iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.61273

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 12 Issue IV Apr 2024- Available at www.ijraset.com

Empowering the Blind: An AI Driven Indoor Assistance for Visually Impaired

Asst. Prof. Mrs. Shraddha. P. Mankar[1], Riya Gawande[2], Dipanshu Sankhala[3], Siddhi Vispute[4], Nikhita Watpal[5]
Information Technology Savitaribai Phule Pune University

Abstract: ***This review paper presents the design and development of a comprehensive mobile application aimed at enhancing the daily lives of visually impaired individuals. The proposed mobile app offers real-time assistance for various visual recognition tasks, including object detection, distance estimation, currency recognition, barcode detection, color recognition, and emotion analysis. The primary objective of this application is to provide visually impaired users with a powerful tool to navigate their surroundings, identify objects and currency, scan barcodes, discern colors, and even gauge the emotions of those they interact with. The app integrates pre-trained deep learning models for object detection and facial emotion analysis, computer vision techniques for distance estimation, and optical character recognition for currency recognition. User feedback and engagement are central to the ongoing improvement of the application, ensuring that it remains a valuable resource for the community it serves. Ethical considerations and privacy concerns are also addressed, with a focus on data security and user privacy. By presenting the development and functionalities of this app, this review paper not only contributes to the field of assistive technology but also underscores the importance of harnessing cutting-edge technology to improve the quality of life for visually impaired individuals.***

Keywords: Visually impaired, object detection, distance estimation, emotion recognition, audio output

I. INTRODUCTION

In a world where millions contend with visual impairments, comprehending and navigating indoor environments presents a unique set of challenges. The visually impaired often encounter difficulties moving autonomously within enclosed spaces, where obstacles and objects can vary widely in position and type. Traditional aids, such as white canes, though helpful in some instances, do not provide real-time information about indoor objects and surroundings.

Recognizing the specific challenges faced by visually impaired individuals in indoor settings, this project aims to introduce a pioneering Android application tailored to enhance their independence and confidence. Our application is designed to assist with object detection, distance estimation, currency recognition, barcode scanning, color identification, and emotion analysis, with a Leveraging the device's built-in camera, the application employs a combination of deep learning models, including object detection using You Only Look Once (YOLO) V3 and Single Shot Detector (SSD), distance estimation with the Mono-depth algorithm, currency recognition based on OCR, barcode detection, color identification, and emotion analysis. It detects, describes, and recognizes indoor objects, provides real-time information about object distances, and delivers comprehensive object-related details through audio output, which can be seamlessly channeled through headphones or the device's speaker.

An essential characteristic of this system is its suitability for indoor environments, eliminating the need for external cameras or sensors. By emphasizing indoor functionality, this project addresses the unique challenges faced by the visually impaired within enclosed spaces, enhancing their ability to navigate and interact with their surroundings. The central aim of this project is to showcase how advanced computer vision techniques can significantly improve the quality of life for visually impaired individuals during indoor activities. We will present a comprehensive overview of our application, detailing its various modules, functionalities, and its potential to empower and support the visually impaired in indoor settings.

The project implemented an Android-based application designed to solve the problems faced by visually impaired people while navigating indoor environments. Traditional tools such as the white cane are limited in providing real information about indoor products and problems. The app uses the Android device's built-in camera and computer vision technology to fill this gap. The app facilitates interaction with visually impaired users by providing real-time information about indoor equipment, distance and other relevant details through voice output. More importantly, it is designed for an indoor environment and should not have external cameras or sensors. The project aims to increase the independence and confidence of blind people in prisons by focusing on in-house work.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

The main goal of the project is to show that the use of computer vision technology can improve the quality of life of visually impaired people while navigating at home. From an overview of the app's structure and features, the project focuses on the app's ability to improve and support visually impaired people manage the indoor environment.

II. MOTIVATION

The motivation behind this project, aimed at assisting visually impaired individuals with their indoor activities, is multifaceted and deeply rooted in a commitment to making a positive impact on the lives of those with visual impairments. Several key factors drive the project's motivation. First and foremost, the project is motivated by the goal of empowering independence among visually impaired individuals. Vision loss often poses significant challenges to performing everyday tasks, which can lead to a greater reliance on assistance from others. This project seeks to change that by providing technological solutions that enhance self-sufficiency.

By offering tools and features tailored to the unique needs of the visually impaired, it aims to foster greater autonomy in daily life. this system. Raspberry pie has to be powered constantly with the help of a power bank which entails the user to carry a power bank along with him/her all the time. Moreover, this system tells the user the name of the detected objects.

