iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VIII Month of publication: August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73684

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

AI-Driven Vision Assistance for Visually Impaired

Hoysala Y Devanga[1], R Jagadishwar Reddy[2], Sai Sanjaya M[3], Shiva Kumar M[4]
Department of Computer Science and Engineering-ICB K S Institute of Technology

Abstract: ***ThisprojectintroducesanAI-basedreal-timeobjectdetectionandvoiceassistancesystemdesigned to support visually impaired individuals in navigating their surroundings independently and safely. The solution combines a mobile application with a portable camera module, enabling continuous environmental awareness without relying solely on the smartphone's built-in camera. By utilizing advanceddeeplearningalgorithmssuchasYOLO orSSD,thesystemdetectsandidentifiesobjects in real time.Once an object is recognized, the application uses a Text-to-Speech (TTS) engine to provide immediate voice feedback, informing the user of the object's identity. The integration of a portable camera enhances flexibility and usability, allowing for hands-free operation and more accurate object capture. The system is optimized for performance on mobile devices, ensuring low latency, high accuracy, and energy efficiency. This project aims to deliver a cost-effective, accessible,andpracticalassistivetechnologysolutionthatimprovesautonomy,mobility,andquality of life for visually impaired users***

I. INTRODUCTION

AI-driven assistive technologies have emerged as powerful tools to support individuals with visual impairments, addressing key challenges in navigation, object recognition, and access to visual information. One such innovation is the AI-Driven Vision Assistance mobile application, which functions as a smart guide to enhance environmental awareness for visually impaired users. By leveraging artificial intelligence, the application enables real-time detection of objects, facial recognition, and text reading, all conveyed through intuitive voice feedback. This mobile-based solution aims to fosterindependenceand improve thequality oflifeforusersby minimizing reliance on human assistance or conventional tools such as white canes. Unlike high-end wearable or specializedassistivedevices,smartphone-basedapplicationsofferamoreaccessibleandcost-effective alternative, makingthem increasinglyviable forwidespreadadoptionineverydayscenarios.Assuch, AI-Driven Vision Assistance exemplifies the shift towards affordable, intelligent, and user-centric assistive solutions in the current research landscape.

II. PROBLEM STATEMENT

Visually impaired individuals face significant challenges in navigating their surroundings independently duetothe lackofaccessibleandaffordablevisualassistancetools.Whileadvancementsincomputervisionandartificialintelligencehaveledtothedevelopmentofobjectdetectionsystems, many existing solutions either require expensive hardware, lack real-time performance, or are not optimized for mobile platforms. Furthermore, current assistive technologies often do not provide intuitive feedback mechanisms, such as natural voice output, that can effectively communicate environmental information to users in real-time. There is a clear need for a low-cost, mobile-based solution that can accurately detect and identify objects in the user's environment and convey this information through voice feedback, thereby enhancing situational awareness and autonomy for the visually impaired.

III. OBJECTIVES

- Todevelopamobileapplicationthatutilizesthesmartphonecameratodetectandrecognize objects in real-time using deep learning-based computer vision algorithms.
- To integrate an efficient object detection model (e.g., YOLO, SSD) optimized for mobile platforms, ensuring accurate and fast performance.
- To implement a Text-to-Speech (TTS) system that provides clear and natural voice output to convey the names of detected objects to the user.
- To ensure real-time performance with minimal latency for smooth and uninterrupted user experience.
- To create a user-friendly interface designed with accessibility in mind, making it easy for visually impaired users to interact with the application independently.
- Toevaluatethesystem'sperformanceindifferentreal-worldenvironmentsandlighting conditions to ensure reliability and robustness.

- To provide an affordable and scalable solutionthat can be deployed on commonly usedAndroid smartphones, eliminating the need for specialized hardware.

IV. LITERATURE SURVEY

1) *AI-SenseVision:ALow-CostArtificial-Intelligence-BasedRobustandReal-Time Assistance for Visually Impaired People*

Source:IEEETransactionsonHuman-MachineSystems Year: 2024

Summary:

This paper presents AI-Sense Vision, a portable, standalone assistive device using deep-learning object detection and ultrasonic sensors to support visually impaired people in real-time. The system offers gesture-based mode switching, object recognition, obstacle alerts, and audio feedback through an embedded speaker or earphones. It is trained on an extended COCO-based datasetwithadditionalobjectsrelevanttoVIPsandtestedacrosscomplexenvironmentswith impressive map scores. The device functions offline, with all data processing done on a single-board computer.

