



---

# Voice-guided Mobile Assistance for the Visually Impaired

---

Yash Khopkar<sup>1\*</sup>, Avantika Deshmukh<sup>2</sup>, Prof. Gufran Ansari<sup>3</sup>

<sup>1\*,2</sup>MCA Science, MIT WPU, Pune, Maharashtra, India.

<sup>3</sup>MIT WPU, Pune, Maharashtra, India.

Corresponding Email: <sup>1\*</sup>ykhopkar3@gmail.com

**Received:** 28 September 2023 **Accepted:** 17 December 2023 **Published:** 01 February 2024

*Abstract: In today's digital age, mobile apps have transformed our daily lives, but for people with visual impairments, these apps often present accessibility challenges. This research addresses the need for improved solutions by focusing on "voice-guided assistance" for Android smartphones. The existing options for visually impaired users are fragmented, and this study aims to create an integrated mobile application, "Voice-Companion," designed to enhance their digital experience. "Voice-Companion" is a specialized Android application designed for visually impaired individuals, developed in Java within the Android Studio environment. It leverages the Android OS to provide comprehensive non-visual access. With features like voice commands, object detection, messaging, a voice-activated calculator, location and time announcements, text-to-speech capabilities, and battery status updates, the user-friendly app bridges accessibility gaps, empowering visually impaired users for equal participation in the digital realm. Its modular, user-centric design emphasizes accessibility through seamless integration of voice commands and gesture recognition, facilitating efficient interactions with mobile devices. This research aims to create "Voice-Companion," enhancing smartphone accessibility through seamless voice command integration and gesture recognition. Anticipated results include improved digital engagement and quality of life for users with visual impairments.*

**Keywords:** Voice-Guided Assistance, Text-to-Speech Integration, Android Accessibility, Object Detection, Optical Character Recognition (OCR), Visual Impairment.

## 1. INTRODUCTION

In this era of rapid technological progress, mobile apps have become an essential part of our daily existence, revolutionizing how we communicate, work, and interact with the world. Nevertheless, individuals with visual impairments often find it difficult to access the advantages of these apps due to accessibility challenges. Visually impaired people encounter



distinct obstacles when it comes to using mobile devices, including challenges with visual content, complex user interfaces, inconsistent accessibility standards, small text sizes, touchscreen interaction, limited voice control, a lack of descriptive information, restricted access to visual content, and compatibility issues. These barriers can hinder their ability to fully participate in the digital world, despite mobile apps becoming an essential part of our daily existence. Visual impairment encompasses a wide spectrum of conditions, ranging from partial vision loss to complete blindness, impacting various aspects of an individual's life. Modern smartphones, especially those running on the Android operating system, offer a wide array of features and functionalities that can be leveraged to bridge the accessibility gap for visually impaired users.

This research paper aims to explore potential solutions for individuals facing visual impairments. It focuses on the utilization of "voice-guided assistance" within smartphones, particularly those operating on the Android platform. This technology can make it easier for people with vision problems to use these apps and be a part of the digital world, just like everyone else.

### **Contribution**

The primary contribution of this research is: -

1. We have developed the "Voice-Companion" Android application with the aim of addressing the accessibility challenges encountered by individuals with visual impairments.
2. This application provides an integrated solution with features such as voice commands, object detection, messaging, a voice-activated calculator, location and time announcements, text-to-speech capabilities, and battery status updates, all aimed at enhancing the digital experience for the visually impaired.
3. We have designed a system architecture outlining the seamless integration of multiple features within the application.

This paper is organized as follows: In Section II, we provide a comprehensive literature review of related work in the field of assistive technology for visually impaired individuals, including the latest developments. Section III outlines the methodology used in the development of the "Voice-Companion" application, with proposed steps presented in flow chart form. Section IV presents the results of the application, including data tables, a comparison table, and a graph to illustrate its performance. Finally, in Section V, we conclude by discussing the significance of "Voice-Companion" in promoting inclusivity and accessibility for visually impaired individuals in the digital age.