The motivation behind the program stems from a strong commitment to promoting the participation, independence and empowerment of visually impaired people. Despite advances in technology, navigating the home environment is still a challenge for visually impaired people. Traditional service tools, while useful in some cases, often do not provide real-time information about indoor equipment and issues.

By creating an Android application suitable for indoor use, we aim to close this important gap and offer practical solutions that will improve the lives of visually impaired people. We are motivated by our belief that all people, regardless of their abilities, should be equal and be able to manage their environment safely and freely.

In addition, our activities are directed to create solutions to new problems that are beneficial to people's lives by using the power of today's technologies, especially computer vision and deep learning. Using the capabilities of smartphones and image processing equipment, we strive to create efficient and effective tools that will increase the independence and quality of life of visually impaired people. Finally, our passion is rooted in our belief in the power of technology to break down barriers and create greater unity.

Through this project, we hope to help visually impaired people move around in a more comfortable, safe and dignified environment, allowing them to participate and contribute to their communities.

III. RELATED WORK

The study on the Android application mentioned above for indoor navigation and object search for the visually impaired consists of research and technological studies aimed at solving a similar problem. Some of the main areas of work are:

1) *Technological services for the visually impaired:* Various services have been developed to assist the visually impaired with navigation and navigation. This includes devices such as wearable cameras, smart sticks, and navigation apps. Research in this area focuses on improving the accuracy, usability, and effectiveness of technology to meet user needs.
2) *Computer Vision for Object Detection:* Extensive research has been done on computer vision for object detection, including deep learning-based methods such as YOLO (You Only Look at One Leg) and SSD (Single Shot Detector). This technology, which is used in many areas including object recognition, can provide instant detection of objects in complex environments.
3) *Distance and Depth Estimation:* Examining distance estimation and depth perception using relevant computer vision techniques to provide instant information about the distances of objects in the indoor environment. Depth algorithms and stereo vision are often used to estimate the depth of images captured by cameras.
4) *Accessibility and Inclusive Design:* Accessibility and inclusive design studies focus on the use of technology to provide easy access to users with disabilities, including the visually impaired. This includes research on designing user interfaces, developing assistive technologies, and ensuring digital content is accessible to all users.