Drawback:

The system is not implemented as a mobile application, which limits its accessibility. It depends oncustomhardwareandmanualassembly,whichmaynotbescalableformassadoption.Obstacle detectionisbasicanddoesnotclassifythenatureofhazards. Real-timeperformanceslowsdown in non-GPU systems, possibly affecting responsiveness in dynamic scenarios.

Comparison:

Unlike AI-Sense Vision's standalone hardwareprototype, ourproject delivers the same real-time visual assistance through a mobile app, increasing scalability and reducing deployment barriers.

2) *AnAI-BasedVisualAidWithIntegratedReadingAssistantfortheCompletely Blind*

Source:IEEETransactionsonHuman-MachineSystems Year: 2020

Summary:

This research proposes a low-cost, wearable visual aid mounted on eyeglasses, designed specifically for completely blind individuals. The system uses Raspberry Pi 3B+ along with a camera and ultrasonic sensors for object detection and distance measurement, integrated with a readingassistantusingTesseractOCRandeSpeak.Theentiresystemprovidesaudiofeedbackin real-time and was evaluated with 60 blind participants across various indoor settings. It showed improved mobility and usability compared to the traditional white cane.

Drawback:

The device is hardware-specific and not implemented as a mobile application, limiting accessibilityandeaseofdistribution. Performanceofthereadingassistantdeclinesinpoorlighting and fails to accurately interpret tables or images. It lacks advanced navigation features like GPS or wet floor detection, and the response speed may vary depending on object complexity and environmental conditions.

Comparison:

Unlike this hardware-dependent eyeglass-mounted solution, our mobile app provides a portable, easily accessible alternative forreal-time object and text recognition using a smartphone camera.

3) *AnAssistiveVisionforaVisually ChallengedPerson using Alexa*

Source: 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)

Year: 2024

Summary:

This paper introduces an assistive vision system that leverages Amazon Alexa for visually challenged individuals to identify objects, read printed text, and navigate their surroundings.

The system integrates computer vision, deep learning (Efficient Net and CNNs), and voice command interfacestodeliverreal-timefeedbacktotheuserthroughaudio.Itenablesobjectrecognition, text-to-speech conversion, and smart navigation via Alexa-enabled devices, aiming to improveautonomy and life quality.

Drawback:
Thesystem is highlydependent oncontinuousinternet connectivityandAlexa-compatibledevices, which limits its usability in low-connectivity or low-income settings. Users with speech impairments may find it difficult to interact with the system. It also cannot handle dynamic or complexenvironmentalchangeseffectively,andrelianceonexternalonlinedatabasesmayreduce reliability in real-time situations.

Comparison:
UnlikethisAlexa-dependentsetup,our mobileapplicationprovidesa fully offline, camera-based object and text recognition system with audio output, ensuring accessibility regardless of connectivity or device ecosystem.

4) EmpoweringtheBlind:AI-AssistedSolutionsforVisuallyImpairedPeople
Source:2023IEEEInternationalSmartCitiesConference(ISC2) Year: 2023

Summary:
This paper presents a smart AI-based assistive stick designed to help visually impaired users recognize faces, objects, and colors using OpenCV and Python. The system employs the Haar Cascade and LBPH algorithms for real-time object and facial recognition, providing voice alerts tousersthroughaRaspberryPi.Theaimistoimprovedailymobility,safety,andinteractionusing computer vision, replacing simple vibration cues with detailed verbal descriptions of the surroundings.

Drawback:
The solution relies on custom hardware (a smart stick with Raspberry Pi), which increases cost and complexity for mass adoption. It requires trained datasets for specific environments and individuals,limitingadaptability. Thesystem'seffectivenessisconstrainedby lightingconditionsandcamerapositioning,anditdoesnotoperateonastandardmobileplatform,reducingportability.

Comparison:
Whilethissystemusesasmartstickandembeddedhardware,ourprojectleveragesexistingmobile devices for scalable and app-based real-time object and text detection with voice output.

5) SmartSolutionsforVisualImpairmentbyAI-BasedAssistive Devices
Source: IEEE, 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI) Year:May2024

Summary:
This paper presents a systematic review of recent advancements in assistive technologies for visually impaired individuals, focusing on AI-powered wearable devices, computer vision, deep learning,andedgecomputing. Itexploresvarioussystems likesmartglasses,smartcanes,and AR/VRtoolsthatenablenavigationandobjectrecognitionwithreal-timefeedbackthroughaudio or tactile signals. The review categorizes technologies based on effectiveness and identifies performance trends and user challenges. It emphasizes the role of edge computing and AI in improving latency and detection accuracy.

