SIGHTBOOST: EMPOWERING VISUAL ACCESSIBILITY

Gaganraj , Karthik T S , Praveen N , Hussain Afroz
Prof Dr. Kempanna M

Department of Artificial Intelligence and Machine Learning, Bangalore Institute of Technology, Bangalore

Abstract

SightBoost is a smartphone-based assistive solution aimed at enhancing the quality of life for visually impaired and blind individuals. It provides rapid and cost-effective solutions, promoting independence, ease of use, and overall usability. The technology is designed to assist blind individuals in performing daily activities, including reading product labels, exploring unfamiliar spaces, and identifying the appearance of objects in their surroundings. By utilizing the smartphone camera, SightBoost aims to break barriers and contribute to the creation of a more accessible and equitable world for individuals with visual impairments.

Keywords - *Visual Impairment, Assistive Application, Text-to-Speech, Color Recognition, Object Detection, Barcode Detection, Inclusive Design, Camera Integration, Gestures and Interaction, Real-Time Processing*

I. INTRODUCTION

- SightBoost, a cutting-edge application, is at the forefront of a technological revolution aimed at significantly improving the lives of individuals with visual impairments. This innovative solution is driven by a profound commitment to fostering inclusivity and dismantling longstanding accessibility barriers that have hindered independence. With its advanced suite of modules, SightBoost endeavors to redefine the visual experience for users, elevating their ability to navigate and engage with the world around them.

The app's commitment to inclusivity is evident in its array of groundbreaking features. Among them is Text-to-Speech Conversion, enabling seamless auditory access to written content. The provision of hands-free interaction amplifies user convenience, while Object Recognition enhances situational awareness, allowing for a more profound understanding of one's environment. Furthermore, the incorporation of color recognition adds an extra layer of functionality, contributing to a richer

SightBoost transcends conventional applications by seamlessly integrating these cutting-edge modules. This holistic approach empowers users with visual impairments, enabling them to confidently and independently navigate through various environments, challenging and overcoming traditional accessibility constraints.

Beyond being just an application, SightBoost envisions a future where visual impairments are not viewed as limitations but as gateways to newfound autonomy. The app stands as a
revolutionary step towards creating a more inclusive society,
where every individual, regardless of visual abilities, can engage with the world with confidence and independence. Join us in embracing this vision as SightBoost paves the way for a brighter and more accessible future.

II. RELATED WORKS

Sight enhancement technologies has delved into the development of applications like SightBoost, emphasizing its role in revolutionizing visual experiences for individuals with impaired vision. Recent research has explored the efficacy of SightBoost-like solutions in providing comprehensive sight enhancement, with a focus on modules like Text-to-Speech Conversion, hands-free interaction, and object recognition for improved accessibility. Studies in sight augmentation have acknowledged the significance of SightBoost's cutting-edge features, recognizing its potential to empower users and challenge traditional constraints.

[1], introduce a Voice Based Email Generating System Using Artificial Intelligence, aiding visually impaired individuals in accessing emails via a voice-based interface. The system employs AI technology, NLP concepts, and RSS to convert spoken commands to text, facilitating email composition and reading. Continuous refinement and user feedback drive improvements, but

limitations in accuracy, security, technical constraints, and integration with existing email platforms persist. Addressing these concerns is crucial for enhancing usability for visually impaired users.

[2], introduce the Blind Assistance Device using AI, employing a real-time Convolutional Neural Network (CNN) for facial expression recognition and text-to-speech conversion. The system, hosted on a Raspberry Pi with a camera, detects faces, emotions, gender, and converts text to speech. Their methodology involves designing real-time CNNs for facial and emotion detection using TensorFlow on Raspberry Pi. Limitations include potential misinterpretation of emotions, challenges for visually impaired users, text-to-voice struggles with complex formats, hardware constraints impacting real-time performance, and a complex user interface for less tech-savvy users.