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

IV. PROPOSED WORK

A. *Architecture:*

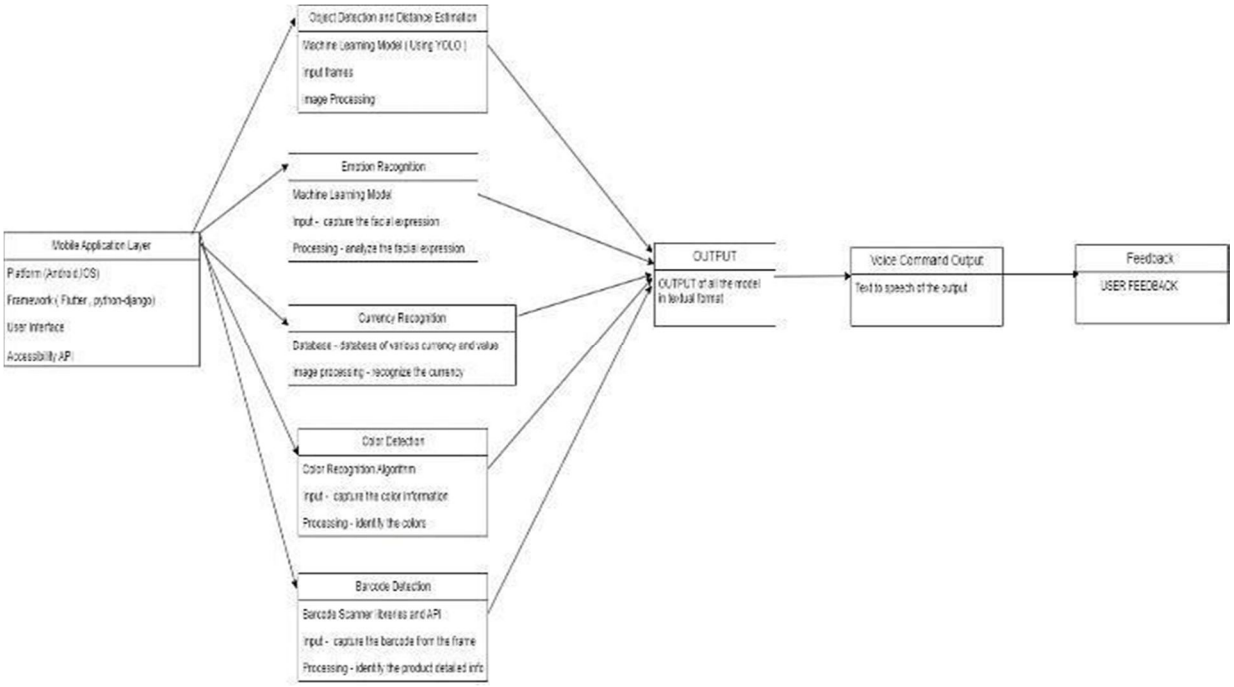


Fig. 1. System's architecture

B. *Activity diagram*

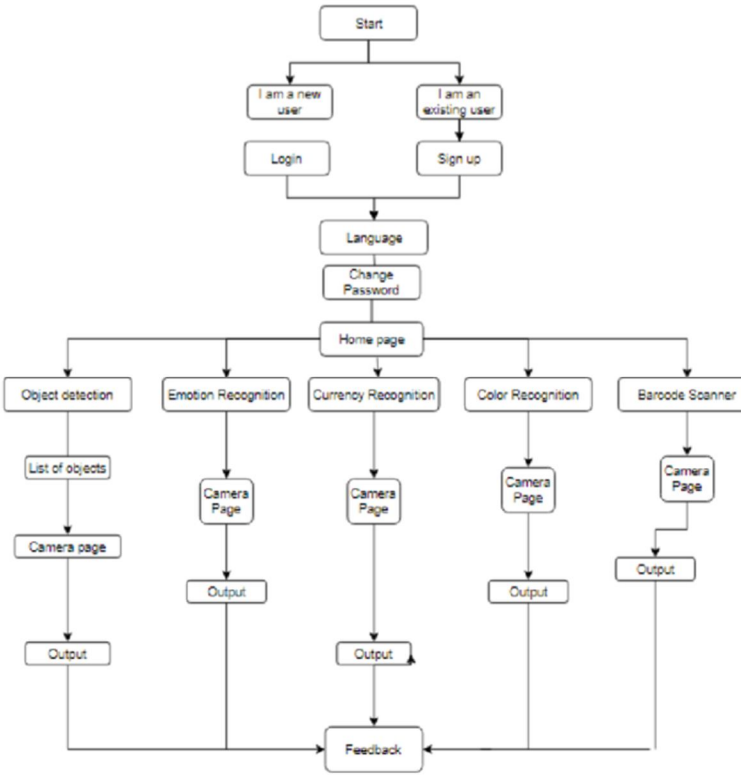


Fig. 2. System's activity diagram

©IJRASET: All Rights are Reserved

1) Object detection

i. Object detection is the principal objective of this system. It encompasses two key elements: object classification and object localization. Object detection entails categorizing objects into predefined classes. In essence, object classification assigns a label to the object within the user's surroundings, thereby providing real-time information about the object's identity.

ii. Object detection is an important task in computer vision that involves identifying and locating objects of interest in images or videos. With the rapid development of deep learning and convolutional neural networks (CNN), object detection has made great progress in recent years and has been widely used in many fields such as driving, surveillance, robotics, healthcare, and virtual reality.

iii. Object positioning: The first step in object detection is to determine the position of the object in the input image. This involves identifying regions of interest (ROIs) that may contain artifacts. The most common placement methods include sliding window methods, regional bidding methods (such as search engines, EdgeBoxes), or new technologies such as Regional Bidding Networks (RPN).



Fig. 3. Object Detection

2) Distance calculation

i. The distance of the object from the visually impaired person will be provided. The OpenCV (Open source Computer Vision Library) will be utilized to compute the distance. The Triangle Similarity Law will be utilized to compute the distance.

ii. Knowing the object's distance from the visually impaired user would be more advantageous than simply knowing the object's name since it will help the user get a sense of the relative space surrounding him. The calibration module must be run each time the program is started for the first time since the device's camera needs to be calibrated before it can calculate the distance.

iii. Finding the camera lens's focal length is the primary goal of calibration since it will be required for other calculations. Finding the camera lens's focal length is the primary goal of calibration since it will be required for other calculations.