Drawback:
While the reviewed assistive technologies demonstrate significant potential, several limitations hinder their practical effectiveness. Many systems rely heavily on stable internet connectivity, especiallythoseusingcloud-basedAIservices, whichcanlimitusability inofflineorlow-network environments. Additionally, sensor-based devices often face calibration challenges and tend to producefalsepositives,particularlyincomplexordynamicsurroundings.Wearablesolutionslike smart glasses frequently suffer from short battery life and bulkiness, affecting continuous usage and usercomfort.

Furthermore, high development and deployment costs, as well as social stigma associated with wearing noticeable assistive gear (like Google Glass), present barriers to widespread adoption. These limitations highlight the need for more user-friendly, reliable, and cost-effective solutions.

Comparison:
Unlikethisreview'semphasisonwearableorinfrastructure-heavysolutions,yourprojectfocuses on a lightweight, real-time mobile app that provides object detection and audio feedback without requiring additional hardware, increasing accessibility and affordability.

6) AssistingtheVisuallyImpairedPeopleThroughAI-AssistedMobility
Source:InternationalJournalofInnovationsinScience&Technology(IJIST),SpecialIssue Year: 2024

Summary:
This paper introduces *"Visually"*, a mobile application designed to assist visually impaired individuals by integrating real-time object detection (YOLOv5), face recognition, and currency recognition, all with Text-to-Speech (TTS) feedback. It uses TensorFlow Lite for on-device processing,ensuringofflineavailability,andsupportscross-platformfunctionalitythroughFlutter. The app was trained on an augmented dataset tailored for everyday objects and achieves high accuracy(90–99%)inidentifyingrelevantitemslikepeople,laptops,vehicles,andfurniture.The paper also benchmarks its performance against existing apps like Seeing AI and Be My Eyes.

Drawback:
Despiteitsstrengths,the"Visually"applicationfacesafewlimitations.Theperformanceofobject detection, while generally strong, may vary in highly cluttered or dynamic real-world environmentsduetovisualnoise. TheuseofYOLOv5 requiresmoderatecomputationalresources, whichmaylimit smoothperformanceonolderorlow-endsmartphones.Whileofflinefunctionality isamajoradvantage,italsomeanstheappmaynotbenefitfromcloud-basedmodelimprovements or real-time updates unless manually refreshed. Additionally, the app currently depends heavily on audio feedback, which might not be ideal in noisy environments without haptic support.

Comparison:
Comparedtoyourproject,whichemphasizesreal-timeobjectdetectionviacameraandimmediate voice feedback through a lightweight mobile app, "Visually" offers similar features but includes additional modes like face and currency recognition, with higher complexity and resource needs.

7) VisuallyImpairedAssistancewithLarge Models
SourceIEEESmartWorldCongress(SWC) Year: 2024
Summary:
ThispaperintroducestheVIALM(VisuallyImpairedAssistancewithLargeModels)framework, whichevaluateshowstate-of-the-artlargevision-languagemodels(VLMs)—likeGPT-4,LLaVA, and CogVLM—can assist visually impaired individuals. It proposes a benchmark task where a user provides an image and a verbal request, and the model generates detailed, tactile-aware guidance. The study uses 200 annotated scenarios and evaluates model outputs on correctness, clarity, and actionability. Results show GPT-4 and LLaVA perform best, but still face key challenges in real-world grounding and fine-grained instruction.

Drawback:
Despite leveraging powerful large-scale models, the study identifies notable shortcomings in currentvision-languagesystems.Mostmodelsstrugglewithenvironmentalgrounding,producing vague or incorrect instructions in complex environments like supermarkets. Even GPT-4, the top performer, showed a 25.7% grounding failure rate and a 32.1% lack of actionability, often generating verbose, less actionable outputs. Other models such as BLIVA and MiniGPT-v2 performed worse, lacking both detail and tactile guidance. The models also generally failed to adapttotactile-centricneedsofvisuallyimpairedusers,revealingagapinmultimodalintegration and user-context awareness.

Comparison:
While this study explores advanced LLMs for generating detailed, task-specific assistance, your project offers a lightweight, mobile-based real-time object recognition and voice feedback system focused on practical, immediate usability without relying on large-scale generative models.