[3], proposed a wearable device comprising a headset with a camera, microphone, and earphones, coupled with a portable processing device. It contains five component modules accessible via voice commands and is built using a Python interpreter. Deep learning methodologies like image captioning, object detection, and OCR form its architecture. The methodologies encompass rule-based chat-bots, image captioning with convolutional architecture, and face recognition via pre-stored databases. The image captioning module generates descriptions but might need filtering. The face recognition identifies stored individuals only, and limitations of the chat-bot module aren't discussed.

III. EXISTING DRAWBACKS

Existing literature on sight enhancement technologies has delved into the development of applications like SightBoost, emphasizing its role in revolutionizing visual experiences for individuals with impaired vision. Recent research has explored the efficacy of SightBoost-like solutions in providing comprehensive sight enhancement, with a focus on modules like Text-to-Speech Conversion, hands-free interaction, and object recognition for improved accessibility. Studies in sight augmentation have acknowledged the significance of SightBoost's cutting-edge features, recognizing its potential to empower users and challenge traditional constraints, marking a pivotal advancement in assistive technology for the visually impaired. The device was limited to specific locations, hindering its widespread use. Its obstacle detection may struggle to identify and name obstacles. Connectivity issues and limited sensory range in existing aids pose adoption challenges. Environmental factors, like traffic sounds, can disrupt obstacle detection.

The system's completely relies on its technology components like visual positioning, obstacle detection, and human-machine interface. If these components fail, the user's navigation experience could be affected. Environmental factors like change in light, complex layouts, and crowded spaces can affect the system's performance. This can impact the accuracy of obstacle detection and path planning. The major drawback of this system was accuracy and recognition issues with diverse accents and voice commands, privacy and security concerns regarding unauthorized access to sensitive information, technical limitations such as processing speed and device compatibility.

IV. PROPOSED METHODOLOGY

Firebase Natural Language for Text-to-Speech:

The integration of Firebase SDK forms the foundational step in enhancing the Android app's capabilities for visually impaired individuals. By seamlessly incorporating the Firebase SDK into the app, we unlock the powerful Natural Language Processing (NLP) capabilities offered by Firebase. Following this integration, the subsequent stage involves text processing through Firebase Natural Language. This step is crucial for processing and analyzing text inputs, enabling the identification of language nuances that contribute to more accurate and nuanced communication. Additionally, the Firebase Natural Language API is harnessed for its ability to adapt pronunciation, enhancing the clarity of synthesized speech. The TTS (Text-to-Speech) functionality is then implemented, leveraging the Firebase Natural Language API to convert processed text into high-quality, natural-sounding speech. Ensuring real-time integration is paramount to providing users with instantaneous auditory feedback as they navigate the app. This real-time feedback mechanism not only enhances user experience but also contributes to the overall accessibility and usability of the application for individuals with visual impairments.

Firebase ML Vision for Color Recognition and Object Detection:

The integration of the Firebase ML Vision SDK marks a pivotal advancement in the Android app, ushering in powerful machine learning models dedicated to image processing. By incorporating this SDK seamlessly, the app gains access to a sophisticated suite of tools and resources essential for enhancing the visual experience for visually impaired users. One key functionality enabled by this integration is color recognition, achieved through the analysis of images using the Firebase ML Vision API. This empowers the app to provide users with audible descriptions of the colors present in their immediate surroundings, thereby enriching their understanding of the environment.

In parallel, the integration extends to object detection capabilities offered by Firebase ML Vision, allowing the app to identify and describe objects captured by the smartphone camera. This comprehensive approach not only aids in recognizing individual objects but also contributes to an enhanced comprehension of the user's overall surroundings. Real-time image processing becomes a top priority in this integration, ensuring that color recognition and object detection occur seamlessly and responsively. The emphasis on real-time processing adds a layer of immediacy to the app's functionalities, aligning with the goal of creating a fluid and intuitive experience for visually impaired individuals navigating their surroundings.