Fig. 4. Output for distance estimation

3) Currency Recognition

i. To implement currency recognition in the app, collect a diverse dataset of currency images, preprocess them, and extract relevant features. Train a machine learning model, like a CNN, to classify denominations. Integrate the model into the app for real-time recognition using the device's camera. Ensure accessibility with text-to-speech or audio feedback for the visually impaired. Test rigorously, make the system adaptable to various currencies, and gather user feedback for ongoing improvements, creating an inclusive tool.

ii. Image collection: The first step is to collect images of the results to be seen. This can be done using a camera, scanner or other imaging device. In practical applications such as mobile banking applications, images are often captured using the device's camera.

iii. Pre-Processing: Pre-processing is done to improve the quality of the images obtained and to prepare them for feature extraction. Preprocessing techniques include transforming, noise removal, contrast enhancement, and normalization.

iv. Integration with applications: Integrate the value recognition model into the target application, such as a mobile application or accounting system. The application uses an authentication model to instantly verify and distribute currency values.



Fig 5. Currency Recognition

i.

4) Emotion Recognition

ii. To implement emotion recognition in the app for visually impaired individuals, a multifaceted approach is essential. This involves several key steps. First, gather a diverse dataset of facial images displaying a range of emotions, ensuring representation of happiness, sadness, anger, and other emotional states, and including images of different individuals for variety.

iii. Preprocess these images by resizing, normalizing lighting conditions, and enhancing contrast to optimize their suitability for analysis. Next, extract relevant features from these images, which could include facial landmarks, texture, and color patterns associated with different emotions. The core of the system involves training a machine learning model, commonly a Convolutional Neural Network (CNN), specifically designed for multi-class emotion classification.

iv. Image capture: The first step is to capture a photo or video frame containing a human face. This can be done using a camera, webcam or other imaging device. For real-time applications such as emotion recognition or video analysis, images are often captured continuously from real-time streams.

v. Face Detection: Perform face detection and extract faces from input images or photos. video post. This step involves identifying regions of interest (ROIs) that may contain human faces. face detection algorithms such as Viola-Jones, Histogram of Directed Gradins (HOG), and deep learning such as convolutional neural network (CNN).

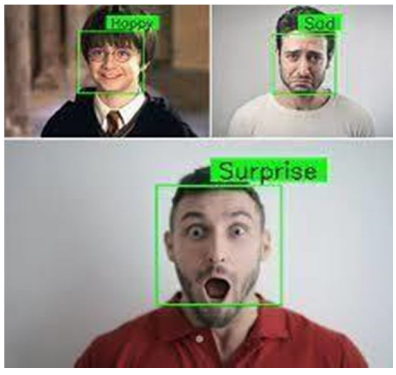


Fig 6. Emotion Recognition

5) Color Recognition

i. Implementing color recognition in the app for visually impaired users involves a structured approach. First, gather an extensive dataset of color images representing a wide range of colors and variations. Preprocess these images by standardizing lighting conditions, resizing, and enhancing contrast to prepare them for analysis.

ii. Image capture: The first step is to capture a photo or video frame with the color of the object or scene to be known. This can be done using a camera, scanner or other imaging device. For real-time applications, images are typically captured continuously from real-time streams.

iii. Preprocessing: Preprocessing of the input image to improve its quality and prepare it for color recognition. This will include things like resizing, removing noise, increasing contrast, and changing color settings. Preprocessing helps standardize input data and improve the efficiency of subsequent steps.

iv. Color space conversion: Convert the input image from the RGB color space to a color space suitable for color analysis. Color systems commonly used in color recognition include RGB, HSV (hue, saturation, value), LAB (CIELAB), and YUV. Each color space has its own advantages and can be chosen according to the specific needs of the application.

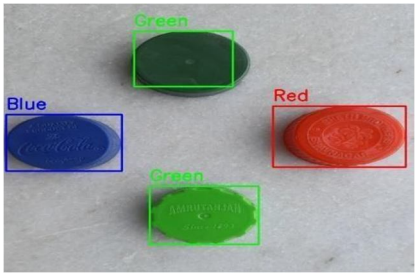


Fig 7. Color Recognition

6) Barcode detection

i. Implementing barcode detection in the app for visually impaired users involves several key steps. Begin by integrating a barcode scanning library or API that's capable of recognizing and decoding barcodes. These libraries often come with pre- trained models for barcode recognition.

ii. Leverage the device's camera to capture real-time images of barcodes. When a barcode is detected, the app can overlay feedback on the screen, such as the decoded information or a description of the product or item associated with the barcode. For accessibility, the app should provide audio feedback through text-to-speech, conveying the barcode's content to users.

iii. Image capture: The first step is to capture the image or video frame containing the barcode to be scanned. This can be done using a camera, scanner or other imaging device. For real-time applications, images are typically captured continuously from a real-time video stream.

iv. Pre-Processing: Pre-processing the input image to improve its quality and prepare it for barcode scanning. This will include functions such as conversion, noise removal, contrast enhancement and edge detection. Preprocessing helps standardize input data and improve the efficiency of subsequent steps.

v. Edge Detection: There is often a difference between black and white lines in barcodes. First use an edge detection algorithm (such as Sobel, Canny, or Gaussian Laplace) to detect edges in the image. Edges represent the transition between dark and light in the barcode.

vi. information for a variety of purposes, such as identifying products, finding items for sale, or accessing online resources.



