iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57531

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 11 Issue XII Dec 2023- Available at www.ijraset.com

Smart Eye for Visually Impaired People

Mamata Kishor Jain
Student, M.S. Ramaiah Institute of Technology, Karnataka, India

Abstract: This paper presents a novel mobile application tailored to enhance the daily experiences of individuals with visual impairments. Leveraging cutting-edge real-time object detection and Optical Character Recognition (OCR) techniques, the app, constructed using the TensorFlow algorithm and the comprehensive Common Objects in Context (COCO) dataset, aims to empower users by providing invaluable assistance through their mobile devices. The application utilizes the device's camera to seamlessly identify objects in real-time via the object detection algorithm, while simultaneously extracting and vocalizing text from captured images through OCR technology. A crucial aspect of the app lies in its commitment to user-friendliness, ensuring a straightforward interface tailored specifically for the needs of visually impaired individuals. The efficacy of the system was rigorously evaluated, focusing on both accuracy and real-time performance. Results conclusively demonstrate the app's effectiveness and efficiency in offering essential support to visually impaired users, marking a significant stride toward fostering independence and ease in their daily lives. Remarkably, this mobile application stands out for its affordability and accessibility, providing a practical and inclusive solution for individuals with visual impairments. By delivering audio feedback about identified objects and recognized text, this innovative app contributes to a more accessible and autonomous lifestyle for its users, contributing positively to the broader landscape of assistive technology for the visually impaired.
Keywords: Encompassing text recognition, object detection, Android, OCR, TensorFlow and COCO, Blind people.

I. INTRODUCTION

The paramount significance of eyesight among human senses cannot be overstated, as it serves as the linchpin for comprehending and engaging with the surrounding environment. Unfortunately, the World Health Organization (WHO) reports that over 285 million people grapple with eyesight challenges or visual impairment. Such conditions pose formidable obstacles to routine activities, such as identifying objects, reading text, and safely navigating thoroughfares. Visual impairment, therefore, emerges as a pervasive disability that profoundly restricts an individual's capacity to interact with the world.

In response to the pressing need for innovative solutions, this paper introduces "REBEL" (Real- Time Object Detection Learning), an Android application harnessing the capabilities of deep learning, specifically object detection and optical character recognition (OCR), to empower individuals with visual impairments. By integrating the robust TensorFlow algorithm and the comprehensive COCO dataset, REBEL emerges as a groundbreaking mobile tool designed to ameliorate the challenges faced by visually impaired individuals. TensorFlow, an open-source framework, serves as the technological backbone, facilitating the development of machine learning models tailored for object detection and OCR tasks. The COCO dataset, a cornerstone in computer vision applications, enriches the app's capabilities with a vast array of over 3.3 thousand images and 2.5 million meticulously labeled object instances spanning 80 distinct classes. REBEL's prowess lies in its real-time object detection and OCR functionalities, providing visually impaired users with a seamless and dynamic means of identifying objects and reading text through their mobile devices. By seamlessly integrating deep learning technologies into an accessible Android application, REBEL stands as a promising catalyst for enhancing the autonomy and independence of individuals grappling with visual impairments. The potential impact of this innovation extends to facilitating daily tasks, enabling information access, and surmounting challenges that were once considered insurmountable. This paper delves into the intricacies of REBEL's development, validating its efficacy, and illuminating the transformative possibilities it holds for the visually impaired community.

II. RELATED WORK

This section delves into pioneering efforts in the realm of object detection models. Galvez et al. [2] underscore the precision achievable in object classification and detection through recent advancements in deep neural networks within image processing. Xiang et al. [3] focus on Single Shot Detection (SSD), an algorithm lauded for its speed, yet acknowledging its limitations in identifying smaller objects due to a lack of contextual awareness beyond proposal boxes. Introducing a shorthand for Single Shot Multi-Box Detector (CSSD), the paper explores two variants, demonstrating how multi-scale context modeling significantly enhances detection precision.

Adnan Othman et al. [4] advocate for the integration of OpenCV libraries and deep learning techniques for real-time video recognition and object detection, utilizing Raspberry Pi 3 for implementation. Their approach incorporates enhancements like multi-scale features, default boxes, and depth-wise separable convolution, contributing to heightened accuracy in object detection and recognition.

EUN Kim et al. [5] focus on domain-specific data augmentation, reporting a substantial 30% improvement in average accuracy for on-road object detection. Leveraging Convolutional Neural Networks (CNN) in a live environment, their findings affirm the suitability of their model for real-time applications.

Chengcheng Ning et al. [7] spotlight Single Shot Multi-Box Detector (SSD) as one of the fastest object detection algorithms but acknowledge a performance gap. They propose a technique to enhance the algorithm's classification accuracy without compromising speed.

Jones and Viola [8] introduce integral images as a novel representation, facilitating faster object detection. They advocate for AdaBoost and Cascade methods to efficiently discard background regions.