### **Problem Statement**

Visually impaired individuals face significant accessibility challenges that make it difficult for them to effectively engage with the digital world. The existing solutions available to them are scattered, failing to provide a complete answer to their various needs. To address this gap in support, our research aims to create an integrated mobile application designed specifically for visually impaired users. This application is meant to provide a unified and inclusive



solution that helps these individuals navigate and use digital resources more effectively. This, in turn, will improve their overall quality of life and their ability to participate in today's digital world.

### **Literature Review**

In the paper "Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired People," Aaron Raymond See, Bien Grenier Sasing, and Welsey Daniel Advincula introduce a smartphone tool for visually impaired individuals. It uses depth imaging, deep learning for obstacle detection, and integrates Google Maps for navigation. The system is effective indoors and outdoors, offering accurate obstacle detection, low latency, and positive user feedback [1]. Abhigyan Baruah, Aryan Dev, Jasowanta Das, and Santanu Kumar Misra present an "Android-Based Assistance System for Visually Impaired Person Using Object Recognition." Their Android system employs smartphone cameras and neural networks for object recognition, achieving impressive accuracy. However, it has limitations related to lighting and object size [2]. Abhishek More, Tushar Gayakwad, Mansi Suryawanshi, Suraj Kshirsagar, and Prof. Pragati Deole introduce an "Android Application for Visually Impaired People Based on AI Technology." Their app uses AI for optical character recognition, object recognition, and more. "VISION EYE" empowers visually impaired individuals with object and currency recognition and IoT integration potential [3].

Kuei-Chun Liu, Ching-Hung Wu, Shau-Yin Tseng, and Yin-Te Tsai present "Voice Helper," an Android-based assistive system for visually impaired users. It offers features like message reading, voice dialling, and QR code recognition, addressing currency recognition challenges and anticipating advancements in assistive tech [4]. Sriraksha Nayak and Chandrakala C. B. introduce an "Assistive Mobile Application for Visually Impaired People." Their mobile app enhances communication, information access, and independent navigation through speech technology [5]. Felipe Mendonça Gouveia, Byron Leite Dantas Bezerra, Cleber Zanchettin, and Joao Raul Jardim Meneses present a "Handwriting Recognition System for Mobile Accessibility to the Visually Impaired People," facilitating access to handwritten content [6]. Joe Yuan Mambu, Gerent Keyeh, Elisa Anderson, Andria Wahyudi, and Billy Dajoh present "Blind Reader," a mobile app using augmented reality for object recognition by visually impaired users, enabling independent object identification through image capture and audio descriptions [7]. In their work titled "A Survey on Assistive Technology for the Visually Impaired," Shivang Tripathi, Chandrani Halder, and D. Vanusha provide an overview of recent advancements in technologies aimed at aiding individuals with visual impairments. The current technologies empower users to detect and distinguish objects, navigate spaces, recognize text, and receive feedback through vibrations and audio cues. The research emphasizes mobile applications capable of object recognition and categorization through captured images, product differentiation, and indoor navigation guidance for individuals with visual impairments [8]. In their paper titled "An Intelligent Voice Assistance System for The Visually Impaired People," Surabhi Suresh, Sandra S, Aisha Thajudheen, Subina Hussain, and Amitha R introduce a novel assistive technology designed for individuals with visual impairments. This system employs Raspberry Pi due to its cost-effectiveness, compact size, and ease of integration. Utilizing a camera and ultrasonic sensors, the system gauges the