Fig 8. Barcode Detection

V. ALGORITHM

A. Object Detection:

1) YOLOv4, or "You Only Look Once version 4," is a state-of-the-art object detection algorithm that belongs to the family of single-stage detectors.
2) It was introduced as an evolution of previous versions such as YOLOv3. YOLO algorithms are known for their speed and efficiency in real-time object detection tasks.
3) YOLOv4 is a one-stage object detection algorithm, meaning it directly predicts bounding boxes and class probabilities from input images in a single pass through the network.
4) This makes YOLO models faster compared to two-stage detectors like Faster R-CNN.
5) YOLOv4 typically uses a powerful backbone network such as CSPDarknet53 or YOLOv4-tiny, which extracts features from the input image.

B. Distance Estimation:

1) The distance estimation of the object in frame is calculated in inches.
2) To make the object distance estimation easier and accessible to the user, we use an algorithm to divide the inches into an average foot length so that the user can take that many amount of steps to reach that object.
3) The distance in inches is not easily measurable by the user, therefore, this distance in inches is converted to the number of steps the user has to take to reach that object.

VI. RESULTS

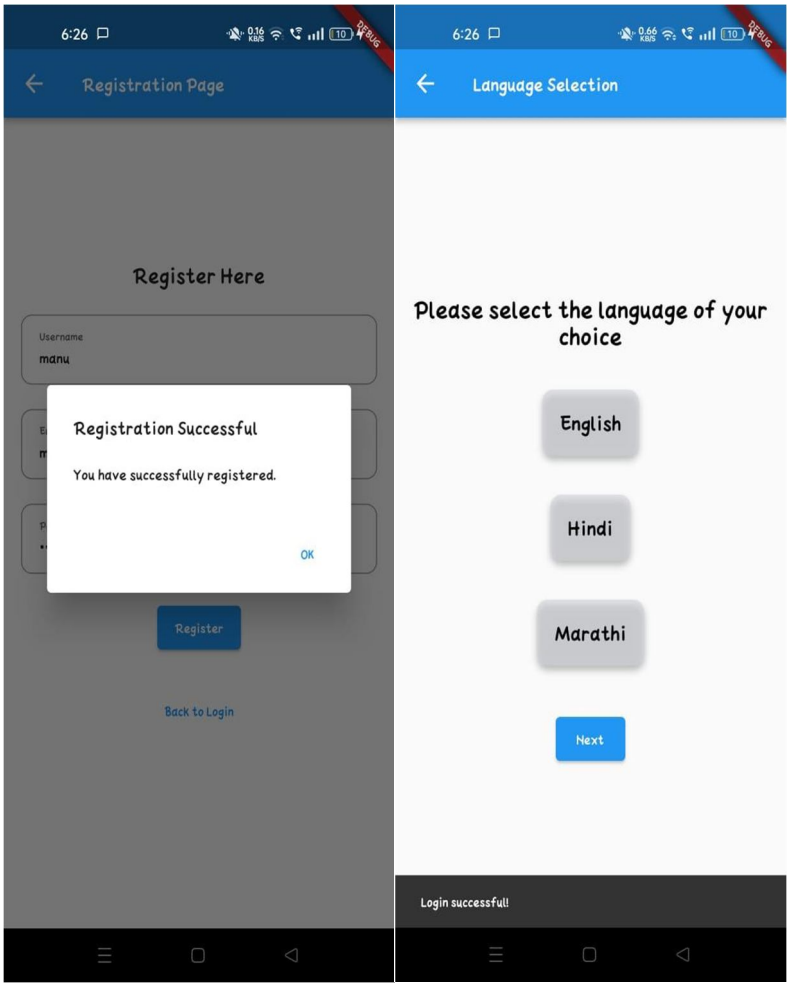


Fig 9. Login Page Fig 10. Language Selection Page

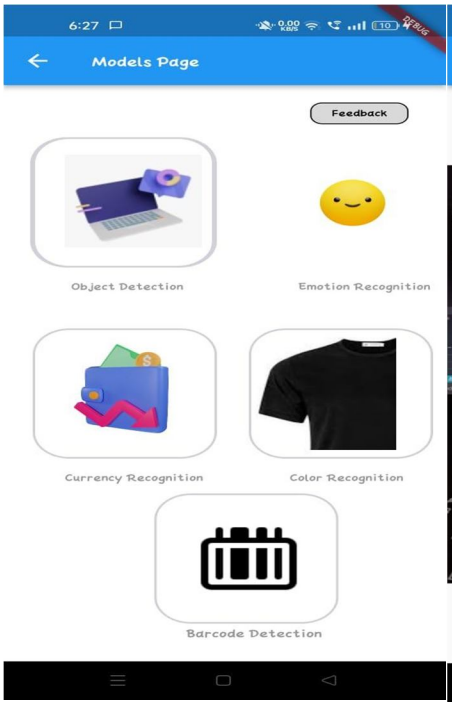


