International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 08 | Aug 2021 www.irjet.net p-ISSN: 2395-0072

The Digital Eye Android Application For Assisting Visually Impaired

Manasi Powar[1], Shivani Patil[2], Shwetali Nikam[3], Komal Kurundwade[4], Sujata.A.Pardeshi[5]

[1,2,3,4]Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Kolhapur, Maharashtra, India.
[5]Assistant Professor, Department of Computer Science and Engineering, Sanjay Ghodawat University, Kolhapur, Maharashtra, India.

Abstract - *In today's environment, independent living becomes necessary, but blind people are severely limited. Blind people are at a disadvantage because they need manual assistance to gain information about their surroundings. Visual information provides the basis for many tasks. Therefore, it is difficult for blind people to complete most of the daily chores. Advances in technology have made it possible to assist the visually impaired. Mobile devices, well known as portable smart devices, are now widely used in our daily life. These devices are very useful for assisting people with visual impairments and making their lives easier. This paper introduces an Android application that recognizes nearby objects and provides audio instruction to the user when those objects are recognized. The hearing ability of the user tries to fulfill the seeing ability.*

***Key Words*: MobileNet-SSD, Object Detection, Text-To-Speech, Image Processing**

1.INTRODUCTION

As of 2015 there were 940 million people with some degree of vision loss. 246 million had low vision and 39 million were blind.The majority of people with poor vision are in the developing world and are over the age of 50 years [12]. Among all the senses, the power to examine through the eyes is one in every of the foremost vital senses in human beings. and therefore this loss of ability seriously affects all possible movements that a person can try to make in his/her life. Traditionally, Tools like sticks were used for many years to avoid obstacles. Keeping pet dogs or guide dogs was one in every of the few techniques employed by the visually impaired. Some took the assistance of their family and friends for assistance. However, these techniques had drawbacks. Individuals with trained dogs required cash and time to feed and train the dogs. In today's technologically advanced world the on top of techniques cannot satisfy the wants of visually impaired people. Through the help of technology solutions will be created to rectify the issues visually impaired people face in day-after-day life. it might have a positive impact on their lives and create their day-to-day life easier. The projected system will find the objects within the user' surroundings by taking visual feed through a mobile camera. It can alert the user concerning the obstacles in his pathway through audio output and so helps the user to navigate from one place to a different saving him from tripping anywhere. The main aim of this project is offer a manner to visually impaired individuals to examine their surroundings. This goal is fulfilled by using technology which includes MobileNet SSD, Image Recognition, Object Detection, and TextToSpeech Conversion.

2. LITERATURE REVIEW

Research Advances of Indoor Navigation for Blind People: a brief Review of Technological Instrumentation[1]. This paper provides helpful insights regarding Electronic Traveling Aid (ETA) solutions for visually impaired individuals within the field of indoor orientation and navigation. The projected systems were classified into 3 groups in line with the technology used for navigation: vision-based, non-vision-based, and hybrid technologies. Vision-based systems use live camera feeds or photo images because the main supply of data representing the surroundings. Non-vision-based tools use varied sensors to sense the surroundings and represent it to the user. Hybrid systems mix vision-based approaches with sensor technologies. They were assigned to a separate class, though they might even be part attributed to the vision-based systems.

Real-Time Object Detection and tracking using Deep Learning and OpenCV [5]. In this paper, SSD and MobileNets based algorithms are enforced for detection and tracking during a python environment. Object detection involves detective work the region of interest of an object from a given category of image. This model showed glorious detection and tracking results on the item trained and may be additional utilised in specific eventualities to detect, track and reply to the actual targeted objects within the video surveillance.

Real-Time Object Detection And Identification For visually challenged people using Mobile Platform[6]. This paper presents a comprehensive and comparative analysis of the work that has been done in the sphere of object detection for visually challenged people. They divide the object detection algorithms into 2 categories, first, one class is region-based object detection algorithms and therefore the second category is regression-based object detection

algorithms. the most advantage of regression-based algorithms over region-based algorithms is that regression-based algorithms work quicker compared to region-based algorithms

VisualPal is a mobile app for object detection and recognition to aid the visually impaired [7]. It detects the course of most brightness and predominant shades withinside the image. It made use of Artificial Neural Network Technology together with Euclidean Distance measures together. It captured a video and categorised it into diverse frames. All frames are as compared with preceding frames and responses may be given primarily based totally on saved gadgets information.

3.METHODOLOGY

Dataset

COCO is a large-scale data set for recognition, segmentation and captioning [8]. COCO dataset features:

- Object segmentation
- Context recognition
- Superpixel segmentation
- 330K images (marking> 200K)
- 1.5 million object instances
- 80 object categories
- 91 content categories

The Microsoft Common Objects in Context (COCO)dataset is considerably larger within the range of instances per class than the PASCAL VOC and SUN datasets. MS COCO contains significantly additional object instances per image (7.7) as compared to ImageNet (3.0) and PASCAL (2.3)[9].