Firebase Vision Barcode Detector for Barcode Detection:

The next crucial phase of the app's development involves the seamless integration of the Firebase Vision Barcode Detector API, a strategic move to enhance the app's functionality for visually impaired users. By incorporating this API, the app gains the capability for efficient and accurate barcode detection, enabling users to explore their surroundings with increased independence. Following successful barcode detection, the implementation of auditory feedback becomes instrumental.

Leveraging the Firebase APIs, the app will deliver instant auditory information about the scanned products, providing users with immediate details about the items in their vicinity. This auditory feedback mechanism is designed to be intuitive and informative, contributing significantly to the app's user-friendliness and accessibility. To further augment the app's utility, the proposed methodology underscores the importance of ensuring compatibility with a comprehensive barcode database. This integration allows the app to recognize a diverse array of barcode formats, ensuring versatility in identifying various products. The combination of efficient barcode detection, real-time auditory feedback, and integration with a comprehensive database aligns with the overarching goal of creating a robust and inclusive application for visually impaired individuals.

User-Centric Design and Testing:

Ensuring the app is truly accessible and user-friendly is integral to its success in aiding visually impaired individuals, and this is achieved through a deliberate focus on implementing robust accessibility features. The development team will concentrate on crafting an interface that not only accommodates but prioritizes accessibility. Voice commands and gestures will be intricately woven into the design, facilitating ease of navigation and interaction for users with visual impairments. By considering these features, the app aims to create an intuitive and empowering experience for its users.

To validate the effectiveness and user-friendliness of these implemented features, a rigorous usability testing phase will be conducted. This involves engaging extensively with visually impaired individuals who will provide valuable insights and feedback. The testing process aims to uncover any potential challenges, assess the intuitiveness of navigation, and gauge the overall usability of the app. User feedback will be collected systematically, capturing diverse perspectives and experiences.

The iterative development approach is foundational to the proposed methodology. It involves a continuous refinement process based on the insights derived from usability testing and user feedback. The development team will iterate on features, making continuous improvements to the overall user experience. By employing an iterative approach, the app adapts to the evolving needs and preferences of its users, ensuring that it remains a dynamic and responsive tool for visually impaired individuals.

This cyclical process of testing, gathering feedback, and refining features is essential for creating a genuinely user-centric application that aligns seamlessly with the diverse needs of its target audience.

The system revolves around a central component, the "User Interface," which serves as the primary interaction point for users. Users can trigger various actions through the user interface, including requesting object detection, color detection, text-to-speech conversion, barcode scanning, and flashlight control.

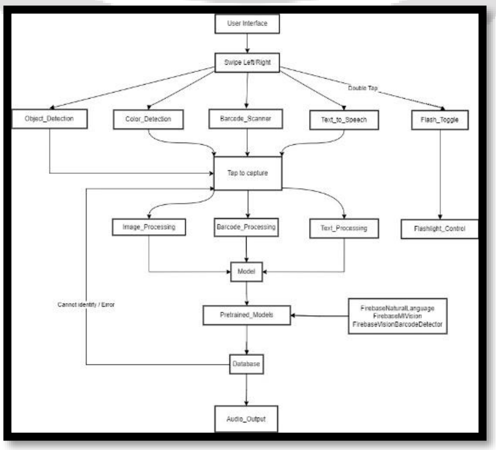

Fig 1. System Architecture of the project showcasing all features and modules

Object Detection: Upon receiving a request from the user interface, the system delegates the task of object detection to the "Object Detection" module. This module interfaces with an "Image Processing" component responsible for analyzing images and detecting objects within them. Detected objects are then passed back to the user interface for presentation to the user.

Color Detection: Similar to object detection, the "Color Detection" module handles requests from the user interface to detect colors within images. It also interacts with the "Image Processing" component to analyze images and identify colors present within them.

Text-to-Speech Conversion: When a user requests text-to-speech conversion through the user interface, the "Text to Speech" module takes charge. It converts the provided text into speech using a "Text Processing" engine and returns the synthesized speech to the user interface for playback.

