International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

REAL TIME OBJECT DETECTION FOR VISUALLY CHALLENGED PEOPLE USING MACHINE LEARNING

Prof. Sneha A. Khaire[1], Dhanraj Nikam[2], Ajinkya Pawar[3], Chaitanya Devkar[4], Shubham Pagar[5]

Assistant Professor[1] BE Students[2,3,4,5]
Department of Computer Engineering,
Sandip Institute of Technology and Research Centre, College of Engineering, Nashik - 422213

Abstract - The latest generation of mobile phones have excellent hardware capabilities and quicker processing, which are strong enough to enable the creation of applications that enable users to connect and engage with the outside world at their own pace. This system reads OCR and utilizes the camera app on your smartphone in addition to the camera.
Optical Character Recognition (OCR) is a tool that converts printed, handwritten, or typed text into machine-readable text from a picture. This system will assist you in taking a photo of or scanning a document that is currently in your possession using the camera on your phone. After the image has been scanned, an application will read any English-language text and turn the results into speech. The speech communication channel is regarded by the authors as one of the most crucial modalities to help the blind and low vision people in the state-of-the-art accessibility area.

***Key Words*:** OCR, camera app, speech, blind, vision

1.INTRODUCTION

The visually impaired experience many difficulties in daily life, with finding things being a typical one. This is in contrast to a normal sighted person. People who are blind have more acute perceptions than sighted people do. When in a familiar setting, they are aware of basic directions and the locations of useful objects. They have trouble finding things when they're in a strange setting. In order to make it simple for users to locate the items they seek; our goal was to develop an application that gives them a general feeling of direction.
Real-time object detection using machine learning can be a useful tool for visually challenged individuals. The process involves using computer vision algorithms to identify and label objects in a live video feed or image, allowing visually challenged individuals to understand their surroundings.

2. LITRATURE REVIEW

There are a variety of techniques for text reading applications, including label reading, voice stick, brick pi reader, and pen aiding, but all of these can conduct text to speech by building datasets. A solution to this issue is the finger reading method, which eliminates previously created and stored datasets and offers a prior response of reading any text provided as input captured image.

In [4], The suggested system is an automatic book reading system, according to the authors. Utilizing Raspberry PI hardware, the text recognition procedure is carried out in this study. Python programming and algorithms like Tesseract are used to identify the characters. The voice output is the product. The image is taken with a web camera that is linked via USB to an ARM microcontroller. The OCR engine in this system uses the TESSERACT library, and the data conversion to audio uses the Flite library.
In [5], The writers suggest that important image processing methods like image acquisition and processing are discussed. Here, text is extracted from a picture using tools from the NI Vision toolkit and LabVIEW.

3. OBJECTIVE

The objectives of system are:

- The objective of real-time object detection for visually challenged people using machine learning is to provide them with a technology-assisted solution that can help them identify and navigate the physical world more effectively. By leveraging machine learning algorithms and computer vision techniques, this technology aims to detect and classify objects in real-time, providing audio or tactile feedback to the user to help them understand and interact with their environment.
- To covert the content of text to speech
- Speech out the detected data by using OCR and Object detection.

4. STPS OF THE PROPOSED SYSTEM

1. Take an image or video input from the user's smartphone camera.
2. Apply pre-processing techniques such as image resizing, normalization, and denoising to enhance the quality of the image.
3. Use object detection algorithms to identify the objects present in the image. This step can be skipped if the app is only interested in OCR of the entire image.
4. If objects are detected, apply OCR only to the relevant regions of interest (ROIs) corresponding to the objects. If no objects are detected, apply OCR to the entire image.
5. Convert the image or ROIs to text using OCR algorithms. This step involves the recognition of characters and their conversion to digital text.
6. Analyze the recognized text to identify the information relevant to the user, such as product names, prices, or labels. This step may involve natural language processing (NLP) techniques to extract meaning from the text.

7. Output the relevant information to the user in an accessible way, such as text-to-speech or a Braille display.

It's worth noting that the performance of object detection and OCR algorithms can be greatly affected by the quality of the input image, lighting conditions, and other factors. Therefore, it's important to test and refine the algorithm in real-world conditions to ensure accurate and reliable results.

5. ARCHITECTURE

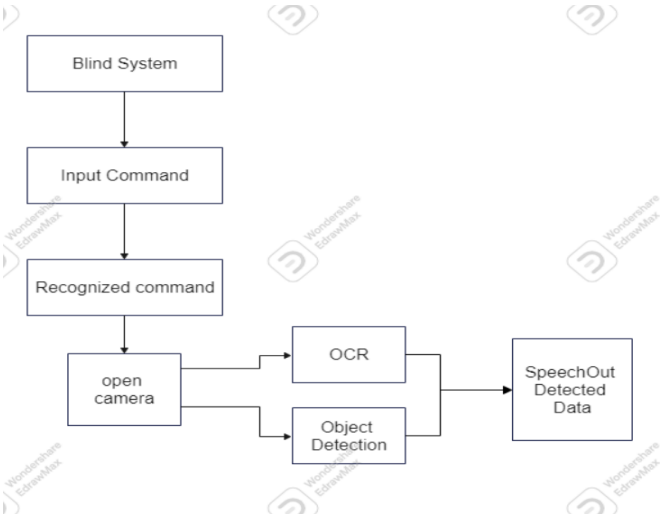

Fig -1: Block Diagram

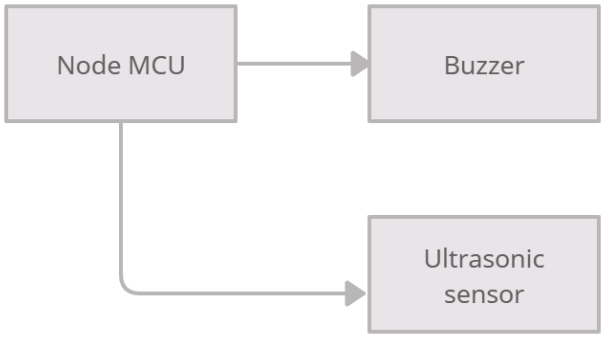

Fig -2: Hardware Working

1. Input Sensors: These are the devices that capture the input data, such as cameras or depth sensors. They are used to collect visual information about the user's surroundings.
2. Preprocessing: The raw data collected from the sensors needs to be preprocessed before it can be used by the machine learning model. This may involve tasks such as image resizing, normalization, and filtering.
3. Object Detection Model: This is the core component of the system, which uses machine learning algorithms to detect and classify objects in real-time. Popular models for object detection include Faster R-CNN, SSD, and YOLO.
4. Postprocessing: Once the objects have been detected, postprocessing is applied to refine the results and extract additional information. This may include tasks such as non-maximum suppression, bounding box regression, and object tracking.
5. Output: The final output of the system is typically provided in the form of audio or tactile feedback, which is used to inform the user about the objects in their environment. This feedback may include object names, positions, and descriptions.

Overall, the architecture of real-time object detection for visually challenged people using machine learning is designed to leverage advanced machine learning algorithms and computer vision techniques to provide users with real-time information about their surroundings, helping them to navigate and interact with the physical world more easily.

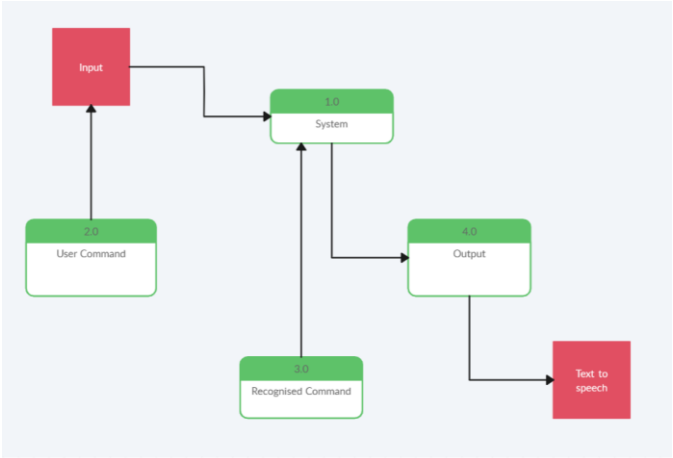

Fig -3: Data Flow Diagram

6. CONCLUSIONS

In conclusion, real-time object detection using machine learning has the potential to greatly improve the quality of life for visually challenged people. By leveraging technologies such as computer vision and deep learning, it is possible to develop algorithms that can recognize and describe the visual world in real-time, allowing users to navigate and interact with their environment more easily.

There are already several promising initiatives underway to develop assistive technologies based on real-time object detection, such as wearable devices that can identify objects and provide audio feedback to the user. As these technologies continue to evolve and improve, we can expect to see even more advanced and sophisticated applications in the near future.

Overall, the use of machine learning for real-time object detection is a promising and exciting area of research, with the potential to transform the lives of visually challenged people and help them to live more independently and confidently.

REFERENCES

[1] Maheshwari, K., Lamba, S., Sharma, R., & Yadav, P. A Survey on Mobile Applications for the Assistance of the Visually Impaired.(1997).
[2] Bhowmick, Alexy, and Shyamanta M. Hazarika. "An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends." Journal on Multimodal User Interfaces 11.2 (2017): 149-172.
[3] David, Robert, et al. "Tensorflow lite micro: Embedded machine learning on tinyml systems." arXiv preprint arXiv:2010.08678 (2020).
[4] Aaron James S, Sanjana S, Monisha M, "OCR based automatic book reader for the visually impaired using Raspberry PI", Vol. 4, Issue 7, January 2016.
[5] Christopher G Relf, "Image Acquisition and Processing with LabVIEW", CRC Press, 2004.
[6] Sonia Bhaskar, Nicholas Lavassar, Scott Green, "Implementing Optical Character Recognition on the Android Operating System for Business Cards".
[7] G. Xydas, G. Argyropoulos, Th. Karakosta and G. Kouroupetroglou, An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents, in: Proc of the 11th Int Conference on Human-Computer Interaction, Las Vegas, Nevada, USA, 22–27 July, 2005.
[8] V. Zarikas, G. Papatzanis and C. Stephanidis, An Architecture for a Self-Adapting Information System for Tourists, in: Proceedings of the Workshop on Multiple User Interfaces over the Internet: Engineering and Applications Trends (in conjunction with HCI-IHM'2001), Lille, France, 10–14 September, 2001.
[9] 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.
[10] 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.
[11] Ashwani Kumar, Ankush Chourasia, "Blind Navigation System Using Artificial Intelligence", March 2018
[12] Aaron James S, Sanjana S, Monisha M, "OCR based automatic book reader for the visually impaired using Raspberry PI", Vol. 4, Issue 7, January 2016.
[13] Christopher G Relf, "Image Acquisition and Processing with LabVIEW", CRC Press, 2004.
[14] Sonia Bhaskar, Nicholas Lavassar, Scott Green, "Implementing Optical Character Recognition on the Android Operating System for Business Cards".
[15] Julinda Gllavata, Ralph Ewerth and Bernd Freisleben, "A Robust Algorithm for Text Detection in Images".
[16] Chucai Yi, Yingli Tian, "Scene Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration", IEEE Transactions on Image Processing, Vol. 23 No. 7, July 2014