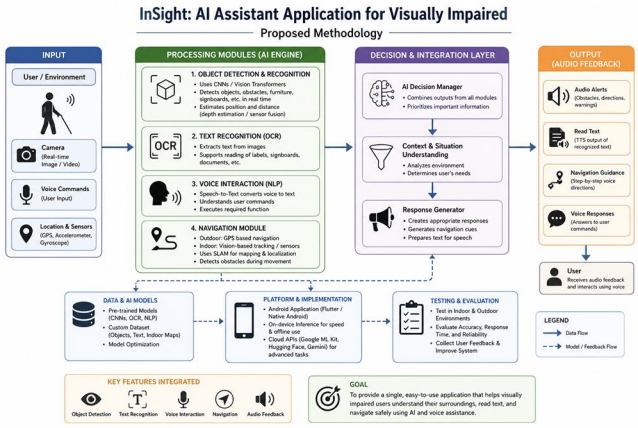


Figure 2.

IV. ANALYSIS

The *InSight: AI Assistant Application for Visually Impaired* is designed as a complete and well-organized solution to help visually impaired people improve their movement, understanding of surroundings, and decision-making in real time. What makes this project different is that it does not solve just one problem; instead, it combines many important features into a single application. Traditional tools like white canes only help in detecting nearby obstacles and do not give full information about the environment. Similarly, many modern apps focus on only one function, like navigation or object detection, which means users often need to switch between multiple apps. InSight solves this issue by bringing together navigation, object detection, text reading, and voice interaction into one system, making it more practical and easier to use.

From a technical point of view, the system is quite practical because it uses hardware that is already available in smartphones, such as cameras, microphones, GPS, and sensors like accelerometers and gyroscopes. This makes the solution affordable and easy to scale, as users do not need any special or expensive devices. The system uses advanced AI models like Convolutional Neural Networks (CNNs) and Vision Transformers to detect objects accurately in real time. These models help identify things like obstacles, furniture, signboards, and even people. The addition of face recognition can also help users recognize familiar individuals, increasing independence. However, running these complex models on mobile devices can sometimes cause issues like slower performance, delay in response, or higher battery usage, especially on low-end phones. To handle this, optimization methods such as reducing model size, using edge processing, or shifting some tasks to the cloud can be used.

One of the most important strengths of the system is its hybrid navigation feature. It uses GPS for outdoor navigation and vision-based or sensor-based techniques for indoor navigation. This is useful because GPS does not work properly inside buildings. By using methods like Simultaneous Localization and Mapping (SLAM), the system can understand indoor environments and guide users effectively. This combination ensures that users get continuous navigation support in all situations. However, SLAM requires a lot of processing power, so it needs to be carefully optimized to work smoothly on mobile devices. The system also combines navigation with obstacle detection, which improves safety by giving real-time alerts and directions.

Another key feature is the use of Optical Character Recognition (OCR) along with Natural Language Processing (NLP). OCR helps the system read text from real-world objects like medicine labels, signboards, or documents. This text is then converted into speech so that users can listen to it easily. NLP makes the system smarter by allowing it to understand user voice commands. For example, users can say “Read this” or “What is in front of me?” and the system will respond accordingly. This makes the application more interactive and reduces the need for manual controls. The project also focuses a lot on improving interaction between the user and the system. Since visually impaired users depend more on sound, the system uses Text-to-Speech (TTS) and spatial audio feedback to provide clear and useful information. For example, it can tell the direction of an object, helping users imagine their surroundings better. This reduces mental effort and allows users to operate the system hands-free, which is very important for convenience and safety. Another advantage of InSight is its modular design. This means that different parts of the system, like object detection, OCR, navigation, and voice interaction, can be developed and improved separately. This makes the system flexible and easy to upgrade in the future. For example, the application can later be converted into wearable devices like smart glasses, which can provide even better assistance using features like dual cameras and depth sensing. The system can also be enhanced by integrating advanced AI platforms like Gemini, Hugging Face, or Groq.

