Real Time Object Detection for Visually Impaired Person Using Tensor Flow Lite

Aditi Hemant Khandewale [+], Dr. Vinaya V. Gohokar and Pooja Nawandar

MIT WORLD PEACE UNIVERSITY, India

Abstract. Real Time Object Detection Using Tensor Flow Lite system has been developed to help visually impaired people for navigation and surrounding objects detection. This system is based on raspberry pi, a single board compute model and Tensor Flow lite framework. The algorithm developed is tested for detecting objects like a table, a chair, a TV, a laptop, a mouse, a cell phone, a bottle etc. This system is capable of detecting people as well as objects. The detection accuracy of 70% is achieved. The testing is done in varying light, background, and distance in indoors as well as outdoor scenarios. This system uses Google based sample quantized SSDLite-MobileNet-v2 object detection model, which is trained of the MSCOCO dataset and converted to run on TensorFlow Lite. The information regarding the detected object is converted into audio for guiding the visually challenged person

Keywords: Tensor flow lite, MS-coco, Raspberry pi, object detection, gtts

1. Introduction

There are huge number of visually impaired people, may be partially or fully blind. According to the World Health Organization (WHO) [1], the number of people of all ages visually impaired is estimated to be 285 million, among which 39 million are blind, this is global statistics. These people suffer many problems during their day to day activities. This project is an attempt to help them, detect objects in the surrounding and make them aware of things around them with the help of audio message. A blind person is helped in navigation by detecting the obstacle on his path. It is even possible to provide information regarding staircase in front of him. This may help to avoid accidents.

It can be a powerful guide for blind people making himself-sufficient and independent. This system is developed using raspberry pi model 3B+, web camera and Bluetooth earphones. An open source end to end machine learning platform, TensorFlow is used for developing intelligent object detection algorithm.

TensorFlow lite, the version of TensorFlow is specifically used because it can run Tensor Flow models on IoT, embedded and mobile devices. TensorFlowlite is having low latency and smaller binary size which makes it easy to design device at the edge of network. This improves latency, connectivity, and privacy. Google text to speech (gtts) a python library interface with Google translate text to speech API. It can read unlimited length of speech.

This model is tested with different light conditions. Even in the absence of light or dim focus of light this model is giving accurate results. Multiple objects in same frame are detected correctly. It is tested for up to 5 objects. The system detects objects accurately till distance 10 to 12 feet's.

Tensor flow framework:

TensorFlow is open source framework for machine learning and used for many applications developed by Google with Google dataset.

[+] Corresponding author. Tel.: + 91 7030449501.

E-mail address: aditikhandewale9@gmail.com.

TensorFlow has developed dataset for many applications, it is having flexible architecture written in C++ or python, there are many frameworks available like Caffe, Keras, Theano, Torch etc. but among them TensorFlow have easy model building and robust machine learning production anywhere.

TensorFlow lite can be used for mobile, IoT embedded models. It has two main components as in Fig. 1

1) Interpreter
2) Converter

Interpreter in which predictions are done on the basis of input data and converter generates flat buffer file which is then deployed by client device. It internally usesTensorFlow lite interpreter file. TensorFlow lite supports APIs for different languages like Java, Swift, Objective-C, C++, and Python. In TensorFlow lite there are Model optimization tools, which include quantization that can reduce size and increase performance of models without sacrificing accuracy. TensorFlow lite have Efficient model format includes FlatBuffer which is optimized for small size and portability. FlatBuffer is an efficient cross platform serialization library for C++, Rust, C#, Java, JavaScript, Typescript. FlatBuffer is good in Memory efficiency and speed .
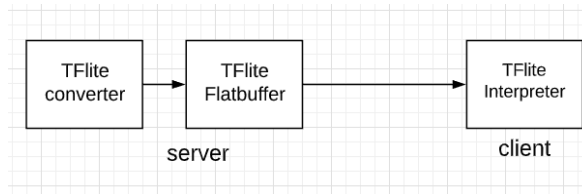

Fig. 1: Components in TensorFlow

TensorFlow lite is suitable on platforms like android, iOS and raspberry pi. There are 80 objects in quantized COCO SSD Mobile Net v1 model. This includes objects a blind person comes across in day to day life like a table, a chair, a cell phone etc. it can even detect person.

2. Related Work:

Table 1: Comparative study of research work for visually impaired people for object detection

papers	Technologies used	Platforms used
(2), (3)	You Only Look Once (YOLO), robust estimator a RANSAC-based paradigm (2) CNN (3)	Windows 7(2) Raspberry pi (3)
(4), (5)	OpenCV	Raspberry pi (4)
(6)	ultrasonic sensor, SLAM based navigation along with barcode reading	Raspberry Pi
(7), (8)	multilingual neural text-to-speech (TTS), R-CNN algorithm	Android
(9)	R-CNN, RNN,	
(10)	Tensor flow	Raspberry Pi.