Fig 11. Main Page

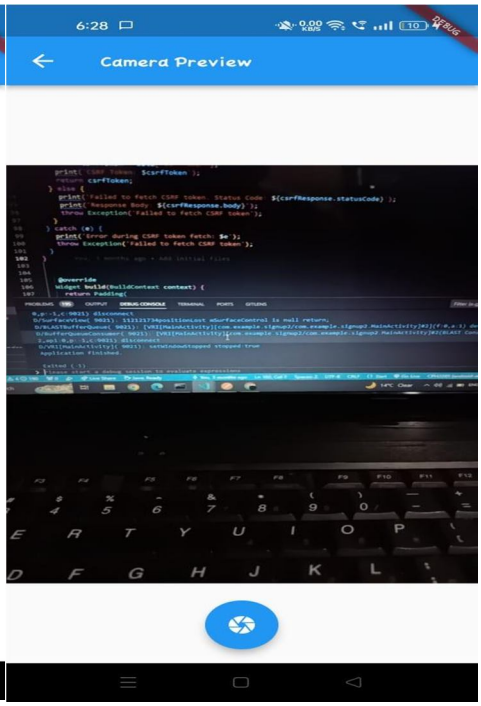


Fig 12. Camera Preview Page

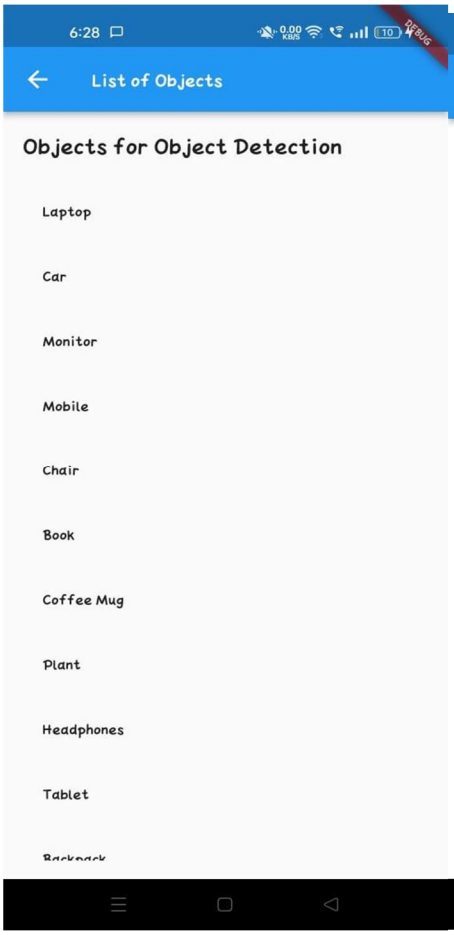


Fig 13. List of Objects



Fig 14. Object Detection Option

VII. FUTURE SCOPE

The future scope for an app designed for visually impaired people, featuring modules such as object detection, emotion recognition, color recognition, currency detection, and barcode detection, holds significant promise. With ongoing advancements in machine learning and AI, these modules can become more accurate and efficient, greatly enhancing their utility for the visually impaired. Furthermore, integration with wearable devices like smart glasses can provide a hands-free experience, and expanding the app's capabilities to include navigation and wayfinding would be invaluable. Language and OCR integration, a supportive user community, and educational modules can further enhance the app's functionality. Maintaining a user friendly interface, ensuring data privacy, and adhering to global accessibility standards are key considerations, as well as collaborating with relevant organizations for wider adoption and support. It is important to continuously improve the accuracy and efficiency of target detection algorithms. This will include improving the fundamental structure of machine learning, optimizing real-time processing algorithms, and reducing errors and omissions.