Even though the system has strong potential, its success will depend on proper real-world testing and user feedback. Different conditions such as low light, crowded areas, noise, and device limitations can affect how well the system works. Therefore, continuous testing and improvement are necessary to ensure good performance. Factors like speed, accuracy, and reliability must be carefully managed so that the system works smoothly in everyday situations.

In conclusion, InSight is an innovative and effective solution that aims to improve accessibility for visually impaired individuals. By combining multiple AI technologies into one platform, it provides a more complete and efficient system compared to existing solutions. With proper optimization and future improvements, it has the potential to greatly improve independence, safety, and overall quality of life for its users.

A. Major Finding

After reviewing different research papers and existing assistive technologies, several important gaps and opportunities were identified in the field of visual assistance for visually impaired people. These observations played a major role in shaping the design of the InSight system.

1) Limited Understanding in Traditional Assistive Tools

Basic tools like white canes and guide dogs mainly help users avoid obstacles while walking. However, they do not provide detailed information about the surroundings. For example, they cannot recognize objects, read text, detect moving objects, or explain the environment such as directions or room structure. Because of this, users still face difficulties in navigating independently, especially in complex or crowded areas [1].

To solve this problem, InSight introduces an AI-based perception system that not only detects objects but also explains them through audio, giving users a better understanding of their environment.

2) Lack of Integration in Existing Digital Solutions

Many current assistive apps focus on only one feature, such as object detection, navigation, or text reading. Some apps are useful, but they do not provide all features in one place. Also, certain platforms depend on human assistance, which may not always be available in real time.

Because of this, users often need to switch between multiple applications, which can be confusing and inefficient [2].

InSight improves this by combining multiple features—object detection, OCR, navigation, and voice interaction—into a single application, making it more convenient and independent to use.

3) Smartphones as an Effective Platform

With recent improvements in smartphone technology, devices now come with powerful processors, high-quality cameras, and built-in AI support. This allows complex AI models to run directly on mobile devices without needing extra hardware [3].

Compared to expensive wearable devices, smartphone-based solutions are more affordable and widely available. This makes them suitable for large-scale use, especially in countries where cost is an important factor.

InSight uses this advantage by being designed as a mobile application, making it easy to access and scalable.

4) Use of AI and SLAM for Better Navigation and Detection

Advanced AI models like YOLO and MobileNet SSD have made it possible to detect objects quickly and accurately in real time [4], [5]. These models help identify obstacles, furniture, and even people in different environments.

For navigation, SLAM (Simultaneous Localization and Mapping) is very useful, especially indoors where GPS does not work properly. It helps the system understand the environment and track the user's position at the same time [6].

By combining these technologies, InSight provides continuous guidance both indoors and outdoors, making navigation more reliable.

5) Importance of Simple and Natural Interaction

Studies show that visually impaired users prefer systems that are easy to use and do not require too much effort to understand. Features like voice input, speech output, and directional audio make interaction more natural and comfortable [7].

Some wearable systems already use audio cues to guide users effectively [8].

InSight follows this idea by using voice commands and spatial audio, allowing users to interact with the system without using their hands and helping them better understand directions.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

B. Proposed System Objectives

Based on these findings, the following goals are defined for the InSight system:

1) Build a Unified AI-Based Assistive Application

The main goal is to develop a single mobile application that combines different assistive features like navigation, object detection, and information access. This removes the need to use multiple apps and provides a smoother user experience [2].