user's proximity to obstacles. It integrates an image-to-text conversion feature and offers auditory feedback. The primary objective is to assist visually impaired individuals in navigating both indoor and outdoor environments by identifying objects and providing timely alerts. Moreover, the system includes a reading aid for those with visual impairments and a pulse oximeter that can send an SMS or call for emergency assistance when the individual is at risk [9]. Ms. Kavya. S, Ms. Swathi, and Mrs. Mimitha Shetty introduce a system titled "Assistance System for Visually Impaired using AI," with the objective of aiding individuals with visual impairments in their daily activities. The system is implemented through an Android mobile application, prioritizing features like a voice assistant, image recognition, currency identification, e-book capabilities, and a chat bot. The application is proficient in guiding users via voice commands to identify objects in their surroundings and analyze text, enabling recognition of content in printed documents. This technology offers a promising means for visually impaired individuals to engage with others and maintain an independent lifestyle [10]. R. R. Varghese, P. M. Jacob, M. Shaji, A. R, E. S. John, and S. B. Philip introduce a novel assistive technology for individuals with visual impairments in their work titled "An Intelligent Voice Assistance System for Visually Impaired using Deep Learning." Their system utilizes a Raspberry Pi 3 Model B+ for its cost-effectiveness, compact design, and ease of integration. The setup incorporates a camera and sensors for obstacle detection, employing robust image processing algorithms to identify and categorize various objects. The combination of a camera and ultrasonic sensors enables the system to calculate the user's distance from obstacles. Additionally, an image-to-text converter is integrated, followed by providing auditory feedback. This entire setup can be conveniently affixed to a standard pair of eyeglasses, ensuring it remains lightweight, compact, and user-friendly. Through a study involving 60 individuals with complete blindness, the authors compared the proposed device to the traditional white cane in terms of performance. The assessments were conducted in controlled settings mirroring real-world scenarios. The outcomes revealed that the proposed device offered enhanced accessibility, comfort, and ease of navigation for visually impaired individuals when compared to using a white cane [11].

Jiawen Li, Lianglu Xie, Zhe Chen, Liang Shi, Rongjun Chen, Yongqi Ren, Leijun Wang, and Xu Lu introduce "An AIoT-Based Assistance System for Visually Impaired People," a system designed to enhance convenience for individuals with visual impairments. This system is devised and executed utilizing the Artificial Intelligence of Things (AIoT) framework. It remains cost-effective and versatile, integrating features like object detection, obstacle distance measurement, and text recognition through wearable smart glasses. Furthermore, it encompasses heart rate detection, fall detection, body temperature measurement, and humidity-temperature monitoring facilitated by an intelligent walking stick. The overall hardware cost is approximately \$66.8. A voice assistant is employed to communicate detection results to users. The assessment of performance demonstrates high accuracies, with object detection at 92.16% and text recognition at an impressive 99.91% during experiments involving the wearable smart glasses [12]. The authors propose a novel visual aid using Mobile Edge Artificial Intelligence (MIEI) for the visually impaired. This system integrates deep learning, point cloud processing, model optimization (OpenVINO, TensorFlow Lite), smart depth sensors, and an AI-driven voice interface, enabling advanced



perception tasks on cost-effective, low-power mobile platforms. Unlike traditional GPU-based setups, this design bypasses high-cost, power-intensive hardware for real-time deep learning inference. It surpasses existing systems by assessing traffic, detecting hangouts, crosswalk obstacles, moving objects, and sudden elevation changes, among other features [13]. Gagandeep Singh, Omkar Kandale, Kevin Takhtani, and Nandini Dadhwal introduce "A Smart Personal AI Assistant for Visually Impaired People." This paper outlines a technology-driven approach to aid individuals with visual impairments, utilizing computer science, voice recognition, and image recognition. The proposed solution involves creating an Android application capable of providing assistance through voice commands, recognizing and analyzing the user's environment, and offering relevant responses. The application leverages the Google Vision API for image recognition and includes features for text recognition from .pdf files, subsequently utilizing a Google-powered text-to-speech engine for auditory presentation [14].

## **2. METHODOLOGY**

"Our 'Voice-Companion' Android app was developed using Java as the primary programming language within the Android Studio environment. This choice ensures the app's robustness and adaptability, allowing it to operate smoothly on a wide range of Android devices. At the core of 'Voice-Companion', we've established a modular and user-friendly architecture with a strong focus on accessibility. The heart of our app revolves around the seamless integration of voice commands and gesture recognition, positioning these features as the central means for visually impaired users to interact with the application.

This all-in-one solution will encompass a voice assistant, OCR technology for converting printed text to speech, object detection capabilities for environment interaction, call functionality, voice messaging, a voice-activated calculator, location announcements, battery status updates, and real-time date and time announcements. These integrated features make 'Voice-Companion' a comprehensive and invaluable tool for visually impaired users, enhancing their daily lives and digital interactions."