The subsequent discussions encompass various approaches, from Ramadhani et al.'s reduction of set vectors for early rejection to Sung and Poggio's view-based face and non-face model. The evolution of classifiers, such as neural networks, SVM, AdaBoost, Bayes, and the introduction of SVM training by Osuna et al., set the foundation for subsequent research in object detection.

The paper's structure unfolds with an initial exposition on object detection and deep learning in the first section, followed by a survey of related work in the second section. The third section comprehensively reviews existing object detection techniques, while the fourth section illuminates the proposed model. Experimental results are presented in the fifth section, and the paper concludes in the sixth section, summarizing the research findings.

III. METHODOLOGY

The proposed system addresses the challenges faced by visually impaired individuals through the development of a sophisticated smart eye application.

The system leverages computer vision and deep learning technologies to empower visually impaired users with independent object detection, navigation assistance, and text-to- speech capabilities.

At the core of the system is an advanced object detection algorithm, integrated into the smart eye application. The algorithm utilizes real-time image processing techniques through the device's camera. When activated, the application captures images and employs a pre-trained deep neural network model for object detection.

Detected objects are then announced audibly through a speech synthesis module, providing users with real-time information about their surroundings. This functionality significantly enhances the user's situational awareness, enabling them to identify and comprehend objects without external assistance.

Furthermore, the system incorporates a navigation feature to assist visually impaired users in safe and independent mobility. Upon object detection, the application not only vocalizes the identified object but also offers navigation guidance. Through a combination of spatial awareness and predefined navigation rules, the application suggests directional cues to guide the user safely through their environment. This integration of object detection and navigation aims to enhance the user's autonomy in daily activities such as walking.

Another pivotal aspect of the system is its text-to-speech functionality, designed to facilitate independent reading for visually impaired users.

Users can capture images of book pages using the device's camera, and the application employs Optical Character Recognition (OCR) technology to extract text from the images. The extracted text is then converted into audible speech, allowing users to comprehend the content of the book without relying on external assistance. This feature significantly contributes to the users' ability to access and enjoy written content independently.

In essence, the proposed smart eye system combines object detection, navigation assistance, and text-to-speech capabilities to provide a comprehensive solution for visually impaired individuals.

By harnessing cutting-edge technologies, this system not only enhances the users' awareness of their surroundings but also promotes their independence in navigating and accessing written information, ultimately contributing to an improved quality of life for the visually impaired.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

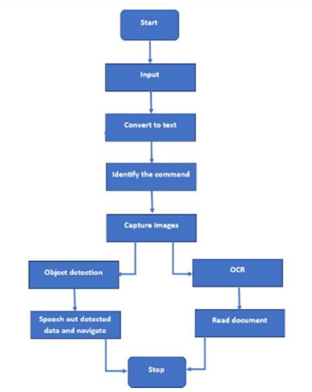


Fig 1: Activity Diagram

1) Module 1: Optical Character Recognition (OCR)

Utilizing the camera module, this system captures images containing text, achieving high-precision text extraction. Employing a Text-to-Speech engine, the extracted text undergoes conversion into audible speech, facilitating auditory accessibility for the user.

2) Module 2: Object Detection

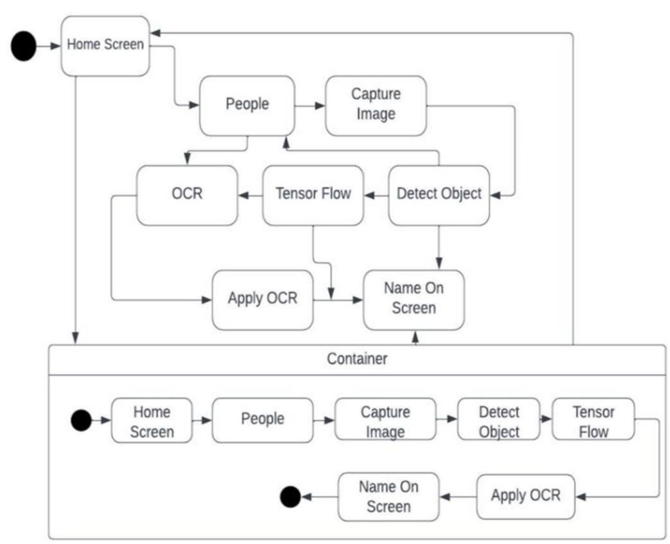
This module employs the camera to capture images, extracting objects with a refined perception. Leveraging a Text-to-Speech engine, it converts the recognized text into audible speech. Object detection, a facet of computer vision and image processing, entails identifying objects within digital images and videos. This identification includes specifying object location through bounding boxes and determining their respective types or classes. Object detection enhances the understanding of surroundings for individuals with visual impairments, promoting independence.

Input: The REBEL system captures object and text images from the camera, processing them using the TensorFlow algorithm.

Output: The REBEL system utilizes the Text-to- Speech engine to audibly relay the extracted text or speak the object name. Additionally, it displays the object name and accuracy on the mobile screen.

3) Module 3: Text Capture

The Optical Character Recognition (OCR) system within this module facilitates image capture or selection from the gallery. Following image selection, users can specify the desired language. The system subsequently employs a Text- to-Speech (TTS) engine to discern text data from the image with precision. This capability extends to capturing text images via the camera and extracting the text accurately. The extracted text is then audibly presented to the user through the Text-to-Speech engine.



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

4) Module 4: Object Detection

This module involves the capture of images through a camera, followed by the application of TensorFlow algorithms for data matching. Upon identifying a match, the object's name and accuracy are displayed on the screen. Simultaneously, audio feedback is generated utilizing the Text-to-Speech (TTS) technique to convey the result. The system exhibits proficiency in detecting various objects, including but not limited to persons, cell phones, bottles, and more.

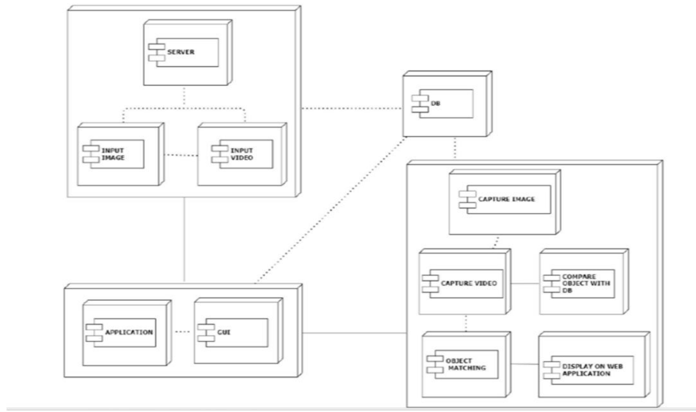


Fig 3: Component Diagram

A. Software and Hardware Requirements User Interfaces

- Operating System: Windows OS
- Programming Language: Java, XML
- Development Environment: Android Studio

1) Hardware Interfaces

- 4-bit Microsoft Windows® 8/10/11.
- x86-64 CPU architecture; 2nd generation Intel Core or newer, or AMD for a Windows Hypervisor.
- 8 GB RAM or more.
- Minimum 8 GB of available disk space (IDE + Android SDK + Andr).
- Minimum screen resolution of 1280 x 800.

2) Software Interfaces

- Operating System: Windows OS
- Programming Language: Java, XML
- Development Environment: Android Studio
- 3.3.4 Communication Interfaces:
- This project is compatible with all types of Android phones with active internet connectivity.

B. User Characteristics

1) Flash Screen.
2) Home Screen. A OCR
Object Detection.

➢ *OCR*

- Camera
- Capture Image
- Apply OCR
- TTS

REBEL: Real time object detection Learning

- *Object Detection*
- Camera On
- Detect Object
- Tensor Flow
- Name on Screen-TTS
- Show the Result.

C. Applications

1) Cost-Effective
2) Easy to use.
3) User friendly.
4) Reduce the hardware component.
5) Reduce the hardware handling.
6) Eco-friendly