8) MobileApplication:MobileAssistanceforVisuallyImpairedPeople –SpeechInterface System (SIS)
Source:IEEE8thInternationalConferenceonInformationTechnologyandMultimedia(ICIMU) Year: 2020

Summary:
This paper presents a Systematic Literature Review (SLR) on the development of speech-driven mobileapplicationstoassistvisuallyimpaired(VI)users.ItemphasizestheuseofSpeechInterface Systems (SIS) combining speech recognition, object detection, and distance estimation to aid in navigation and object identification. The review spans multiple techniques including Text-to- Speech (TTS), Scene-to-Speech (STS), and speech recognition models like HMMs and neural networks, concluding that smartphones are viable platforms for inclusive assistive technology.

Drawback:
While the review showcases various advancements in speech-based assistance for the visually impaired, it reveals several limitations. Many existing systems lack real-time processing efficiency, particularly when combining multiple functions like object detection and navigation. Thereliance oncloud APIs (e.g., GoogleSpeech Recognition)raises privacyconcerns andlimits offline usage. Another issue is limited language support, which poses barriers for non-English- speaking users. Furthermore, many systems depend heavily on command-specific input, lacking natural conversational interfaces, which can hinder usability in diverse, spontaneous situations.

Comparison:
Unlikethisreview'semphasisonspeech-commandsystemsandliteraturesurveys,yourprojectis focused on real-time, camera-based object detection and immediate audio feedback via a standalone mobile application, offering more direct, context-aware interaction for visually impaired users.

9) WearableVisionAssistanceSystemBasedonBinocularSensorsforVisuallyImpaired Users
Source:IEEESensorsJournal Year: 2022
Summary:
This paper introduces a wearable vision assistance system using a binocular RGB-depth sensor mounted on glasses to assist visually impaired users with real-time scene understanding. It incorporates deep learning for object detection (YOLOv4), scene captioning, and OCR to inform usersofobstacles,objecttypes,andsurroundings.Voicecommandstriggerspecificfunctions,and the processed results are conveyed via audio output through headphones. The system runs on an NVIDIA Jetson TX2 for edge computation, balancing mobility with computational efficiency.

Drawback:
While the system demonstrates robust performance in diverse environments, it comes with practical limitations. The need for specialized hardware like the NVIDIA Jetson TX2 and a binocular sensor setup reduces portability and makes it less accessible to the average user. Additionally, wearing a headset and sensors may lead to user discomfort and social acceptability concerns. The system also consumes a considerable amount of power, limiting its usage duration inreal-worldsettings.Lastly,real-timeperformancemaydegradeunderpoorlightingoroccluded scenes, affecting object recognition accuracy.

Comparison:
Unlikethiswearable,hardware-dependentsystem,yourmobile-basedappoffersamoreaffordable and accessible solution by leveraging existing smartphone cameras and speakers to deliver real- time object detection and voice feedback.

10) NavigationAssistancefortheVisuallyImpairedUsingRGB-DSensorWithRange Expansion
Source:IEEESystemsJournal Year: 2016
Summary:
ThispaperpresentsawearablenavigationsystemforvisuallyimpairedindividualsthatcombinesRGBD(depth+visual)sensordatatodetectobstacle-freepathsinreal-time.Thecoreinnovation lies in fusing range and color information to expand the effective navigation range, overcoming depth sensor limitations. Audio feedback is delivered via stereo beeps and voice commands, allowing users to interpret proximity and orientation of obstacles. The system was validated in various real-world indoor environments with high precision and recall in floor segmentation.

Drawback:
Despite strong performance, the system has practical limitations. It relies on wearable RGB-D sensors and a backpack-mounted laptop, makingit bulky andless user-friendly for everyday use. Thesystem runs at arelatively low frame rate (~0.3 fpsfor full processing), limiting its real-time responsiveness. Moreover,thedepth sensoris sensitiveto sunlight, reducing accuracy in brightly litoroutdoorscenarios.Audiocues,whileeffective,canoccasionallyinterferewiththeperception of environmental sounds, potentially affecting user awareness in dynamic surroundings.

Comparison:
In contrast to thishardware-heavy wearable system, your mobile app provides a more portable anduser-friendlyreal-timesolutionbyusingjustthesmartphonecameraandvoicefeedbackfor object detection.