Algorithm

A. Mobilenet Algorithm:

The Mobilenet model relies on depthwise seperable convolutions that could be a type of factorized convolutions that factorize a standard convolution into a depthwise convolution and a 1x1 convolution known as a pointwise convolution. For MobileNets the depthwise convolution applies one filter to every input channel. The pointwise convolution then applies a 1x1 convolution to mix the outputs of the depthwise convolution. a standard convolution each filters and combines inputs into a replacement set of outputs in one step. The depthwise separable convolution splits this into 2 layers, a separate layer for filtering and a separate layer for combining. This factorization has the impact of drastically reducing computation and model size[2].

B. Single Shot Detector (SSD):

SSD is a popular object recognition algorithm developed by Google. It's supported the VGG16 architecture. Therefore, SSD is simpler and easier to implement. The SSD method is based on a feed forward convolutional network, which creates a set of bounding boxes of fixed size and an estimate of the existence of feature class instances in these boxes, followed by a non-maximum suppression step to obtain the final detection [3].

C. MobileNet-SSD:

To reduce the computational complexity of the VGG16-SSD detector, Google applied the Mobilenet network model to replace the VGG16 network, improving the real-time performance of the SSD detector. The existing Mobilenet-SSD network architecture uses the second-generation Mobilenet network, called Mobilenet-v2, as the backbone network model of the SSD detector. The Mobilenet-SSD detector inherits the design of VGG16-SSD that the front-end Mobilenet-v2 network provides six feature maps with different dimensions for the back-end detection network to perform multi-scale object detection. Since the backbone network model is changed from VGG-16 to Mobilenet-v2, the Mobilenet-SSD detector can achieve real-time performance and is faster than other existing object detection networks[4].

Voice Generation Module

Pyttsx3 plays an important role in the speech engine. Pyttsx3 is a Python conversion library that can convert text to speech. This library is suitable for Python 2 and 3. Get a reference to the pyttsx.Engine instance, which is a factory function. pyttsx.init() is called by the application. Pyttsx3 is a tool that can easily convert text to speech. Audio commands are generated as output [10].

PyTorch is mainly a machine learning library. PyTorch is mainly used in the audio field. PyTorch helps load audio files in standard MP3 format. It can also adjust the speed of measuring sound. Therefore it is used to control the properties of the sound. Such as frequency, wavelength and waveform. Many of the available sound synthesis options can also be checked using the Pytorch function [11].

4.RESULT ANALYSIS

Figure 1

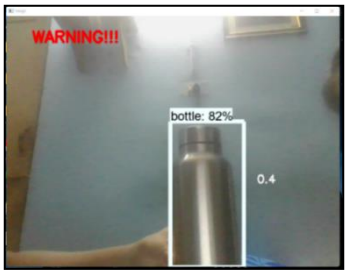


Figure 2

Figure 3

Figure 4

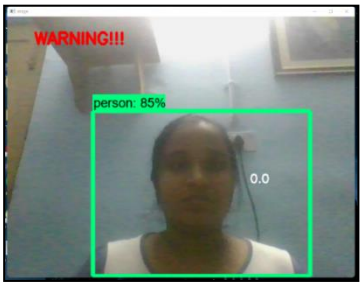
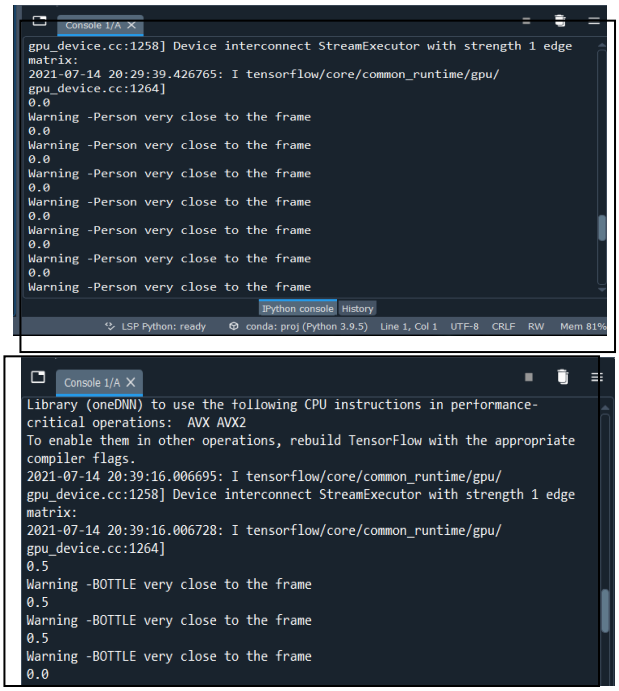


Figure 5

As shown in Figure 1, the accuracy of a bottle is 82% and the final distance is 0.2 units from frame and warning is generated for distance as it is too close and voice output heard that its class name is a bottle. It states (warning - bottle very close to the frame). As shown in figure 2, the accuracy of tv is 93% and no distance-based warning as it is at a safer distance rather class recognition voice is generated and the name of the object can be heard as it is a tv. As shown in Figure 3, the accuracy of remote is 94% and no distance-based warning is generated as it is at a safer distance and class recognition voice can be heard and the class is remote. Figure 4 shows the accuracy of a person is 85% and warning is generated because the distance is too close to the frame and voice output is generated that is a class name is a person. Figure 5 shows the accuracy of a cup is 96% and no distance-based warning is generated as it is at a safer distance and class recognition voice can be heard and the class is a cup.

Console View



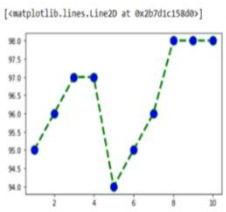
As shown in the figure above, the system calculates the distance and displays its output on the console, and blind people can hear these distance values and object-based warnings.

Performance Analysis

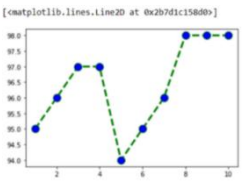
Graphs are premeditated performance metrics to accuracy. It is seen as follows:

- X-Axis: Time Frame(in Sec.)
- Y-AXIS: Accuracy Metrics(in percentage)

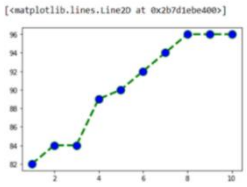
1.Cup



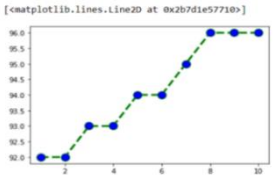
2.Remote



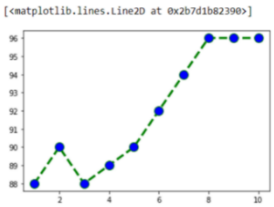
3.Bed



4.Chair



5.TV



The projected system with success detects ninety objects, labels them, and conjointly shows its accuracy. The model also calculates the distance from the thing to the camera and offers voice feedback. The dataset was tested on 2 totally different models, SSD MobileNet V1 and SSD inception V2. However, the SSD MobileNet V1 model showed less latency and was quicker in detective work objects

5.CONCLUSION

Most of the existing systems use additional hardware devices, adding a burden to the visually impaired. There is a need for such a system that may not seem like an added burden to them and that is also part of their lives. This paper presents a solution for the visually impaired that is affordable and easily accessible. The solution is an Android application for the visually impaired to detect objects with a MobileNet-SSD algorithm that helps them by informing them about their surrounding. After the object is recognized, the system informs the visually impaired by generating a voice. It solves the problem of keeping additional special devices.

REFERENCES

[1] D. Plikynas, A. Žvironas, M. Gudauskis, A. Budrionis, P. Daniušis and I. Sliesoraitytė, "Research advances of indoor navigation for blind people: A brief review of technological instrumentation," in IEEE Instrumentation & Measurement Magazine, vol. 23, no. 4, pp. 22-32, June 2020,

[2] Howard, Andrew G. et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." ArXiv abs/1704.04861 (2017): n. pag.

[3] Liu W. et al. (2016) SSD: Single Shot MultiBox Detector. In: Leibe B., Matas J., Sebe N., Welling M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9905. Springer, Cham.

[4] Y. -C. Chiu, C. -Y. Tsai, M. -D. Ruan, G. -Y. Shen and T. -T. Lee, "Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems," 2020 International Conference on System Science and Engineering (ICSSE), 2020.

[5] G. Chandan, A. Jain, H. Jain and Mohana, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018.

[6] Joshi, Neeraj et al. "Real-Time Object Detection And Identification For Visually Challenged People Using Mobile Platform." (2021).

[7] Shagufta Md. Rafique Bagwan, Prof. L.J. Sankpal," VisualPal: A Mobile App for Object Recognition for the Visually Impaired", IEEE International Conference on Computer, Communication, and Control (IC4-2015).

[8] COCO-Common Objects in Context https://cocodataset.org

[9] Lin, Tsung-Yi et al. "Microsoft COCO: Common Objects in Context." ECCV (2014).

[10] Pyttsx3.(2020,July6).PyPI.https://pypi.org/project/pyttsx/.

[11] PyTorch documentation — PyTorch 1.9.0 documentation.https://pytorch.org/docs/stable/index.html

[12] Visual Impairment - Wikipedia. (2015, May 23). Visual impairment Wikipedia.https://en.wikipedia.org/wiki/Visual_impairment.