

Navigation Aid for the Blind and the Visually Impaired People using eSpeak and Tensor Flow



Nishkala H M, Anu S H, Ashwini B V, Kavya C M, Monika B S

Abstract: Applications of science and technology have made a human life much easier. Vision plays a very important role in one's life. Disease, accidents or due some other reasons people may lose their vision. Navigation becomes a major problem for the people with complete blindness or partial blindness. This paper aims to provide navigation guidance for visually impaired. Here we have designed a model which provides the instruction for the visionless people to navigate freely. NoIR camera is used to capture the picture around the person and identifies the objects. Using earphones voice output is provided defining the objects. This model includes Raspberry Pi 3 processor which collects the objects in surroundings and converts them into voice message, NoIR camera is used to detect the object, power bank provides the power and earphones are used here to output the message. TensorFlow API an open source software library used for object detection and classification. Using TensorFlow API multiple objects are obtained in a single frame. eSpeak a Text to Speech synthesizer (TTS) software is used to convert text (detected objects) to speech format. Hence using NoIR camera video which is captured is converted into voice output which provides the guidance for detecting objects. Using COCO model 90 commonly used objects are identified like person, table, book etc.

Keywords: COCO model, eSpeak, NoIR camera, Raspberry Pi 3, TensorFlow API

I. INTRODUCTION

According to the human sense organs Vision plays a very vital role in the human's life. Some people face this vision problem by birth or due to some road incidents. In the day to day life the visually impaired people face many difficulties in the vision senses. For the visually impaired people it was hard to

confess the things in the everyday life. For this problem this model helps out the visually impaired people in the things identification and gives some hearing details about the things which are identified.

For real time object observation and recognition this smart cap helps the visually impaired people to guide alone themselves. The Raspberry Pi processor which is in the advanced system, the processor was filled accompanied by the instruction of Convolutional Neural Network model (CNN), which was advanced by make use of TensorFlow. The NoIR camera is connected by the processor and that processor was encrypted in Python. The image was captured by NoIR camera in live time and it provides the Raspberry Pi processor process the image. The COCO model is used to identify and differentiate the thing by the use of Python code. If the NoIR camera captures the image and that image was covered by squares, the camera detects the image with accuracy and name of the image. The text file which consists of the classification of the identical objects. The classification reference be made of class name and class id of the identified things. Text to Speech Synthesizer (TTS) is a software called eSpeak, which changes the text file to audio. This Raspberry Pi system was easily transportable and it can be comfortably carried by the blind people.

II. RELATED WORKS

Many models have been invited to guide the visually impaired people. Most related model of them is intelligent electronic eye for blind people [1]. Using image and obstacle sensors visual information about the surrounding is gathered. Required voice output is obtained by AVR microcontroller. By solar photo voltaic module, piezoelectric source and by body temperature electricity which is developed to provide electric power to this [2]. Here for object identification and distance calculation ATmega2560 based Arduino Mega 2560 is used. Traffic signals status can also be analyzed in this model. Computer vision guidance for indoor navigation of visionless people is explained in this paper [3]. Mobile apps and remote handling computer are used for indoor navigation to help the blind.

III. PROPOSED SYSTEM

This model works on the basis of TensorFlow and a software called eSpeak which convert text to speech format. We can arrange many classes which is occurred in an image and can also identify the location of an image with a single object detection model.

Manuscript received on February 10, 2020.

Revised Manuscript received on February 20, 2020.

Manuscript published on March 30, 2020.

* Correspondence Author

Nishkala H M, Assistant Professor, Department of CS&E, SJM Institute of Technology, Chitradurga, India. Email: nish07scorpio@gmail.com

Anu S H, CS&E, SJM Institute of technology, Chitradurga, India. Email: anuappu03@gmail.com

Ashwini B V, CS&E, SJM Institute of technology, Chitradurga, India. Email: bvashwini2006@gmail.com

Kavya C M, CS&E, SJM Institute of technology, Chitradurga, India. Email: kavyacmutter@gmail.com

Monika B S, CS&E, SJM Institute of technology, Chitradurga, India. Email: monikavyakunta30@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Table- I: Various dataset for Object identification

Model Name	Speed	Outputs
ssd_mobilenet_v1_coco	Fast	Boxes
ssd_inception_v2_coco	Fast	Boxes
rfcn_resnet101_coco	Medium	Boxes
faster_rcnn_resnet101_coco	Medium	Boxes
faster_rcnn_inception_resnet_v2_atrous_coco	Slow	Boxes