Barcode Scanning: The system offers barcode scanning functionality facilitated by the "Barcode Scanner" module. This module utilizes the "Image Processing" component to process images containing barcodes, extract relevant information from them, and return the scanned barcode data to the user interface.

Flashlight Control: Users can toggle the device's flashlight on or off through the "Flash Toggle" module. This module interacts directly with a "Flashlight Control" component responsible for managing the device's flashlight functionality.

Image Processing and Preprocessing: The "Image Processing" component handles the bulk of the image analysis tasks, including object detection, color detection, and barcode scanning. It interfaces with a "Model" component, which likely houses machine learning models or algorithms necessary for these tasks. Additionally, both the "Image Processing" and "Text Processing" modules interact with a common "Model" component, indicating a shared resource for processing tasks.

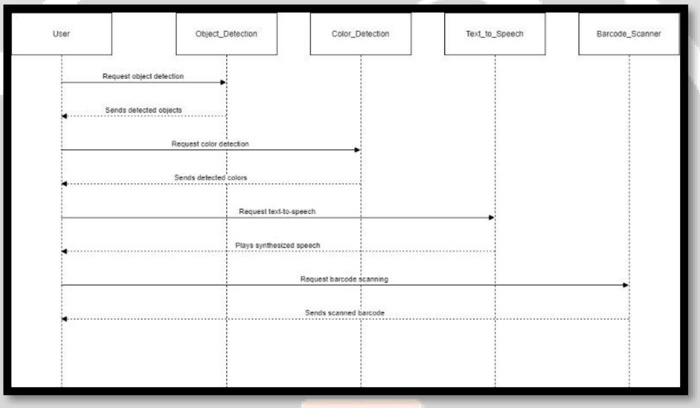

Fig 2. The provided sequence diagram outlines the interactions between different participants in a system designed to handle various tasks including object detection, color detection, text-to-speech conversion, and barcode scanning.

The sequence begins with the user initiating requests for various functionalities. Upon receiving a request for object detection, the "Object Detection" module processes the provided data, usually an image, to identify any objects within it. Following the completion of the object detection process, the identified objects are returned to the user. Similarly, when the user requests color detection, the "Color Detection" module analyzes the given data, which may include images or color samples, to determine the colors present. Subsequently, the detected colors are relayed back to the user. In the case of text-to-speech conversion, the "Text to Speech" module transforms the input text into synthesized speech, which is then played back to the user once the synthesis is complete. Finally, when the user initiates barcode scanning, the "Barcode Scanner" module processes the provided data, such as images containing barcodes, to extract barcode information. Upon completion of the scanning process, the extracted barcode data is sent back to the user.

V. LIMITATIONS AND FUTURE ENHANCEMENTS

Color recognition faces challenges in maintaining accuracy under varying lighting conditions, potentially resulting in inaccuracies when identifying colors. Similarly, object detection may encounter limitations in precision, especially within complex environments featuring numerous objects or challenging backgrounds. These challenges underline the need for continual refinement to ensure reliable performance across diverse scenarios.

Dependency on Quality of Camera: The effectiveness of object detection and barcode recognition is intricately linked to the quality and capabilities of the device's camera. Lower-end or older devices may experience diminished performance, impacting the app's ability to provide accurate and timely information. Addressing this dependency involves optimizing algorithms and functionalities to accommodate a range of camera specifications, ensuring a consistent user experience across devices.

Language Dependency in Text-to-Speech: The text-to-speech module's effectiveness may exhibit variability when extended to languages beyond the initially targeted ones. This can result in diminished clarity and naturalness in speech synthesis for users communicating in non-primary languages. Future enhancements should prioritize expanding language support to enhance inclusivity and improve the overall user experience.

Limited Barcode Compatibility: Barcode detection may encounter difficulties in recognizing less common or custom barcode formats, constraining the app's utility for specific product identification. To overcome this limitation, efforts should be directed towards creating a more extensive and adaptive barcode database, allowing for broader compatibility and increased functionality across a diverse range of products.