Integration of advanced sensors such as LiDAR (light detection and ranging) or radar with cameras can provide more in-depth information, improving the ability to accurately identify and separate objects, especially in a challenging or challenging environment. weather. In addition to simple object detection, future iterations of the system may include semantic understanding and context recognition. This means not only identifying objects, but also understanding their relationships, functions and effects in different situations.

Integrating object detection systems with navigation tools will provide greater assistance to the visually impaired. This may include using GPS and map data to generate audio or feedback to guide users around obstacles or to specific locations. Designing features that provide immediate feedback or descriptions of the environment can improve the user experience. This can include not only the identification of objects but also elements such as lighting, textures and spatial layout. Providing the user with a customizable interface based on the user's preferences and needs is crucial to ensuring interaction.

Future work will include creating a better information management system, support for multiple languages, and better compatibility with assistive devices.

Integration of object detection with wearable devices such as smart glasses or haptic vests can provide users with a more interactive and intuitive experience. This will enable hands-free working and promote environmental awareness. Optimizing object detection algorithms used at the edge can increase responsiveness and reduce dependence on cloud services. This includes the construction of heavy structures and efficient pipelines suitable for low-energy products with low-cost components.

It is crucial to collect continuous feedback from visually impaired users and incorporate their input into the design process. This ensures that the system remains user-friendly and addresses the needs and challenges faced by its users. Partnering with organizations and communities dedicated to accessibility and disability rights can help ensure that diagnostic products are designed and used in ways that respect and promote the freedom and independence of the visually impaired.

VIII. CONCLUSION

In conclusion, the app designed for visually impaired individuals, featuring modules such as object detection, emotion recognition, color recognition, currency detection, and barcode detection, represents a beacon of hope for a more inclusive and accessible future. With advancements in technology and AI, the app's potential for improving the quality of life for the visually impaired is substantial. Its evolution may lead to greater independence, better social interaction, and enhanced everyday experiences. As the app continues to grow and adapt, it has the power to break down barriers and provide new found opportunities, ensuring that the visually impaired community can navigate the world with confidence and autonomy. This app embodies the principles of innovation, compassion, and inclusivity, offering a brighter future for those who rely on its support.

REFERENCES

[1] Ruffieux S, Hwang C, Junod V, Caldara R, Lalanne D and Ruffieux N. (2023). Tailoring assistive smart glasses a according to pathologies of visually impaired individuals: an exploratory investigation on social needs and difficulties experienced by visually impaired individuals. Online publication date: 1-Jun-2023.