Table 1 compares the research work carried out using computer vision based object detection system for visually impaired people. A Frame-work assisting the Visually Impaired People for Common Object Detection and Pose Estimation in Surrounding Environment is proposed in [2] which was performed on

Windows 7 using You Only Look Once (YOLO). A robust estimator a RANSAC-based paradigm technologies was used by the researchers. In [3] Design and Implementation of an Embedded Real-Time System for Guiding Visually Impaired Individuals is discussed which uses same technology.

A Vision Module for Visually Impaired People by Using Raspberry PI Platform is proposed in [4-5] which uses Open cv on raspberry pi. In [6] Assistant for Visually Impaired using Computer Vision platform on Raspberry Pi as well as they used ultrasonic sensor, SLAM based navigation along with barcode reading. Real Time object detection and multilingual speech synthesis is proposed in [7-8] and research work was performed on android using multilingual neural text-to-speech (TTS) and R-CNN algorithm.

Intelligent Platform for Visually Impaired Children for Learning Indoor and Outdoor Objects is proposed in [9], they have used R-CNN, RNN. In [10] Designing an Obstacle Detection and Alerting System for Visually Impaired People on Sidewalks was developed by authors and was performed on Raspberry Pi used Tensor flow.

After detailed study of computer vision-based object detection techniques employing AI, it was decided to develop raspberry pi based object detecting system using TensorFlow lite. Tensor flow lite framework is suitable as it is lightweight solution for mobile, internet of things and embedded devices. Tensor flow lite is preferred because of fewer complexities and easy to perform on platform like raspberry pi.
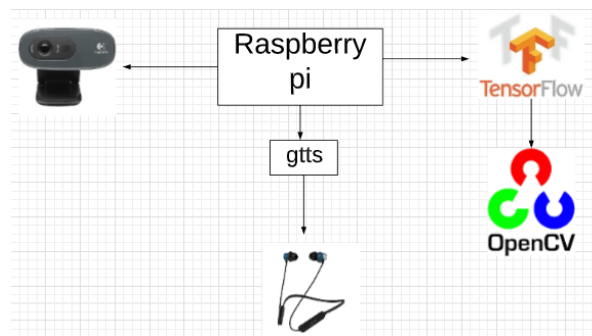
2.1. Methodology

Fig. 2: "Block diagram"

The block diagram of Object detection system for blind people is described in Fig. 2. In this architecture, raspberry pi model 3b+ is used. Specifications are given in Table 2
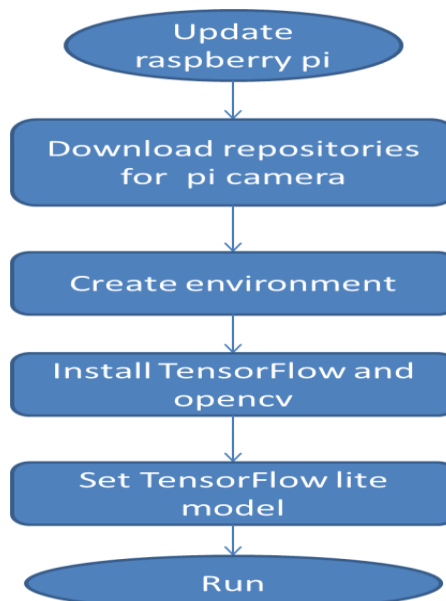
Raspbian stretch operating system is installed.

Fig. 3: Workflow

Workflow for implementation of object detected model is explained in Fig. 3 first raspberry pi is updated. Next step is to install dependencies for pi camera, then it is needed to create environment. An environment is created in order to avoid version complexities and to isolate package installation from the system. In that environment install TensorFlow and open cv. The next step is to set TensorFlow model with dataset and last but important step is test results.

Image capturing is done with the help of Web camera or pi cam. The model takes input image. It is expected to have image with 300x300 pixels and there are 3 channels per pixel (Red, Green, Blue). Andobject detection model can identify which of a known set of objects might be present and provide information about their positions within the image. The detected object model has flattened buffer of 270,000 byte values (300x300x3) and model is quantised representing value between 0 to 255.

A quantized neural network model is used which has an of 8-bit integer value, because it runs faster and speed up the actions. Frames per seconds (FPS) in Tensor Flow lite model is observed up to 4.4 which is comparatively more than TensorFlow which is 3.

Output of model has 4 arrays (0,1,2,3), in which 0 represents location that is bounding box [top, left, bottom, right]. 1 represents classes it is as integer of 10 integers, indicating index of class label. 2 represents scores it is an array of 10 floating values between 0 and 1, as it is a probability to indicate class detection. 3 indicates number and detections it is an array of length 1 indicating total number of detection results.

OpenCV supports the deep learning frameworks such as TensorFlow. OpenCV library is used for real time computer vision developed by Intel, includes statistical machine learning library which contains SVM(support vector machine), DNN(Deep neural network), K-NN, naïve neural network etc used for many real time applications like emotion recognition, face recognition, object detection, mobile robotics, motion tracking etc. refer section 1 for TensorFlow details.