### **Voice Commands and Gestures**

Efforts were dedicated to ensuring the user-friendliness and responsiveness of these features. Users can effortlessly issue voice commands and execute simple swipes, simplifying navigation within the app. By prioritizing intuitive and straightforward interactions, 'Voice-Companion' aims to empower visually impaired individuals by providing them with a natural and efficient way to access its functionalities and features. The fluid integration of voice commands and gestures enhances the overall usability and accessibility of the application.



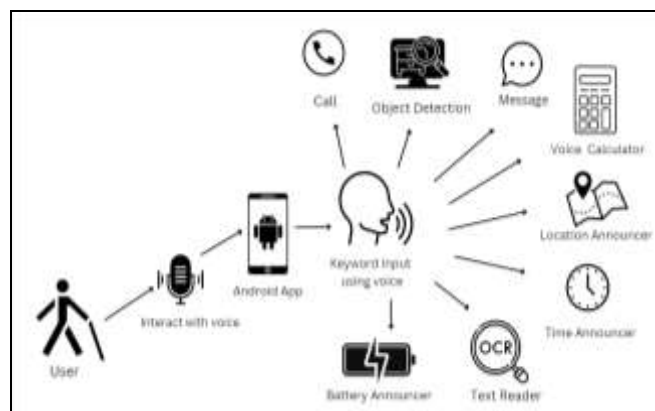


Fig 1: System Architecture

The integration of various features within the application was a pivotal aspect of its development. Here's a brief overview of how these features were seamlessly integrated:

### Calling

Our application features a dedicated "Calling" module designed with the specific needs of visually impaired users in mind. This module simplifies the process of making phone calls using voice commands. Users can effortlessly initiate calls by either stating the contact's name or directly providing the contact's phone number. Once the necessary information is provided, the app swiftly initiates the call, offering a seamless and hands-free calling experience. To boost user confidence and interaction, we've integrated haptic feedback, ensuring users receive tactile confirmation when their call is successfully launched.

### Object Detection

Our app employs cutting-edge technology, powered by TensorFlow Lite models and the device's camera, to achieve real-time object detection. This dynamic capability enables the app to identify a diverse array of objects within the user's environment. Whether it's everyday items, tools, or obstacles, our app excels in delivering precise and prompt recognition. Upon detecting an object, "Voice-Companion" swiftly responds by providing a clear and informative auditory description. This feature transcends mere object identification; it imparts users with a deeper understanding of their surroundings. By enriching their environmental awareness, our app empowers visually impaired individuals to interact with the world more independently and confidently. Furthermore, our object detection capabilities encompass a wide spectrum of objects, ranging from household items to outdoor elements. This comprehensive recognition ensures that users can effortlessly engage with their surroundings, fostering greater autonomy and accessibility.

### Message

Our app makes messaging a breeze with the "Messaging" module. Users can send text messages using their voice, making communication quick and easy. What's more, we've added a special feature – you can record a voice message, and our app will turn it into text before sending it. This personalizes your messages and adds an expressive touch. Whether



you want to message someone by saying their name or providing their phone number, our "Messaging" module ensures it's convenient and easy, just the way you like it.

### **Voice Calculator**

Visually impaired users can activate the Voice Calculator with a left swipe gesture, allowing them to perform calculations effortlessly by providing voice input. The app employs robust speech recognition technology to convert spoken words into text, ensuring accurate interpretation of mathematical expressions. It utilizes the powerful 'ExpressionBuilder' from the 'net.objecthunter.exp4j' library to precisely evaluate mathematical expressions and provide an audible output, making complex calculations accessible and straightforward for visually impaired individuals.

### **Location Announcer**

Our app includes a handy "Location Announcer" feature that provides users with their current location using GPS technology. Once the app fetches the user's location, it promptly announces it aloud, ensuring users are always aware of where they are. This feature is particularly useful for visually impaired users, as it enables them to navigate confidently and independently. With just a simple command, users can get instant location information, making their daily journeys more convenient and secure.