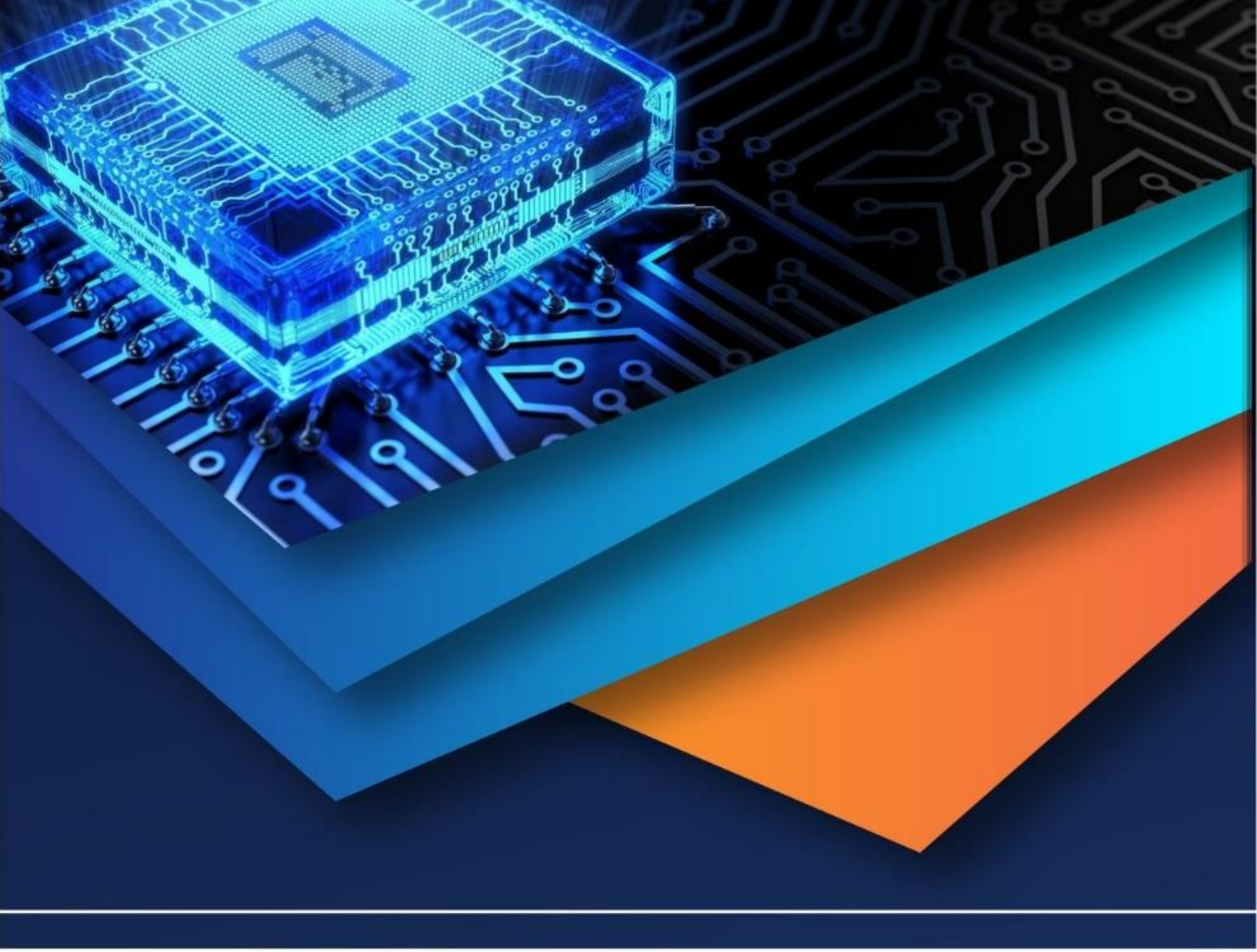
[2] Jafri, Rabia & Ali, Syed & Arabnia, Hamid & Fatima, Shameem. (2013). Computer vision-based object recognition for the visually impaired in an indoors environment: a survey.

[3] Murad, M., Rehman, A., Shah, A.A., Ullah, S., Fahad, M., Yahya, K.M.: RFAIDE—an RFID based navigation and object recognition assistant for visually impaired people. In: 7th Inter- national Conference

[4] CICERONE- A Real Time Object Detection for Visually Impaired PeopleTherese Yamuna Mahesh[1], S S Parvathy[2], Shibin Thomas[2].

[5] Ruxandra Tapu, Bogdan Mocanu, Andrei Bursuc, Titus Zaharia; Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops, 2013, pp. 444-451
[6] Shin, Byeong-Seok, and Cheol-Su Lim. "Obstacle detection and avoidance system for visually impaired people." In International Workshop on Haptic and Audio Interaction Design, pp. 78-85. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
[7] Rodríguez, Alberto, J. Javier Yebes, Pablo F. Alcantarilla, Luis M. Bergasa, Javier Almazán, and Andrés Cela. "Assisting the visually impaired: obstacle detection and warning system by acoustic feedback." Sensors 12, no. 12 (2012): 17476-17496.
[8] Zhang, Qian, and Wei Qi Yan. "Currency detection and recognition based on deep learning." In 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-6. IEEE, 2018.
[9] Ajmal, Aisha, Christopher Hollitt, Marcus Frean, and Harith Al-Sahaf. "A comparison of RGB and HSV colour spaces for visual attention models." In 2018 International conference on image and vision computing New Zealand (IVCNZ), pp. 1-6. IEEE, 2018.
[10] Hansen, Daniel Kold, Kamal Nasrollahi, Christoffer Bøgelund Rasmussen, and Thomas B. Moeslund. "Real-time barcode detection and classification using deep learning." In International Joint Conference on Computational Intelligence, pp. 321-327. SCITEPRESS Digital Library, 2017.
[11] Kosti, Ronak, Jose M. Alvarez, Adria Recasens, and Agata Lapedriza. "Emotion recognition in context." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1667-1675. 2017.
[12] Natanael, Gafencu, Cristian Zet, and Cristian Foşalău. "Estimating the distance to an object based on image processing." In 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), pp. 0211-0216. IEEE, 2018.

©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)