Google text to speech (gtts) is used here which uses (TTS) API. This library is used to read name of object detected. It can read unlimited number of letters and digits.

Table 2: Specifications are as follows:

Raspberry pi	RPI 3B+
RPI camera	REV1.3
Web camera	Logitech HD720P
SD card	32 GB
Tensor Flow	TensorFlow lite
Operating system	Raspbian stretch

2.2. Analysis

For testing the performance of developed model, it was tested under various scenarios. This includes

- Variation in Light
- Distance from camera
- Background
- Number of objects in frame

There are 80 sample images in the data set, out of which few were tested, minimum 51% and maximum 83% accuracy is observed testing is done indoor and outdoor also, moreover light effects giving better and accurate results, this architecture can correctly label up to 5 objects in one capture. The conducted test showed that up to 10 to 12 feet, this is giving correct output for a number of objects in the range.

Testing was carried out for various classes of same object for ex. Cell phone of 10 different companies were tested out of which 7 were detected accurately. Person was detected correctly till distance of 12 feet's.

Fig. 4: "chair detection"

In Fig. 4, object tested was chair, as a result chair was detected with 68% accuracy.

Fig. 5: "cell phone detection"

In Fig. 5, cell phone was detected with accuracy of 65% accuracy.

Results for same object differ depending on if the object is present in focus of light or outside focus of light

Fig. 6: "mouse detection (in focus of light)"

In the focus of light as shown in Fig. 6. It is giving 83% accuracy of detection of mouse as object and out of focus of light and half captured image it is giving 53% accuracy of same object mouse (Fig. 7), however this rate of accuracy is adequate for blind person
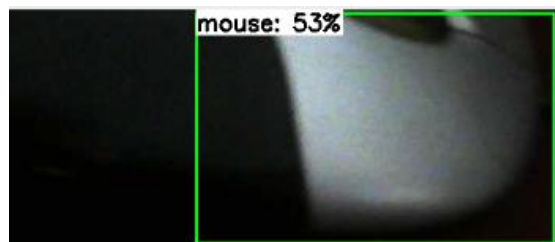

Fig. 7: "mouse detection (out of focus of light)"

This architecture gives good results for distant objects and number of objects

Fig. 8: "Number of objects detection"

In Fig. 8, system is showing correct results for object 10 to 12 feet away from camera as well as number of objects were detected correctly in same frame

3. Conclusion

This architecture is tested and it is reliable for visually impaired person. As it uses tensor flow lite response time is very less. It even consumes less power, so is desirable for portable application. There are only 80 objects in data set, all are basic object any individual may need in day to day life .most attractive feature of this project is as the output is audio form and visually impaired people can use earphones which can easily guide blind person thus making user self-reliant without having to ask for assistance .

4. Acknowledgements

5. References

[1]   . W. Chang, L. Chen, C. Hsu, J. Chen, T. Yang and C. Lin, "MedGlasses: A Wearable Smart-Glasses-Based Drug Pill Recognition System Using Deep Learning for Visually Impaired Chronic Patients," in IEEE Access, vol. 8, pp. 17013-17024, 2020.

[2]   V. Le, H. Vu and T. T. Nguyen, "A Frame-work assisting the Visually Impaired People: Common Object Detection and Pose Estimation in Surrounding Environment," 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, 2018, pp. 216-221.

[3]   S. Duman, A. Elewi and Z. Yetgin, "Design and Implementation of an Embedded Real-Time System for Guiding Visually Impaired Individuals," 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 2019, pp. 1-5.

[4]   L. Tepelea, I. Buciu, C. Grava, I. Gavrilut and A. Gacsádi, "A Vision Module for Visually Impaired People by Using Raspberry PI Platform," 2019 15th International Conference on Engineering of Modern Electric Systems (EMES), Oradea, Romania, 2019, pp. 209-212.

[5]   I. Joe Louis Paul, S. Sasirekha, S. Mohanavalli, C. Jayashree, P. MoohanaPriya and K. Monika, "Smart Eye for Visually Impaired-An aid to help the blind people," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-5.

[6]   P. Vyavahare and S. Habeeb, "Assistant for Visually Impaired using Computer Vision," 2018 1st International

Conference on Advanced Research in Engineering Sciences (ARES), Dubai, United Arab Emirates, 2018, pp. 1-7.

[7] O. Stephen, D. Mishra and M. Sain, "Real Time object detection and multilingual speech synthesis," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-3.

[8] M. A. Khan Shishir, S. Rashid Fahim, F. M. Habib and T. Farah, "Eye Assistant: Using mobile application to help the visually impaired," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, 2019, pp. 1-4.

[9] C. Jayawardena, B. K. Balasuriya, N. P. Lokuhettiarachchi and A. R. M. D. N. Ranasinghe, "Intelligent Platform for Visually Impaired Children for Learning Indoor and Outdoor Objects," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, pp. 2572-2577.

[10] S. Pehlivan, M. Unay and A. Akan, "Designing an Obstacle Detection and Alerting System for Visually Impaired People on Side.