### **Time Announcer**

Our app features a handy "Time and Date Announcer" module designed to provide users with the current date and time. Using this module is as easy as a swipe. Once activated, the app fetches the current date and time, and then announces it aloud. Visually impaired users can stay informed about the time and date, enhancing their daily routines and activities. A simple swipe allows users to repeat the announcement or return to the main page, ensuring an accessible and user-friendly experience.

### **Text Reader**

OCR (Optical Character Recognition) feature is integrated seamlessly. It uses the device's camera to capture text from physical documents or surfaces and then converts that text into speech, allowing users to access printed information independently.

### **Battery Announcer**

Battery Announcer in "Voice-Companion" provides the current device's battery level when users request it. It operates on a user-initiated basis, announcing the battery percentage when asked, ensuring on-demand accessibility and convenience for visually impaired users.

### **Text-To-Speech Engine**

A Text-to-Speech (TTS) engine is a software component that synthesizes text into audible speech, enhancing accessibility and user interaction in applications. It offers language customization and speech rate control to cater to diverse user preferences.

These are the key components and functionalities of a Text-to-Speech (TTS) engine within a mobile application:

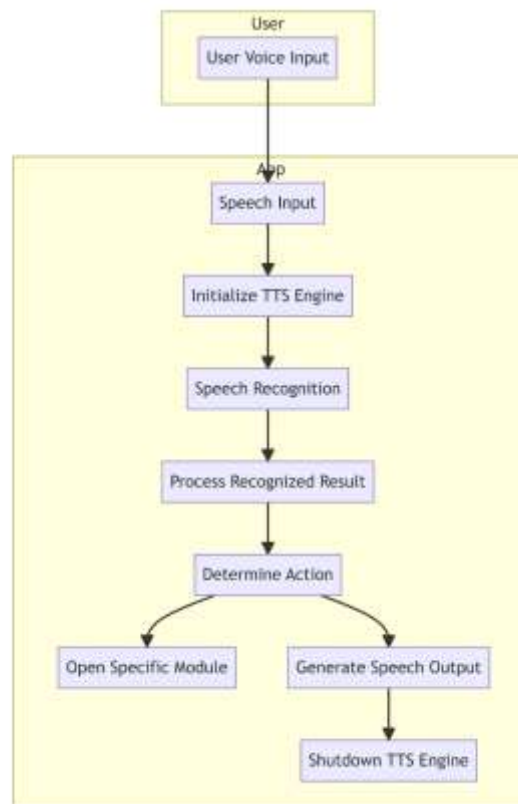


Fig: Text-To-Speech Engine

### TTS Initialization

The TTS engine initialization is a fundamental component of the application. It configures the Text-to-Speech engine, defining settings like language and speech rate, ensuring an optimal auditory experience for users.

### User Interaction

Users engage with the TTS engine by employing touch gestures and voice commands through the application's interface. These interactions enable users to interact seamlessly with the app's auditory capabilities.

### Speech Recognition

Speech recognition is a pivotal feature that transforms spoken words into actionable commands. This technology, often powered by tools like RecognizerIntent, empowers users to provide voice input naturally and efficiently.

### Voice Commands

Upon recognizing user voice input, the TTS engine responds intelligently to voice commands. It interprets and executes commands, making the application responsive to user needs, whether it's launching specific functionalities or delivering information.





**Error Handling**

Table. 1

<b>Feature</b>	<b>Voice-Companion</b>	<b>Existing Solutions</b>
Voice Commands	Highly responsive	Varied
Object Detection	Real-time, precise, slow	Limited accuracy
Messaging	Voice-to-Text	Text-only
Voice Calculator	Accurate results	Limited features
Location Announcer	Live Address, GPS-based accuracy	Basic information
Time Announcer	Real-time updates	Manual checks
Text Reader (OCR)	Medium Accuracy and fast	Slow and error-prone
Battery Announcer	On-demand access	Manual checks
Text-To-Speech Engine	Fast, occasional delays	Standard

The application incorporates robust error-handling mechanisms. In situations where the device lacks support for speech input or errors occurs during voice processing, the application provides informative feedback, enhancing the user experience.