V. RESULTS AND ANALYSIS

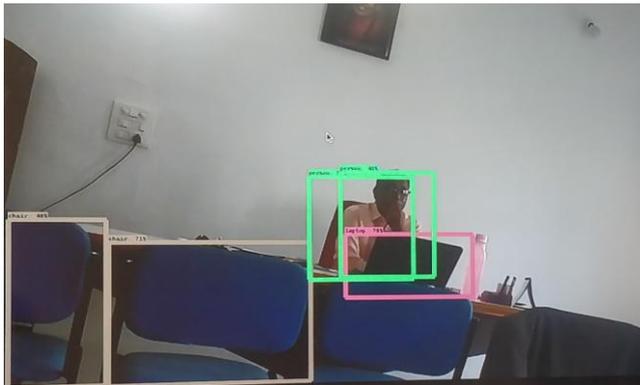


Fig. 4: Final output

The figure represents the detected objects with accuracy. In a single picture one or more objects can be identified like

person, laptop, chair as shown in Fig. 4. It is not able to identify the objects which are not included in the COCO dataset and also there occurs fault in object identification where the system misleads the object name. These are the major drawbacks of the model.

VI. CONCLUSION

This is a simple model in which Raspberry Pi converts the visually captured picture using NoIR camera into voice output. Blind people have to simply wear it without any particular skills of operating it he/she has to only supply power to the device. Proposed model is reasonable and customizable. It is real-time model that guides the blind by providing voice output about the surrounding environment helping to navigate more safely.

This simple architecture is very useful for navigation for the blind. Objects in the picture can be recognised by object detector. SSD mobilenet which is trained by COCO model can identify 90 different classes of objects. Objects can be increased by training the model by showing the image in different angles. As face detections is also included family members and friends can be easily identified by the blind people.

REFERENCES

1. M.Maiti,P.Mallick,M.Bagchi,A.Nayek,“Intelligent electronic eye for visually impaired people”,978-1-53862215-5/17/\$31.00 ©2017 IEEE.
2. Namita Agarwal, Anosh Iyer, Sonalakshi Naidu, Snedden Rodrigues, “Electronic Guidance System For The Visually Impaired -A Framework”, 2015 International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb. 04 – 06, 2015, Mumbai, India.
3. Kabalan Chaccour, Georges Badr, “Computer vision guidance system for indoor navigation of visually impaired people”, 2016 IEEE 8th International Conference on Intelligent Systems.
4. <https://www.tensorflow.org/>
5. Nitin R. Gavai, Yashashree A. Jakhade, Seema A. Tribhuvan, Rashmi Bhattad “MobileNets for Flower Classification using TensorFlow”, 2017 International Conference on Big Data, IoT and Data Science (BIGD), Vishwakarma Institute of Technology, Pune, Dec 2022, 2017.
6. Tsung-Yi Lin, Michael Maire, Serge, Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Doll ´ ar, “Microsoft COCO: Common Objects in Context”, arXiv:1405.0312v3 [cs.CV] 21 Feb 2015.
7. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C & Berg AC (2016) “Ssd: Single shot multibox detector. European conference on computer vision”, Springer: 21-37.
8. Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y & Guadarrama S (2016) Speed/accuracy trade-offs for modern convolutional object detectors. arXiv preprint arXiv:1611.10012.
9. Kruti Goyal, Kartikey Agarwal, Rishi Kumar “Face Detection and Tracking Using OpenCV”, International Conference on Electronics, Communication and Aerospace Technology ICECA 2017.
10. <http://espeak.sourceforge.net/>

AUTHORS PROFILE



Nishkala H M , Assistant Professor in the department of CS&E, SJMIT, Chitradurga. I have pursued my M.Tech in CS&E from VTU in the year 2016. Area of interest includes Machine Learning Techniques and IoT.



Anu S H, Under graduated student B.E, completing graduation in the year 2020, in the department of computer science and engineering, SJM Institute of Technology, Chitradurga, India.



Ashwini B V, Under graduated student B.E, completing graduation in the year 2020, in the department of computer science and engineering, SJM Institute of Technology, Chitradurga, India.



Kavya C M, Under graduated student B.E, completing graduation in the year 2020, in the department of computer science and engineering, SJM Institute of Technology, Chitradurga, India.



Monika B S, Under graduated student B.E, completing graduation in the year 2020, in the department of computer science and engineering, SJM Institute of Technology, Chitradurga, India.