International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

"Third Eye an Android Application for Visually Impaired People"

Atharv Devhare[1], Rushikesh Bhaskar[2], Atharva Sonawane[3] ,Ahhad Nadaf[4,] Prof. M. D. Sanap[5]

[1,2,3,4,5] *Computer Engineering & LoGMIEER.Nashik*

Abstract - The "3rd Eye" Android application is an innovative solution designed to empower visually impaired individuals by providing access to printed information using advanced technologies. The app utilizes document scanning, text-to-speech conversion, and object detection functionalities, all controlled through voice commands. Its primary focus is to cater to the specific needs of visually impaired users, enabling independent navigation and interaction with the world. By leveraging the device's camera, the app effortlessly scans documents, enhances their quality through image processing, and accurately extracts text using OCR algorithms. Additionally, it employs machine learning and computer vision to detect objects in real-time, providing audio descriptions and enhancing spatial awareness. The app's voice command feature allows seamless interaction, responding to user commands through natural language processing. Rigorous testing and user feedback ensure performance optimization and reliability. The 3rd Eye app empowers visually impaired individuals by providing document scanning, access to printed information, and improved environmental comprehension, ultimately enhancing their independence and daily lives.

Key Words: Mobile Phone, Optical Character Recognition, Text to Speech, Object Detection, TensorFlow Lite

1.INTRODUCTION

The visually impaired community faces significant challenges in navigating their surroundings and accessing information independently. Traditional methods of accessing text-based information, such as Braille, are time-consuming and require specialized training. Additionally, identifying objects in the environment can be challenging and may lead to safety concerns. The rapid advancement of smartphone technology and the widespread availability of Android devices provide a unique opportunity to develop assistive applications that leverage computer vision and voice recognition technologies. In response to these challenges, our research focuses on the development of 3rd Eye, an innovative Android application that empowers visually impaired individuals by providing them with a powerful tool for accessing and interacting with their environment. By harnessing the capabilities of OCR technology, the application enables users to effortlessly scan printed text and images and convert them into speech output, thus facilitating easy access to information. Moreover, the application's computer vision algorithms detect and identify objects in real-time, providing users with audio feedback about the location and nature of objects in their vicinity.

The primary objective of this research paper is to present the technical details of the 3rd Eye application, including its architecture, functionalities, and underlying algorithms. We aim to highlight the robustness and effectiveness of the application in enhancing the independence and quality of life for visually impaired individuals. By conducting comprehensive usability tests and performance evaluations, we assess the application's accuracy, reliability, and user-friendliness. Through the development and evaluation of 3rd Eye, we anticipate that visually impaired individuals will gain increased independence, improved access to information, and enhanced navigation capabilities. We believe that this research contributes to the ongoing efforts in developing inclusive technologies that enable individuals with visual impairments to thrive in an increasingly digital and visually oriented world.

2. Literature Survey

1. In [1] The research paper highlights the need for visually impaired individuals to receive proper training in order to use the Third Eye prototype effectively. They must learn how to understand and interpret the signals from the sensors, recognizing them as obstacles or water bodies. With practice, users can become better at using the technology and gain maximum benefit from it. Although the prototype shows promise, more research and testing are necessary to improve. its design and make it more accurate and reliable in real-world situations. Feedback from users will be valuable in identifying any limitations and making the Third Eye more functional for visually challenged individuals. Continuous improvements will be made based on this feedback.

2. In [2] This IJRASET paper introduces "Blindfold, " a smartphone-based object detection application aimed at supporting visually impaired individuals. The application harnesses the power of smartphones to detect objects in real-time and provides audio or haptic feedback to the user. By utilizing the camera and image processing capabilities of smartphones, Blindfold analyzes the live video feed and identifies objects within the frame. The application then translates this information into auditory or tactile cues, allowing the user to perceive and interact with their surroundings more effectively. The paper emphasizes the user-friendly nature of Blindfold, as it leverages the widespread availability and familiarity of smartphones among visually impaired individuals. The application provides an accessible and cost-effective solution for object detection without the need for additional specialized hardware.

3. Architectural Design

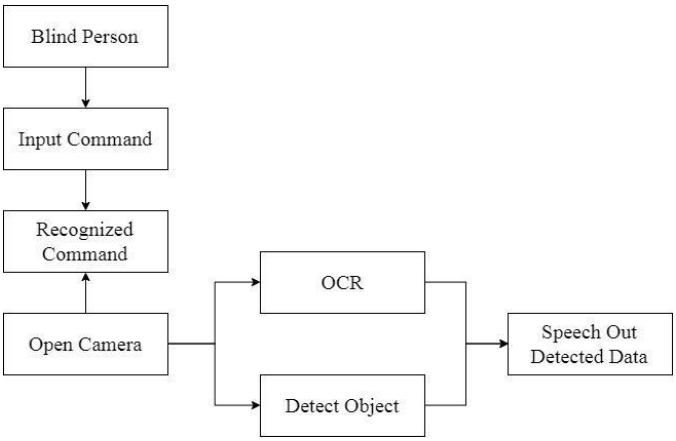

Fig 3.1 Architecture Diagram

This application will be available to the users in the form of an Android application. Modules involved in this Reading System are as follows:

1. **Camera:**
   The built-in camera application on smartphones is used to identify objects present in the frames captured by the smartphone's rear camera. Additionally, it has the capability to recognize and extract text from the frames captured by the smartphone's rear camera.

2. **TensorFlow:**
   An open-source machine learning framework used for implementing the object detection functionality in the application. TensorFlow provides a robust ecosystem for training and deploying machine learning models.

3. **Google Vision API:**
   A cloud-based service provided by Google for optical character recognition (OCR) capabilities. The Google Vision API is utilized to extract text from scanned documents and images.

4. **Speech-to-Text:**
   A technology or service that converts spoken language into written text. The application utilizes a suitable speech-to-text API to process voice commands and convert them into actionable instructions.

4. Methodology

1. Object Detection Algorithm:
The application utilizes a deep learning-based object detection algorithm, such as the Single Shot MultiBox Detector (SSD) or You Only Look Once (YOLO). These algorithms are trained on large datasets and can detect various objects in real-time using the device's camera. They provide bounding box coordinates and class labels for recognized objects.

2. Optical Character Recognition (OCR) Algorithm:
The project incorporates an OCR algorithm to extract text from scanned documents and images. This algorithm applies image processing techniques to recognize and interpret characters within the input image. It leverages machine learning models and pattern recognition algorithms to accurately identify and extract text from different languages and fonts.

3. Speech Recognition Algorithm:
The application employs a speech recognition algorithm to convert spoken commands into text or actionable instructions. This algorithm uses techniques such as Hidden Markov Models (HMM) or deep neural networks to analyze audio input, identify spoken words, and convert them into textual representations.

4. Voice Command Processing Algorithm:
The project includes a voice command processing algorithm that interprets the recognized speech commands and maps them to corresponding actions within the application. This algorithm matches the recognized speech against a predefined set of commands and triggers the appropriate operations based on the recognized command.

5. Image Pre-processing Algorithms:
Various image pre-processing algorithms are employed to enhance the quality and readability of scanned documents and images. These algorithms may include techniques like noise reduction, image denoising, contrast enhancement, and edge sharpening to improve the OCR accuracy and object detection performance.

6. Machine Learning for Model Training:
The application may incorporate machine learning techniques for training models used in object detection, OCR, and other tasks. This involves collecting and annotating large datasets, selecting suitable machine learning models (such as convolutional neural networks), and
training the models using techniques like transfer learning or fine tuning.

These methodologies and algorithms form the core of the 3rd Eye application, enabling it to perform document scanning, object detection, speech recognition, and other critical functionalities. Through the integration of these techniques, the application aims to provide visually impaired individuals with an intuitive and efficient tool for accessing information and interacting with their surroundings.

5. Results

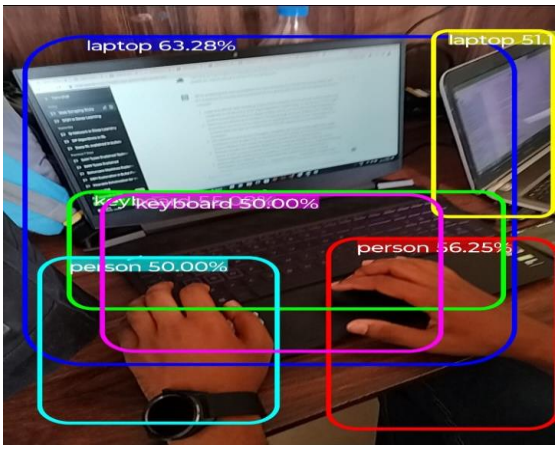

Fig. 5.1 Real Time working of the Object Recognition Module in Third Eye Application

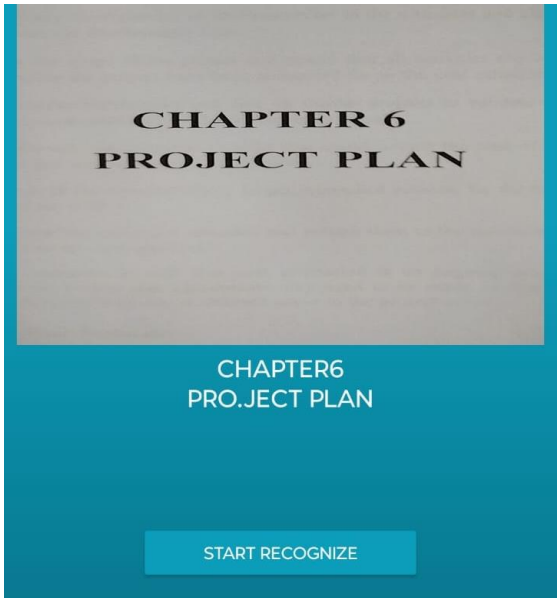

Fig. 5.2 Real Time working of the OCR Module Third Eye Application

6. Future Scope

1. Multi-Language Support:
Adding support for multiple languages in the ”3rd Eye” application can make it more versatile and useful for visually impaired users around the world. This can involve integrating additional language models and text-to-speech capabilities to enable users to scan and extract text from documents in different languages, opening the application to a wider user base.

2. Integration with Other Accessibility Tools:
Integrating ”3rd Eye” with other accessibility tools or features, such as screen readers, braille displays, or navigation aids, can further enhance its usability and effectiveness for visually impaired users.

7. CONCLUSION

In conclusion, the 3rd Eye Android application has addressed the specific needs of visually impaired individuals by providing them with a powerful tool for document scanning, text extraction, object detection, and voice-based operations. The project has successfully developed a robust system architecture, implemented various modules, and utilized technologies like TensorFlow for object detection and Google Vision for OCR. Through extensive testing and validation, the application has demonstrated its reliability and usability.
The project has made significant progress in improving the accessibility and independence of visually impaired individuals. By leveraging the capabilities of mobile technology and advanced image processing techniques, the application has the potential to greatly enhance the lives of visually impaired users, enabling them to perform tasks that were previously challenging or inaccessible.

7. REFERENCES

1. Third Eye for Visually Challenged Using Echolocation Technology / 2021 International Conference on Emerging Smart Computing and Informatics (ESCI) / IEEE
2. Blindfold: A Smartphone based Object Detection Application / International Journal for Research in Applied Science & Engineering Technology / IJRASET/Rishabh Sharma1 , Sampada Lamba2 , Khushi Maheshwari3 , Parul Yadav4
3. AI based Reading System for Blind using OCR / Proceedings of the Third InternationalConference on Electronics Communication and Aerospace Technology [ICECA 2019] /IEEE
4. Speech Technologies for Blind and Low Vision Persons / Technology and Disability 20(2008) 135–156 135 IOS Press / ResearchGate
5. Jisha Gopinath, Aravind S, Pooja Chandran, Saranya S S, "Text to Speech Conversion System using OCR", International Journal of Emerging Technology and Advanced Engineering , Volume 5, Issue 1, January 2015.
6. Raviraj S Patkar , S. Pratap Singh , Ms. Swati V. Birje, "Marker Based Augmented Reality Using Android OS", International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 3, Issue 5, May 2013.
7. Maheshwari, K., Lamba, S., Sharma, R., & Yadav, P. A Survey on Mobile Applications for the Assistance of the Visually Impaired.(1997).
8. Maheshwari, K., Lamba, S., Sharma, R., & Yadav, P. A Survey on Mobile Applications for the Assistance of the Visually Impaired.(1997).
9. Xia, Xiao-Ling, Cui Xu, and Bing Nan. "Facial expression recognition based on tensorflow platform." ITM Web of Conferences. Vol. 12. EDP Sciences, 2017