**Integration with Other App Modules**

The TTS engine serves as a central element, seamlessly integrating with various application modules. These modules encompass a wide range of functionalities, from mathematical calculations and date/time announcements to battery status updates. The TTS engine enhances these modules by providing auditory information and feedback.

**3. RESULTS AND DISCUSSION**

We present the results of the "Voice-Companion" application in this section. The following table provides a comparison of key features and performance metrics of our application with existing solutions:

**Comparison Table: Voice-Companion vs Existing Solution**

**Graphical Representation: Voice-Companion vs Existing Solution**

Voice-Companion's voice command recognition system is highly responsive, ensuring swift and accurate execution of voice commands. In contrast, existing solutions exhibit varying degrees of responsiveness, which can lead to inconsistent user experiences.

Voice-Companion offers real-time and precise object detection, although there might be a slight delay in processing the results. On the other hand, existing solutions struggle with accuracy in detecting objects, which can be a significant limitation.

Voice-Companion's voice-activated calculator ensures accurate and reliable results, making it a valuable tool for users. In contrast, existing solutions lack advanced features, limiting their functionality in mathematical tasks.

Voice-Companion enhances the user experience by providing real-time announcements with live addresses and GPS-based accuracy, improving navigation and location-based interactions.

Existing solutions offer only basic location information, which may not be as informative or helpful.

Voice-Companion ensures users stay updated with real-time time-related information, eliminating the need for manual time checks. In contrast, existing solutions rely on manual checks, which can be inconvenient for users.

Voice-Companion's text reader uses optical character recognition (OCR) with medium accuracy and fast text-to-speech conversion, improving the accessibility of printed text. Existing solutions tend to be slower and more error-prone in text reading, which can frustrate users.

Voice-Companion allows users to access battery status on-demand, providing timely information about their device's power level. The exact alternatives for existing solutions will be determined based on user needs.

Voice-Companion features a fast text-to-speech engine, ensuring efficient conversion of text to speech. Although it may experience occasional delays, it outperforms existing solutions that use standard text-to-speech engines in terms of speed and responsiveness.

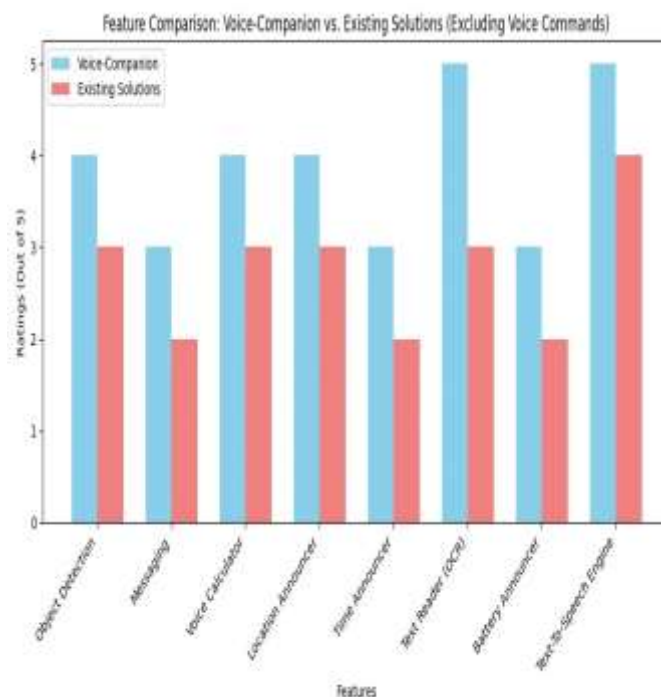


Fig: Voice-Companion vs Existing Solution

#### 4. CONCLUSION

"Voice-Companion" isn't just any app; it's a transformative tool designed for people with visual impairments. Its main goal is to bridge the accessibility gap and make navigating the digital world easier and more efficient for them. This comprehensive solution meets their unique needs and challenges, empowering them to confidently engage in the digital age. By smoothly combining advanced technology with a user-friendly approach, "Voice-



Companion" ensures that visual impairments do not stop or slow down participation or independence. It aims for inclusivity, changing the digital world into a place where everyone, no matter their vision issues, can actively take part and benefit from the opportunities of the digital age. "Voice-Companion" is more than just software; it symbolizes a big step toward a fairer and more inclusive society where everyone's voice is valued, and everyone has the chance to succeed.