Real-Time Processing Resource Consumption: The continuous real-time processing required for color recognition and object detection may impose a significant strain on device resources, potentially impacting battery life and overall performance. Future optimizations should aim to strike a balance between real-time processing demands and resource efficiency, ensuring a seamless user experience without compromising device functionality.

Enhancement involves extending language support for the text-to-speech module. The objective is to cater to a broader user base by encompassing additional languages and dialects. By doing so, the Sight Boost app seeks to break down language barriers, allowing users from various linguistic backgrounds to seamlessly interact with the application.

The incorporation of deep learning models for object detection stands as a significant technological leap for the project. This enhancement aims to improve the accuracy of object detection, enabling the app to identify a wider range of objects with increased precision. By leveraging deep learning, the Sight Boost app strives to enhance its overall effectiveness in providing detailed information about the user's environment. A forward-looking enhancement involves the investigation of augmented reality (AR) integration.

This feature aims to enrich the user experience by providing additional context and information about the surroundings through AR features. This not only enhances the app's usability but also contributes to a more immersive and informative experience for visually impaired users. Additionally, the introduction of gender detection represents a novel and progressive feature.

By identifying genders, the Sight Boost app can offer users valuable information about the individuals around them, fostering a greater understanding of their social environment. This innovative enhancement aligns with the project's overarching goal of promoting independence, accessibility, and inclusivity for individuals with visual impairments.

VI. RESULTS AND DISCUSSION

This section presents some results showing the efficiency of the Sightboost application in Object Detection, Color Detection, Barcode detection, Text to Speech (text recognition)

Object detection: In Figure 3 the object recognition feature, demonstrates a scenario in which the app captures an image depicting a laptop's keyboard, representing a common object encountered in daily life. The SightBoost app swiftly processes the visual input, analyzing the intricate details and structure of the keyboard with precision. Once the image is processed, the application provides valuable auditory feedback through spoken descriptions.

Fig 3. Object Detection

Color Detection: Figure 4, represents a pivotal feature of the SightBoost application, enhancing accessibility for visually impaired users. Upon capturing an image containing the target color, the application swiftly processes the visual data, and recognizes the targeted color in the image. The results are then seamlessly conveyed to the user through an intuitive audio output format.

Fig 4. Color Detection

Barcode detection: In Figure 5, the SightBoost application showcases its capability to recognize barcodes captured within photographs, when a barcode is captured in a photo using the SightBoost app, the application swiftly processes the image data and decipher the encoded information contained within the barcode. SightBoost accurately extracts the numerical data embedded within the barcode, enabling users to access critical information such as product identification numbers, serial codes.

Fig 5. Barcode detection

Text to Speech: In Figure 6, the SightBoost application provides a Text-to-Speech feature, offering a transformative solution for visually impaired individuals to access textual information present within images. When an image containing text is captured using the SightBoost app, the application employs sophisticated Optical Character Recognition (OCR) technology to recognize and extract the textual content embedded within the image.

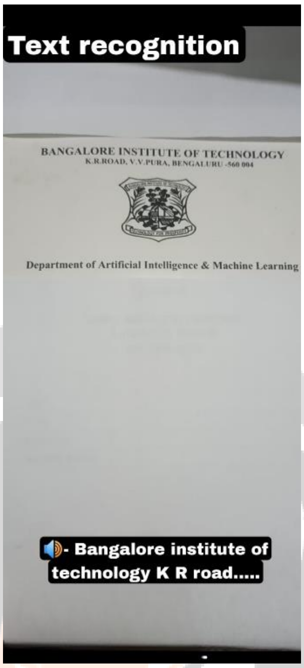

Fig 6. Text to Speech

Through precise analysis and interpretation of the visual data, SightBoost accurately identifies individual characters then seamlessly synthesized it into speech output using a Text-to-Speech (TTS) API integrated within the SightBoost application. Then transforms the recognized text into clear and intelligible audio output.

VII. CONCLUSION

In the development journey of the "SIGHTBOOST" application, our primary focus is on integrating indispensable features such as object recognition, text-to-speech, color recognition, and barcode scanning. We understand the profound impact these functionalities can have on the daily lives of visually impaired individuals, empowering them with increased independence and accessibility.

Our commitment to creating a user-centric application goes beyond the inclusion of essential features. We recognize the importance of cross-platform compatibility as a cornerstone of our development strategy. This approach ensures that "SIGHTBOOST" can be seamlessly accessed and utilized across a diverse range of devices and operating systems. By embracing cross-platform compatibility, we aim to cater to the varied technological preferences of our users, fostering inclusivity and breaking down barriers that may limit accessibility.

Moreover, our dedication to user-centric design extends to ensuring an intuitive and streamlined user experience. The integration of object recognition allows users to effortlessly identify and understand their surroundings, while text-to-speech functionality transforms written information into clear and audible speech. The addition of color recognition and barcode scanning further enhances the app's capabilities, offering users a comprehensive and versatile tool for navigating the world around them. As we move forward with the development of "SIGHTBOOST," our vision is rooted in creating more than just an application; we strive to build a supportive and inclusive community. We actively seek user feedback to continually refine and enhance the app, ensuring that it evolves in tandem with the needs and preferences of the visually impaired community. By embracing cross-platform compatibility and a user-centric approach, "SIGHTBOOST" aims to be a beacon of accessibility, fostering independence and empowerment for individuals with visual impairments.

VIII. REFERENCES

[1] *Rijwan Khan- "Voice Based Email Generating System Using Artificial Intelligence", International Conference on Artificial Intelligence and Smart Communication – 2023.*

[2] *Safnaz – "Blind Assistance Device using AI", International Journal of Engineering Research & Technology - 2021.*

[3] *Geetanjali Kale – "A Comprehensive Assistance System for Visually Impaired Individuals", Institute of Electrical and Electronics Engineers - 2021.*

[4] *Alejandro Boza-Chua – "Inclusive Education: Implementation of a Mobile Application for Blind Students", International Journal of Advanced Computer Science and Applications - 2021.*

[5] *Akshata Sangwai – "Eye(I) Still Know! – An App for the Blind Built using Web and AI", International Journal of Engineering and Management Research - 2021.*

[6] *Shrikesh Suresh - "Vision - Android Application for the Visually Impaired", IEEE International Conference for Innovation in Technology – 2020.*

[7] *Rishi Sayal – "Mobile App Accessibility for Visually Impaired", International Journal of Advanced Trends in Computer Science and Engineering - 2020.*

[8] *Devashish Pradeep Khairnar – "PARTHA: A Visually Impaired Assistance System", 3rd International Conference on Communication System, Computing and IT Applications – 2020.*

[9] *Ranjeet Patil – "Designing Mobile Application for visually impaired and blind persons", Institute of Electrical and Electronics Engineers - 2020.*

[10] *Joe Louis Paul – "Smart Eye for Visually Impaired-An aid to help the blind people", Second International Conference on Computational Intelligence in Data Science - 2019.*

[11] *Bing Li – "Vision-based Mobile Indoor Assistive Navigation Aid for Blind People", Institute of Electrical and Electronics Engineers - 2019*

[12] *Joe Yuan Mambu – "Blind Reader: An Object Identification Mobilebased Application for the Blind using Augmented Reality Detection", 1st International Conference on Cybernetics and Intelligent System – 2019.*

[13] *Sumitra A. Jakhete – "Object Recognition App for Visually Impaired", IEEE Pune Section International Conference (PuneCon) – 2019.*

[14] *Milios Awad - "Intelligent Eye: A Mobile Application for Assisting Blind People", IEEE Middle East and North Africa Communications Conference – 2018.*

[15] *Jawaid Nasreen – "Object Detection and Narrator for Visually Impaired People", 6 th IEEE International Conference on Engineering Technologies and Applied Sciences – 2019.*