V. RESEARCH GAP

Overdependence on Specialized or Wearable Hardware: Wearable Vision Assistance System BasedonBinocularSensors.Manysystemsrelyonbulkyorexpensivesetups(e.g.,JetsonTX2,RGB- D cameras, backpacks), limiting portability, affordability, and mainstream adoption.
Limited Real-Time Performance on Lightweight Devices: Use of large models or low frame rates (0.3 fps) impedes real-time feedback on mobile or low-end devices. Few systems demonstrate optimized object detection on smartphones with sufficient speed and accuracy. VIALM with Large Models.
Lack of Fully Offline Functionality: Some solutions depend on cloud-based services for speech recognition or model inference, which limits accessibility in areas with poor connectivity. Offline- capable systems remain underexplored.
Neglect of User Comfort and Social Acceptability: Wearable solutions often ignore factors like long-term wear comfort, battery life, and social stigma. Lightweight, familiar form factors (e.g., phones) are less commonly used despite high acceptability.
Lack of Focus on Immediate, Low-Complexity Tasks: Many advanced models focus on complex sceneunderstandingorsemanticreasoning,which,whilepowerful,canbeoverkillfordaily-usetasks like identifying a chair, bag, or vehicle in real time.

VI. CONCLUSION

Thereviewofexistingliteraturehighlightssignificantadvancementsinthedomainsofcomputer vision, deep learning, and assistive technologies aimed at aiding visually impaired individuals.
Various object detection models such as YOLO, SSD, and Faster R-CNN have demonstrated high accuracyandefficiencyinreal-timeenvironments.Additionally,text-to-speech(TTS)systemsand mobile integration of these models have enabled practical and portable solutions for visual assistance.Despitetheprogress, many existing solutions either lack real-time performanceon mobile devices or fail to provide accurate contextual awareness in dynamic environments. Furthermore, accessibility,affordability,anduser-friendlinessremaincriticalchallengesinthedeploymentofsuch technologies at scale.This project seeks to bridge these gaps by developing a mobile application that leverages the smartphone camera to detect and identify objects in real time and provide audio feedback through voice commands. By integrating robust object detection algorithms with efficient TTS engines, the proposed solution aims to offer an effective, user-friendly, and real-time visual aid tool for the visually impaired community

REFERENCES

[1] R.C.Joshi,N.Singh,A.K.Sharma,R.Burget,andM.K.Dutta,"AI-SenseVision:ALow- Cost Artificial-Intelligence-Based Robust and Real-Time Assistance for Visually ImpairedPeople,"IEEETransactionsonHuman-MachineSystems,vol.54,no.3,pp.325–336,Jun. 2024.

[2] M.A.Khan,P.Paul,M.Rashid,M.Hossain,andM.A.R.Ahad,"AnAI-BasedVisualAid With Integrated Reading Assistant for the Completely Blind," IEEE Transactions on Human- Machine Systems, vol. 50, no. 6, pp. 507–516, Dec. 2020.

[3] C. Dheeraj, K. S. Reddy, and R. Rajalakshmi, "An Assistive Vision for a Visually ChallengedPersonusingAlexa, "ProceedingsofInternationalConferenceonComputerScienceandEngineering,Sathyabama Instituteof ScienceandTechnology,Chennai,India.

[4] M.S.Qureshi,I.U.Khan,S.M.B.Qureshi,F.M.Khan,andS.Aleshaiker,"Empowering the Blind: AI-Assisted Solutions for Visually Impaired People," Proceedings of the International Conference on Computer Engineering and Technology, 2023.

[5] S. R. Katke and U. Pacharaney, "Smart Solutions for Visual Impairment by AI-Based AssistiveDevices, "FacultyofEngineering andTechnology,DMIHER(DU),India,2024.