IV. CONCLUSION

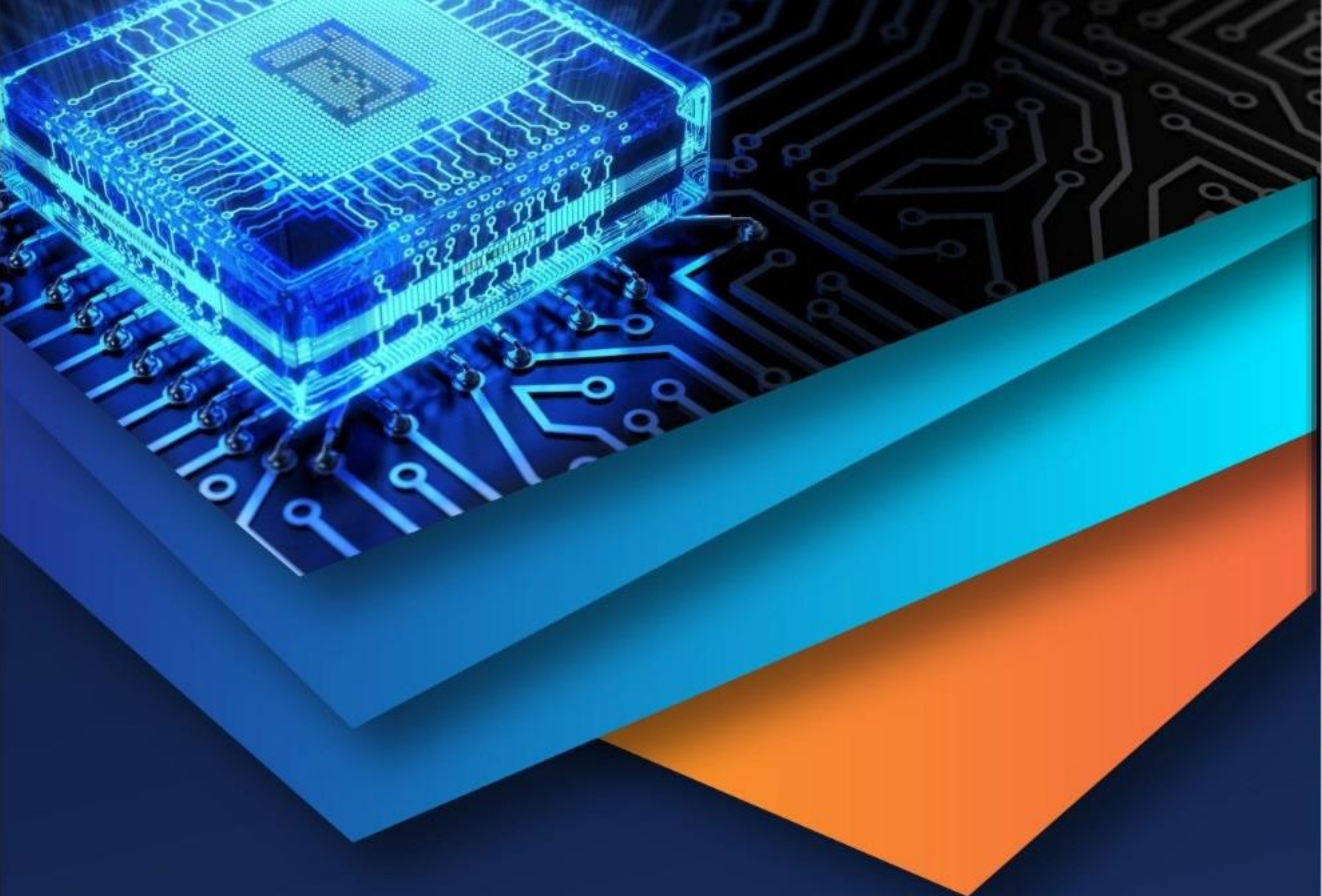
In this paper, our endeavor focuses on developing a groundbreaking Real-Time Object Detection Learning system with voice sensors, aimed at empowering visually impaired individuals to better comprehend their surroundings. This innovative project leverages advanced computer vision techniques, paving the way for a blind visualization system. Looking ahead, we envision expanding our impact by considering an application for language conversion tailored for nursery children. The web-based Object Detection system, with its primary goal of identifying diverse objects across image types, utilizes sophisticated methods, including shape and edge feature extraction. A vast image database ensures precise object detection and recognition. The user-centric interface enhances accessibility, allowing for seamless retrieval of desired images. Notably, the system's unique feature guides users in navigating safely, underscoring its novelty in promoting independence and safety for the visually impaired.

REFERENCES

[1] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: CVPR. (2016).
[2] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra, J. M. Z. Maningo, "Object Detection Using Convolutional Neural Networks", TENCON 2018-2018 IEEE Region 10 Conference, 2018.
[3] P. Poirson, P. Ammirato, C.-Y. Fu, J. Liu, Wei Koeck, A. C. Berg, "Fast single shot detection and pose estimation", International Conference on 3DVision(3DV), 2016.
[4] N.A. Othman, I. Aydin, "A new deep learning application based on movidius ncs for embedded object detection and recognition", 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018.
[5] Wei Xiang Dong-Qing Zhang Heather Yu Vassilis Athitsos, Context-AwareSingle- ShotDetector, 2018 IEEE Winter Conference on Applications of Computer Vision, pp. 1784-1793, 2018.
[6] D. Alamsyah and M. Fachrurrozi, Faster R- CNN with Inception V2 for Fingertip Detection in Homogenous Background Image, IOP Conf. Series: Journal of Physics: Conf. Series 1196 (2019) 012017.
[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, "SSD: Single shot multibox detector", 2015.
[8] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, vol. 1, pp. 511-518, 2001.
[9] C. Papageorgiou, M. Oren, T. Poggio, "A General Framework for Object Detection", Proc. IEEE Int'l Conf. Computer Vision, 1998.
[10] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:1–10.
[11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, "SSD: Single shot multibox detector", 2015.
[12] R. Girshick, "Fast R-CNN," in IEEE International Conference on Computer Vision (ICCV), 2015.

AUTHORS PROFILE

Ms. Mamata Kishor Jain – is a Computer Engineering Student Studying at M.S.Ramaiah institute of technology Bangalore, Karnataka. she is pursuing Master's Degree.



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY