Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 1, August : 2023

MACHINE LEARNING BASED REAL-TIME OBJECT DETECTION LEARNING AND OCR FOR ASSISTING THE VISUALLY IMPAIRED

Dr..Rajiv R. Bhandari, Dr. Khyati R. Nirmal, Aditi Kolhe, Sapana Shinde, Sakshi Jadhav, Sayali Vyavahare Department of Compuyter Engineering, SNJB's Late Sau. KBJ College of Engineering, Chandwad, Dist: Nashik, India
rajivrbhandari@gmail.com, khyatinirmal@gmail.com, sakshijadhav17501@gmail.com, kolheaditi5@gmail.com, sapana8434@gmail.com, sayali.vyavahare17@gmail.com

Abstract
This paper presents the development of a mobile application that uses real-time object detection and Optical Character Recognition (OCR) techniques to assist visually impaired individuals in their daily activities. The application is developed using the Tensor flow algorithm and the Common Objects in Context (COCO) dataset. The object detection algorithm is used to identify objects in real-time through the camera of the mobile device, while the OCR technology recognizes and reads text from images captured by the camera. The application provides audio feedback to the user about the identified objects and the text read by OCR. The user interface is designed to be simple and user-friendly to cater to visually impaired individuals. The application's performance is evaluated accuracy and real-time performance, and the results indicate that the proposed system is efficient and effective in assisting visually impaired individuals. Overall, this mobile application provides an affordable and accessible solution for visually impaired individuals, enabling them to navigate their daily lives with greater ease and independence.

Keywords—Text recognition, Object detection, Android, OCR, Tensor flow, COCO

I. INTRODUCTION

Visual impairment is a common disability that affects a significant portion of the population worldwide. For individuals with visual impairment, tasks such as object recognition and text reading can be extremely challenging, limiting their ability to interact with the world around them. In recent years, advances in machine learning technology, specifically object detection and optical character recognition (OCR) have shown great potential to address this issue.

The paper introduces a simple android app based on object detection and reading text application named "REBEL"(Real Time Object Detection Learning) to help the visually impaired. The development of a mobile application that leverages the power of object detection and OCR can weighty improve the quality of life for individuals with visual impairments. This mobile application uses the Tensor flow algorithm and the COCO dataset to enable real-time object detection and OCR on a mobile device. The Tensor flow algorithm is a popular and powerful open-source framework that enables the creation of machine learning models for a variety of tasks, including object detection and OCR. The COCO dataset is a widely used dataset in OpenCV, containing over 3.3 thousand images with over 2.5 million object cases labeled across 80 different classes. By utilizing the Tensor flow algorithm and COCO dataset, this mobile application can accurately identify objects and read text in real-time, providing a powerful tool for visually impaired individuals to navigate and interact with their surroundings. This application has the potential to improve the independence and autonomy of individuals with visual impairments, enabling them to access information and perform daily tasks that would otherwise be difficult or impossible.

II. LITERATURE SURVEY

Pattern recognition is a field that is constantly evolving, and researchers continue to propose new technologies to address the challenges of object detection and text recognition. With each approach seeking to improve these areas in its unique way, it remains an active area of research. In this survey,

They provide a thorough examination of the various approaches proposed for text recognition, highlighting both their advancements and limitations.
The main objective of this work is to assist visually impaired individuals in reading textbook material and identifying objects in their environment [1]. The system receives input in the form of an image captured by a web camera, which can be used for either textbook reading or object detection, based on user preference. The Raspberry Pi serves as the microcontroller for processing the entire system. Optical character recognition (OCR) software is utilized for textbook reading.
Real-time object detection and audio conversion for the visually impaired is a challenging task[2]. However, with recent advancements in computer vision, it has become possible to develop efficient real-time object detection systems. This paper proposes a straightforward Android application that aids visually impaired individuals in comprehending their environment. The application captures the surroundings using the phone's camera and performs real-time object detection using Tensor Flow's object detection API. The detected objects are then converted into audio output for ease of understanding.
The field of textbook recognition is an important area as it finds colorful useful operations in computer vision similar to document analysis, image hunt, robot navigation etc. Indeed however, a bunch of exploration has been carried out in this area but still there's a room for the enhancement as none of the presented styles are error-free due to the challenges that may include the fountain of the characters, multi-oriented textbook, and the quality of filmland that contain textbook[4].
Recent advancements in hardware and software technologies have opened up new avenues of exploration for computer vision techniques [7]. These techniques involve extracting meaningful information from images to make intelligent decisions. In this project, we have utilized two types of computer vision techniques, namely optical character recognition (OCR) and object recognition, to develop an intelligent database management system for post-examination process control.

III. PROPOSED WORK

The system was implemented as an Android application capable of detecting various objects in real-time, while also including a real-time text reader. The REBEL system involved the development of an object recognition Android application, which utilized Google's Tensor Flow object detection API model, executed using the SSD algorithm. Additionally, the application included a real-time text reader, utilizing Google's TTS machine and Google Play Services Mobile Vision API. The text recognizer class was utilized to detect text from a real-time environment, enabling comprehensive text recognition capabilities in the application. The project demonstrates the effectiveness of integrating multiple technologies and APIs for developing a robust and efficient object detection and OCR system, suitable for real-time deployment in an Android application.

A. System Outline

The project utilized a range of technologies, including Android Studio as the primary integrated development environment for Android applications. The Android platform provided support for real-time processing via Android Intent. To enable high-performance numerical computing, the Tensor Flow library was integrated for running object detection models. Its flexible architecture allowed for seamless deployment across multiple platforms. For real-time object detection, the system utilized the SSD-Mobile-Net COCO model. This model was based on a single-shot detection (SSD) architecture, which predicted bounding box locations and detected objects in the form of bounding boxes. The system included two object detection modules and a real-time text collection module, enabling comprehensive object recognition capabilities. The project demonstrates the potential of combining Tensor Flow and the COCO dataset for developing robust and efficient object detection and OCR systems, integrated into an Android application for real-time processing.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, No. 1, August : 2023

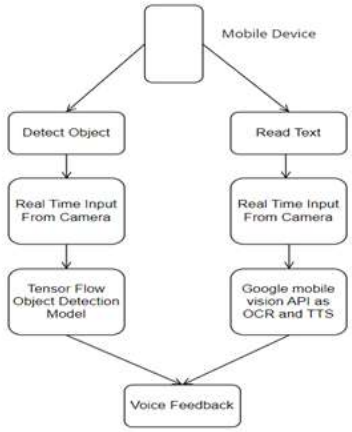

Fig. 1 Workflow of REBEL

B. Background Study

The system was developed by integrating various technologies, including Android Studio, the official integrated development environment (IDE) for developing Android applications. The Android platform supports real-time processing through Android Intent. Tensor Flow library is utilized to run object detection models within the Android application, providing high-performance numerical computing. Its flexible architecture enables seamless deployment across multiple platforms.

For real-time processing, the system utilizes the SSD-Mobile-Net COCO model. This model is based on a single-shot detection (SSD) architecture, which predicts bounding box locations and detects objects in the form of bounding boxes. The system employs two object detection modules and a real-time text collection module, providing comprehensive object recognition capabilities.

Object Detection:

The Tensor flow Object Detection algorithm is a cutting-edge technique that can accurately detect objects in real-time. The algorithm divides the input image into a grid and predicts bounding boxes and class probabilities for each cell. Non-maximum suppression is used to refine the predicted bounding boxes and eliminate redundant detections. The algorithm is trained on large datasets such as COCO, which includes over 330,000 images representing 80 different object categories. This extensive training data enables the algorithm to achieve high accuracy in object detection tasks.

Optical Character Recognition(OCR):

OCR is the process of feting textbook from an image. OCR can help visually disabled people read textbook by converting textbook into speech. OCR can also be used to identify textbook in images, similar as road signs and markers. We use OCR to fete textbook in real- time and give audio feedback to visually disabled people.

Tensor flow:

The Tensor Flow Object Detection API is an open-source framework built on top of Tensor Flow that simplifies the development, training, and troubleshooting of object detection models. It provides a set of pre-trained models that have been trained on various datasets, such as the COCO (Common Objects in Context) dataset, the KITTI dataset, and the Open Images Dataset. These pre-trained models serve as a starting point for users to customize and fine-tune the models for their specific object detection tasks.

C. Algorithm

Here's a algorithm for developing a mobile application that captures an image, identifies objects using the Tensor flow algorithm, and provides audio feedback for blind people:

- Step 1: Open the camera:

Open the camera on the mobile device and display the camera view on the screen.

Provide audio instructions to the user on how to position the camera and capture the image.

- Step 2: Capture the image:

When the user presses the capture button, capture the image from the camera.

- Step 3: Object detection:

Apply the Tensor flow algorithm to the captured image to detect the objects present in the image.

- Step 4: Object identification:

Once the objects are detected, identify the object using a pre-trained machine learning model or database.

Retrieve the object's description, uses, and other related information.

- Step 5: Display object information:

Display the object's information in text format on the screen, including the object's name, description, uses, and other relevant information.

Provide audio feedback of the same information in a clear and concise manner for the blind user.

- Step 6: Text identification (if applicable):

If the image contains text, use OCR (Optical Character Recognition) to extract the text from the image.

Convert the text into speech for the blind user to hear.

- Step 7: Repeat or capture new image:

Provide an option to repeat the process or capture a new image.

- Step 8: Save the image and its information:

Save the captured image and the object information in a database for future reference.

This algorithm provides a general framework for building an application that identifies objects and provides audio feedback for blind people. The details of implementing the algorithm can vary based on the specific technologies used and the requirements of the application.

D. Dataset

The model in this project was trained using the Common-Object-in-Context (COCO) dataset, which is a well-known large-scale image recognition dataset commonly used in computer vision research. The COCO dataset consists of over 330,000 images, with more than 2.5 million labeled object instances belonging to 80 different object categories. Its comprehensive and diverse nature makes it one of the most widely used datasets for object recognition tasks in the field of computer vision.

IV. RESULT AND DISCUSSION

Improved accessibility: The mobile application would allow visually impaired people to have greater access to information and the ability to identify objects in real-time, improving their quality of life and independence.

Efficiency: The real-time object detection feature of the application enables users to quickly and easily identify objects without having to wait for a response, which can save time and increase productivity.

Accuracy: Using TensorFlow algorithm and COCO dataset, the application can achieve a high level of accuracy in object detection and OCR, ensuring that users receive reliable information.

Versatility: The application can be used in a variation of settings and situations, making it a flexible tool for visually impaired individuals.

Innovative: The REBEL presents a new and innovative solution for visually impaired individuals, demonstrating the use of advanced technologies to solve a real-world problem.

Potential for further research: The research paper can work as a initial point for further studies on the application of machine learning and computer vision in assistive technologies for people with disabilities.

Result:

The result of the above discussion is that the paper presents a mobile application that uses object detection and OCR technologies to assist visually impaired individuals in their daily activities. The application is designed to be user-friendly and provides audio feedback to the user about the identified

objects and the text read by OCR. The proposed system is evaluate the accuracy and real-time performance and is found to be efficient and effective in assisting visually impaired individuals.
The use of object detection and OCR technologies in a mobile application offers an affordable and accessible solution for visually impaired individuals to navigate their daily lives with greater ease and independence. The paper highlights the potential of computer vision technologies to boost the quality of life for blind people individuals and demonstrates the practical application of these technologies in a mobile application.
In this project, REBEL will detect the object in real-time by using the Tensor flow object detection model.

Fig. 2 Detected Objects

Fig. 3 Document Reading

Figure 3 displays the output of text recognition. The OCR technology accurately captured the text from the document, and the REBEL application was able to recognize it with precision. It will Read the real-time Document as it is.

V. CONCLUSION

The proposed system implementation for a REBEL application which detect the objects in real-time and OCR for blind people using the Tensor flow algorithm and COCO dataset can help visually impaired people overcome various challenges they face in their daily lives. The implementation involves data collection, data preprocessing, training the Tensor flow algorithm, integrating with OCR, mobile application development, and testing and evaluation. The accuracy and efficiency of the proposed system can be improved by using advanced techniques and testing with real visually impaired people.

VI. REFERENCES

[1] Murali, M., Shreya Sharma, and Neel Nagansure. "Reader and Object Detector for Blind." 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2020.
[2] D. Choi, and M. Kim, "Trends on Object Detection Techniques Based on Deep Learning," Electronics and Telecommunications Trends, Vol. 33, No. 4, pp. 23-32, Aug. 2018, Article.
[3] Dipu, Nadim Mahmud, Sifatul Alam Shohan, and K. M. A. Salam. "Bangla optical character recognition (ocr) using deep learning based image classification algorithms." 2021 24th International Conference on Computer and Information Technology (ICCIT). IEEE, 2021.
[4] Hassan, KM Naimul, et al. "A Dual-Purpose Refreshable Braille Display Based on Real Time Object Detection and Optical Character Recognition." 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON). IEEE, 2019.
[5] Zaki, M. A., et al. "Development of an android app for text detection." Journal of Theoretical and Applied Information Technology 97.20 (2019): 2485-2496.
[6] Salunkhe, Akilesh, et al. "Android-based object recognition application for visually impaired." ITM Web of Conferences. Vol. 40. EDP Sciences, 2021.
[7] Rizvi, Mehdi, et al. "Optical character recognition based intelligent database management system for examination process control." 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST). IEEE, 2019.
[8] Dalmeda, J., Raju, K. G., Prashanth, B., Shwetha, P., & Reddy, D. N. READER AND OBJECT DETECTOR FOR BLIND.
[9] Bazerque, Ana, Davi Moraes, and Marcela Souza. "Using Object Detection Algorithm and Optical Character Recognition to Read Data from alphanumeric tags in text." (2020).
[10] Microsoft “COCO Dataset”, http://cocodataset.org