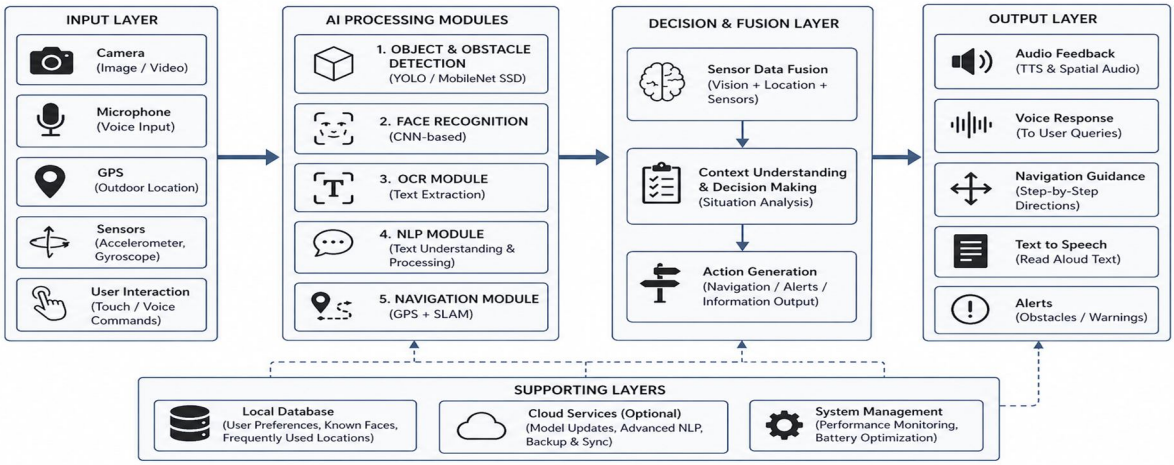


Figure 3: InSight System Architecture

2) Enable Real-Time Detection of Objects, Faces, and Obstacles

The system will use lightweight AI models like YOLO and MobileNet SSD to detect objects and obstacles in real time. It will also identify known faces when possible.

This helps in:

- Avoiding obstacles (safety)
- Recognizing surroundings (awareness)
- Identifying people (social interaction)

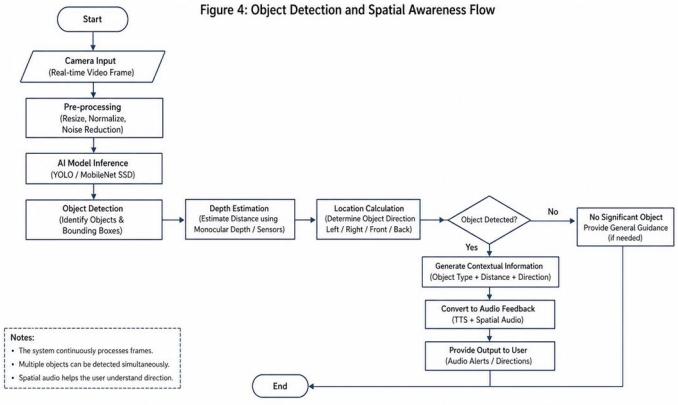


Figure 4: Object Detection Flow

3) Provide Navigation in Both Indoor and Outdoor Environments

The system will use GPS for outdoor navigation and SLAM-based methods for indoor navigation. This ensures users can move easily in any environment.

This approach helps in:

- Handling GPS failure indoors
- Understanding new environments
- Giving step-by-step directions

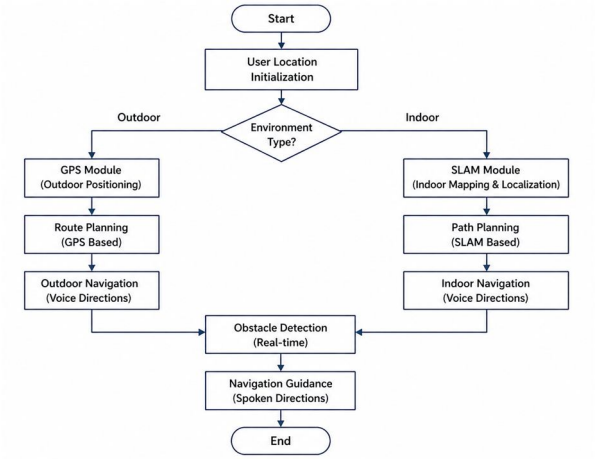


Figure 5: Hybrid Navigation System

4) Make Text Information Accessible Using OCR and NLP

The system will extract text from images using OCR and then process it using NLP. The output will be converted into speech so users can understand it easily.

This will help in:

- Reading signboards and documents
- Understanding medicine labels
- Accessing important information

5) Enable Hands-Free and Easy Interaction

Users will interact with the system using voice commands, and responses will be given through audio. Spatial audio will also indicate the direction and distance of objects.

This results in:

- Less mental effort
- Faster response time
- Better safety during navigation

V. CONCLUSION

The InSight system offers a complete AI-based solution designed to assist visually impaired individuals in their daily lives. It combines important features like real-time environment understanding, indoor and outdoor navigation, text reading, and audio-based interaction into one single mobile application. By using modern technologies such as computer vision, SLAM for mapping, natural language processing, and smartphone-based AI, the system is able to solve many problems found in traditional assistive tools and existing apps that usually focus on only one function. This integrated approach helps users move more safely, understand their surroundings better, and perform tasks independently without relying heavily on others. The system is designed to be user-friendly, cost-effective, and easy to access, making it suitable for a wide range of users. Overall, InSight shows strong potential as a scalable and practical solution that can improve accessibility and support the development of inclusive technology for visually impaired individuals

REFERENCES

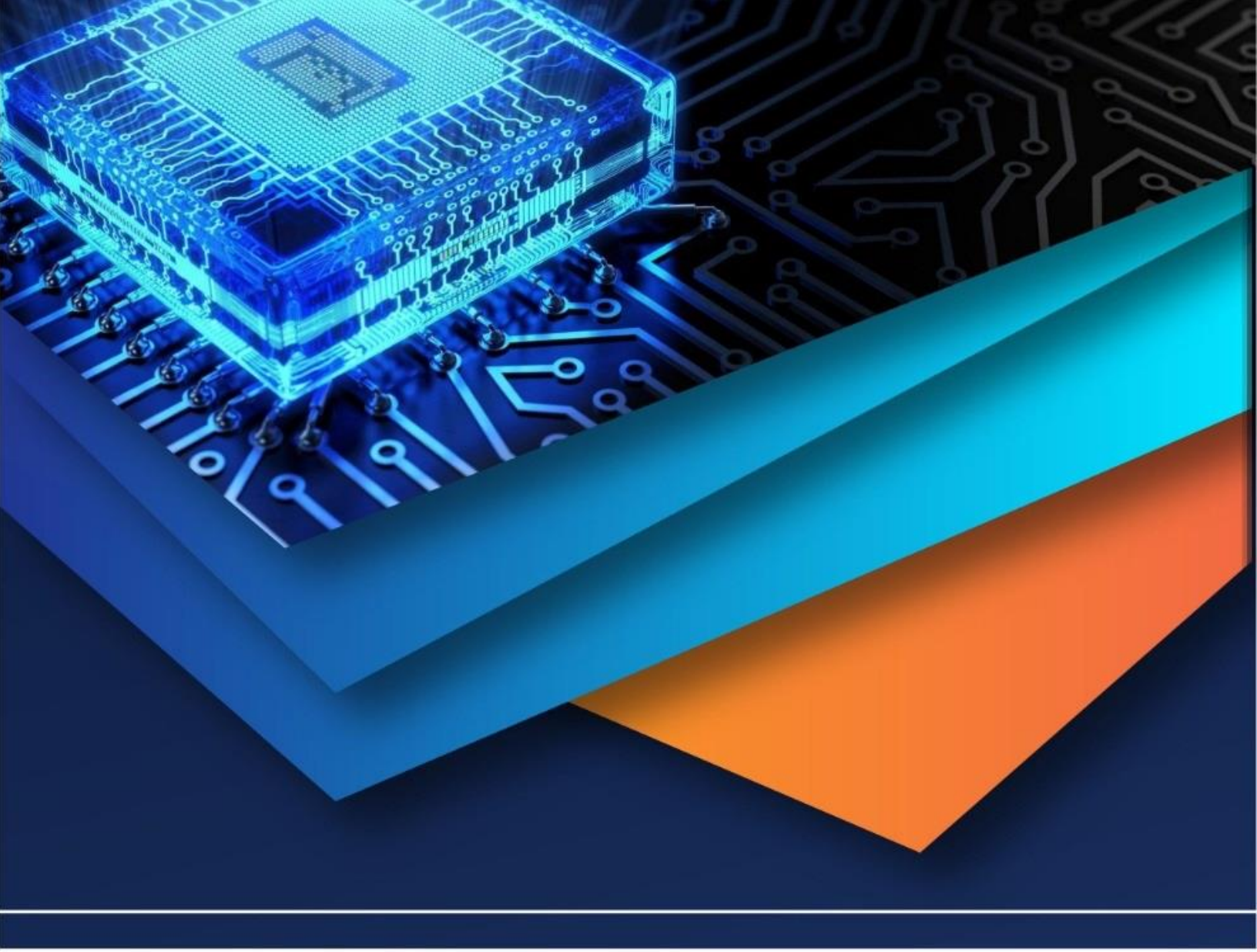
[1] R. Joshi, N. Singh, and A. Sharma, “AI-SenseVision: A Low-Cost Artificial-Intelligence-Based Robust and Real-Time Assistance for Visually Impaired People,” IEEE Transactions on Human-Machine Systems, 2024.
[2] H. Walle, C. De Runz, B. Serres, and G. Venturini, “A Survey on Recent Advances in AI and Vision-Based Methods for Helping and Guiding Visually Impaired People,” Applied Sciences, vol. 12, no. 5, pp. 1–25, 2022.
[3] S. Caraiman, A. Morar, and M. Owczarek, “Computer Vision for the Visually Impaired: The Sound of Vision System,” Proceedings of Computer Vision Conference, 2017.
[4] A. Budirions, D. Pilknyas, and P. Daniušis, “Smartphone-Based Computer Vision Travelling Aids for Blind and Visually Impaired Individuals: A Systematic Review,” Assistive Technology Journal, 2022.

[5] R. Jafri, S. A. Ali, H. R. Arabnia, and S. Fatima, "Computer Vision-Based Object Recognition for the Visually Impaired in Indoor Environments: A Survey," The Visual Computer, Springer, vol. 30, no. 11, pp. 1191–1209, 2014.
[6] D. Croce, L. Giare, and F. Pascucci, "An Indoor and Outdoor Navigation System for Visually Impaired People," IEEE Conference on Intelligent Systems, 2019.
[7] A. J. Ramadhan, "Wearable Smart System for Visually Impaired People," Sensors Journal, vol. 18, no. 3, pp. 1–12, 2018.
[8] S. C. Sethuraman, G. R. Tadkapally, and S. P. Mohanty, "MagicEye: An Intelligent Wearable Towards Independent Living of Visually Impaired," arXiv Preprint arXiv:2301.xxxxx, 2023.
[9] Y. Lin, K. Wang, and W. Yi, "Deep Learning Based Wearable Assistive System for Visually Impaired People," Proceedings of Computer Vision Systems, 2019.
[10] W. J. Chang, Y. X. Yu, and J. H. Chen, "A Deep Learning-Based Wearable Medicines Recognition System for Visually Impaired People," IEEE Access, vol. 7, pp. 17054–17064, 2019.

MORE REFERENCES

[11] Microsoft Corporation, "Seeing AI: Talking Camera App for the Visually Impaired," Available: https://www.microsoft.com/seeing-ai
[12] Be My Eyes Organization, "Be My Eyes – Connecting Blind and Sighted People," Available: https://www.bemyeyes.com
[13] Google LLC, "Lookout – Assisted Vision App for the Blind and Low Vision," Available: https://play.google.com/store/apps/details?id=com.google.android.apps.accessibility.reveal
[14] Envision Technologies, "Envision AI Glasses and Mobile Application," Available: https://www.letsenvision.com
[15] OrCam Technologies, "OrCam MyEye Assistive Device for the Visually Impaired," Available: https://www.orcam.com
[16] Aira Tech Corp., "Aira: Visual Interpreting Service for the Blind," Available: https://aira.io
[17] TapTapSee Inc., "TapTapSee – Object Recognition App for Visually Impaired Users," Available: https://taptapseeapp.com

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)