## 5. REFERENCES

1. Aaron Raymond See, Bien Grenier Sasing, Welsey Daniel Advincula. "Smartphone-Based Mobility Assistant Using Depth Imaging for Visually Impaired People." *Applied Sciences*, 2022.
2. Abhigyan Baruah, Aryan Dev, Jasowanta Das, and Santanu Kumar Misra. "Android-Based Assistance System for Visually Impaired Person Using Deep Learning and Augmented Reality." *Lecture Notes in Electrical Engineering*, 2021.
3. Abhishek More, Tushar Gayakwad, Mansi Suryawanshi, Suraj Kshirsagar, Prof. Pragati Deole. "Android Application for Visually Impaired People Based on AI Technology." *International Research Journal of Modernization in Engineering Technology and Science*, 2022.
4. Kuei-Chun Liu, Ching-Hung Wu, Shau-Yin Tseng, Yin-Te Tsai. "Voice Helper: A Mobile Assistive System for Visually Impaired Persons." *IEEE Transactions on Human-Machine Systems*, 2015.
5. Sriraksha Nayak, Chandrakala C. B. "Assistive Mobile Application for Visually Impaired People." *ResearchGate*, 2020.
6. Felipe Mendonça Gouveia, Byron Leite Dantas Bezerra, Cleber Zanchettin, Joao Raul Jardim Meneses. "Handwriting Recognition System for Mobile Accessibility to the Visually Impaired People." *IEEE Xplore*, 2014.
7. Joe Yuan Mambu, Gerent Keyeh, Elisa Anderson, Andria Wahyudi, Billy Dajoh. "Blind Reader: An Object Identification Mobile-Based Application for the Blind Using Augmented Reality Detection." *IEEE Xplore*, 2019.
8. Shivang Tripathi, Chandrani Halder, D. Vanusha. "A Survey on Assistive Technology for the Visually Impaired." *IRJET*, 2018.
9. Surabhi Suresh, Sandra S, Aisha Thajudheen, Subina Hussain, Amitha R. (2023). "An Intelligent Voice Assistance System for the Visually Impaired People." *International Journal of Engineering Research & Technology (IJERT) Volume 11, Issue 04*.
10. Ms. Kavya. S, Ms. Swathi, Mrs. Mimitha Shetty. (2019). "Assistance System for Visually Impaired Using AI." *International Journal of Engineering Research & Technology (IJERT) RTESIT – 2019 (Volume 7 – Issue 08)*.
11. R. R. Varghese, P. M. Jacob, M. Shaji, A. R, E. S. John and S. B. Philip. "An Intelligent Voice Assistance System for Visually Impaired Using Deep Learning." *2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 2022, pp. 449-453, doi: 10.1109/DASA54658.2022.9765171*.



12. Li, Jiawen & Xie, Lianglu & Chen, Zhe & Shi, Liang & Chen, Rongjun & Ren, Yongqi & Wang, Leijun & Lu, Xu. (2023). "An AIoT-Based Assistance System for Visually Impaired People." *Electronics*. 12. 3760. 10.3390/electronics12183760.
13. J. K. Mahendran, D. T. Barry, A. K. Nivedha and S. M. Bhandarkar. "Computer Vision-Based Assistance System for the Visually Impaired Using Mobile Edge Artificial Intelligence." 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, TN, USA, 2021, pp. 2418-2427, doi: 10.1109/CVPRW53098.2021.00274.
14. Gagandeep Singh, Omkar Kandale, Kevin Takhtani, Nandini Dadhwal. "A Smart Personal AI Assistant for Visually Impaired People." *International Research Journal of Engineering and Technology (IRJET)*, 2023.