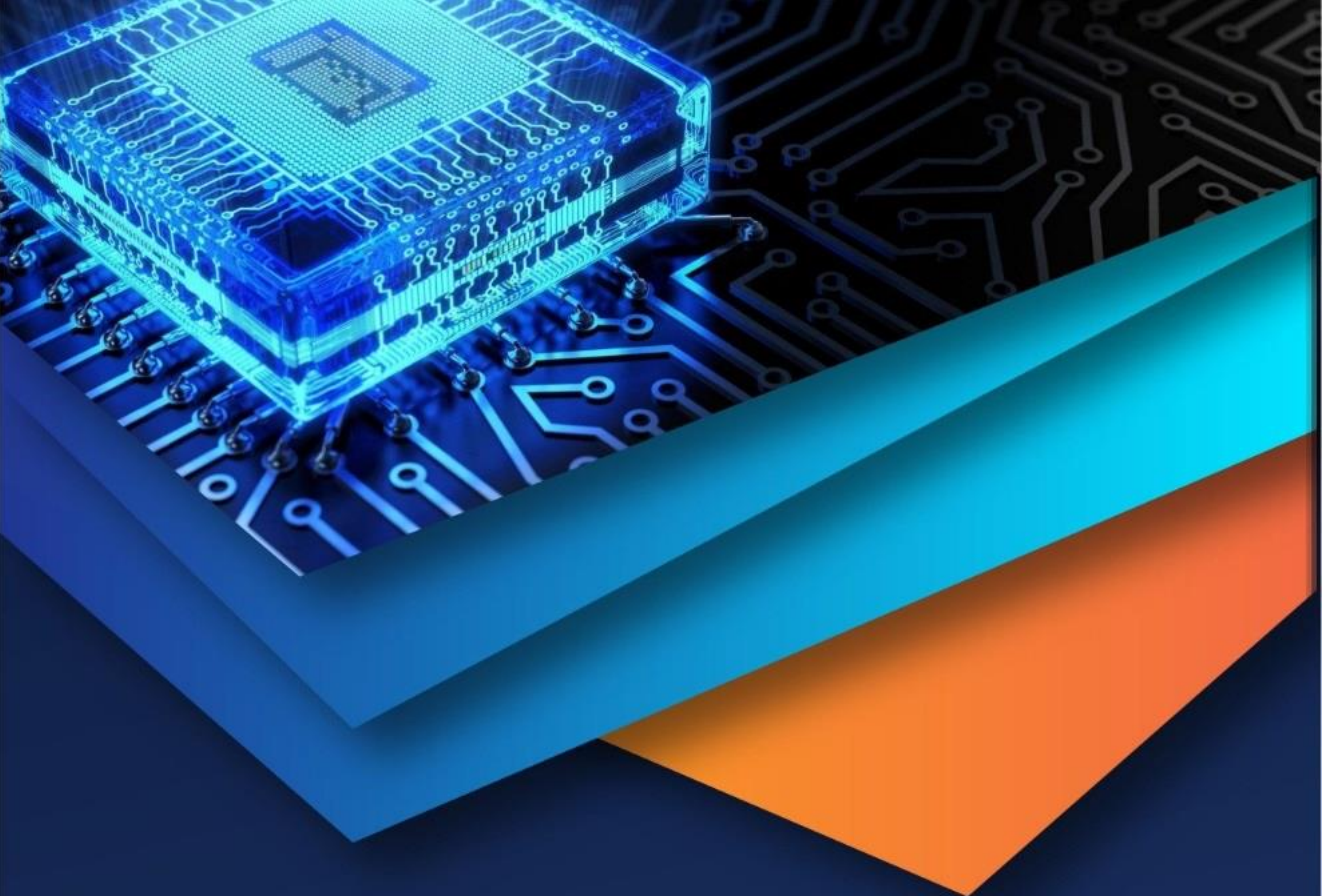
[6]]Y.S.Afridi,M.Sher,andI.Jr,"Visually:AssistingtheVisuallyImpairedPeopleThrough AI-Assisted Mobility," International Journal of Scientific and Technology Research, vol. 13, no. 4, pp. 115–122, May 2024

[7] R.Xiang,Y.Zhao,Y.Zhang,J.Li,M.Liao,andY.Li,"VisuallyImpairedAssistancewith Large Models," IEEE Smart World Congress (SWC), 2024.

[8] A. M. Norkhalid, M. A. Faudzi, A. A. Ghapar, and F. A. Rahim, "Mobile Assistance for VisuallyImpairedPeople–SpeechInterfaceSystem (SIS),"20208thInternationalConference on Information Technology and Multimedia (ICIMU), pp. 329–334, 2020.

[9] AbdulMajidNorkhalid1,a,MasyuraAhmadFaudz"WearableVisionAssistanceSystemBased on Binocular Sensors for Visually Impaired" IEEE Sensors Journal,2022.

[10] A. Aladrén, G. López-Nicolás, L. Puig, and J. J. Guerrero, "Navigation Assistance for the VisuallyImpairedUsingRGB-DSensorWithRangeExpansion ,"IEEESystemsJournal,vol. 10, no. 3, pp. 922–933, Sep. 2016.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY