International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 6 ǁ 2021 ǁ PP. 11-15

Virtual Sight Mobility for Visually Impaired

Bhagyashree Vidhate[1], Bhushan Bari[2], Utkarsha Kale[3], Vandan Vadnere[4], Dr.S.V. Gumaste[5]

[*1,2,3,4,5]*Department of Information Technology, MET Bhujbal Knowledge City, Institute of Engineering, Nashik, Maharashtra, India*

Corresponding Author:Bhagyashree Vidhate

Abstract

The most important human sense is vision. The human perception of the environment is due to the images captured by the human eye. However, for the visually impaired people, there are several devices that have been designed to help them cope with their disabilities. The authors with the help of the emerging technologies like Digital Image Processing and Deep Learning have developed an eyewear. These intriguing technologies are helping visually impaired people grow individually. Digital image processing technique which processes the images with the help of a digital computer is used to process the images captured. Thereafter, the deep learning algorithm, yolo object detection is used to classify the images according to classes and audio signals are passed to the visually impaired person to make them aware about the obstacle in front. Therefore, this paper provides the analysis of the implementation.

Keywords:*Object Detection, Visual Substitution, Feature Extraction, Tensor flow.*

Date of Submission: 14-06-2021 Date of acceptance: 17-06-2021

I. INTRODUCTION

According to WHO, there are almost 1 billion people with vision impairment. The major factors that cause vision impairment and blindness are uncorrected refractive errors and opaqueness [1].

The life of a visually impaired person has always been dependent on others. The advancement in technologies have made this possible to overcome this type of impairment. These will result in replacing the traditional technologies like White Cane, helper support and taking help of dogs. Also, it is not easy for visually impaired to understand the visual aspect like depth, color, position and orientation of an object. There are numerous assistive devices designed to support visually impaired some of them are based on sensors which alert the user using beeping sound and some of those also include computer vision techniques, but are efficient that the blind can count on.

The survey conducted by the authors also mentions about the various vision substitution techniques which are used for the implementation of the proposed system. The vision substitution technique is the technique which uses the hearing aid to provide the output for the captured images from the surrounding. Therefore, the authors focus on the vision substitution methods for the visually impaired. The vision substitution methods can be further subdivided into the following [2]:

1. Electronic Travel Aids (ETA): This is the technique that uses different forms of information and gives the output with the help of sensors or hearing aids for improving the travelling experience of the visually impaired.
2. Electronic Orientation Aids (EOA): This technique refers to the guidance provided by the system at every stage of mobility by the visually impaired user.
3. Position Locator Devices (PLD): The technique includes the use of technology that helps to navigate or can be used to track the location of the visually impaired user.

As the authors focus on the object detection criteria for providing efficiency to the user, ETA technique is used for better mobility of the blind user.In this project, we tried to resolve the issue by using deep learning techniques to detect an object after identifying it. A large dataset consisting of objects such as chairs, doors, tables, bicycles, etc. from daily scenes is generated with the help of a web camera to apply the object detection and recognition technique. Google Object Detection API (Application Program Interface) based on TF(TensorFlow) is used to identify different obstacles [7,8]. The API consists of several models of object detection, but we have used YOLO, which is trained on COCO datasets. By using this model, we did not only identify the type of object but also inform the user on which direction that object is present with the help of a voice [9].

II. LITERATURE SURVEY

The survey conducted on various techniques present is mentioned below.

A. A Comparison of SIFT, PCA-SIFT and SURF:

In this paper authors had done comparison of Scale Invariant Feature Transform (SIFT), Principal Component Analysis (PCA-SIFT), Speeded Up Robust Features (SURF) and outlined the feature spotting strategies. Also, authors compared strategy with scale variation, accuracy in match findings, the necessitated repository space and pivotal point. They found that SIFT needs farther memory space compared to the other two strategies and the matching speed for SIFT is less although it gives good match findings which improves the perfection to detect objects. Furthermore, SIFT is slow and not good at illumination changes, while it's undeviating to rotation, scale changes and affine transformations. Whereas PCA-SIFT show its advantages in rotation and illumination changes and gives lower matches compared to SIFT [3].

B. Visual substitution system for blind people based on SIFT description:

This literature acquainted the nonidentical video substitution systems vOICe, PSVA, Vibe. Authors unravelled on the well and rattling algorithms to recognize and detect objects in images. They had enforced feature - based extraction instead of using raw pixel valuations to frame the faster processing and to reduce in - class variability to the raw input data. The coming work consists on real time detection. In order to meliorate and pump up the object recognition rate. It'll reckon the colour while correlating things and append an auditory translation for correlated objects [4]

C. Object Detection and Identification for Blind People in Video Scene:

In this proposed system, system restores a central event of the visual system which is the identification of surrounding things. This style is based on the local feature extraction concept. The simulation outcomes exploiting SIFT algorithm and crucial points corresponding showed rational preciseness for detecting objects. They held solved for the key point detection in rapid tape recording using affine conversion in SIFT which is invariant to the changes in luminance. In coming works, it'll rate this relationship for background understanding or detecting everything that belongs to a given place or position [5].

III. ARCHITECTURE

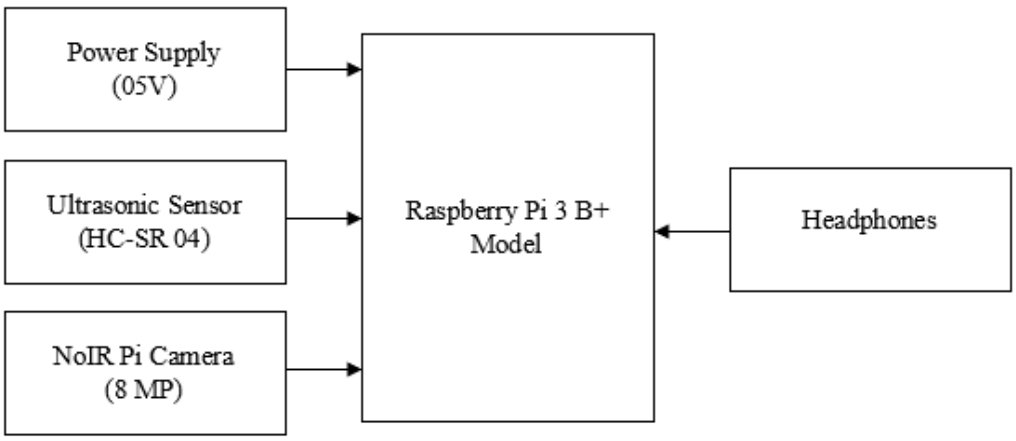

Figure 1**:** System Architecture

The proposed system is Deep learning-based blind aid for the object recognition system which is a very helpful product for Visually Impaired people.

The main purpose of the system is to detect the objects and covert the output into audio format. So, authors have considered five main components to design proposed system such as Raspberry pi 3 B+ Module, Power supply (Power Bank 16000MW), Ultrasonic Sensor (HC-SR04), NoIR Pi camera and Headphones where the controlling of kernel and processing is performed by Raspberry Pi 3 B+ model. The system is developed using Python programming language and OpenCV library and ported on Raspberry Pi3 model B+ platform.

In the above figure, as soon as the power supply is provided, the Raspberry Pi 3 module execute the code and then NoIR camera will start capturing image frames. After that the captured frame will be stored in the

temporary memory for detection purpose. Then the algorithm with the pre-trained model is applied on that captured frames for object detection. The output received will be converted into Speech format using the google text-to-speech API (gtts), which is audible to the blind user through earphones.

Figure 2: Actual Developed Model

IV. METHODOLOGIES

The proposed system uses different python libraries to get an accurate output. The libraries used for implementation are OpenCV, gTTS, pydub.

1. OpenCV:
   OpenCV is a python library which is widely used for Computer Vision. With the help of OpenCV developers can develop real-time computer vision. This library mainly focuses on image processing, analysis including features like object detection.
2. Google Text-to-Speech:
   gTTS is also known as Google Text-to-Speech. gTTS is also a python library and CLI tool to interface with Google translator's text-to-speech API. In the proposed system, this library is used for converting output in speech.
3. Pydub:
   The Pydub library is used for playing, splitting and merging for the given ".wav" file.
4. YOLO:
   YOLO (You Only Look Once) real-time object detection algorithm, which uses COCO (Common Object in Context) data set. YOLO algorithm employs CNN (Convolutional Neural Network) and COCO is large scale data set with most common objects which appear in day-today life. Also, COCO provides object segmentation and captioning dataset. The YOLO object detector is trained on COCO dataset.

In the proposed system, "coco names" file provided to algorithm with most common object names such as a person, ball, chair, etc. And the file is trained to identify objects. Simultaneously, YOLOv3 weights and configuration file loaded with the help of DNN module of OpenCV. For accurate prediction of objects with deep neural network, model need to pre-process the data. Cv2.dnn provides two functions i.e blobFromImage and blobFromImages. These two functions perform scaling, mean subtraction and channel swap. blobFromImage will be used in function detect_objects() that accepts image/frame from video and output as parameters. All functions such as cv2.dnn.NMSBoxes and other pre-processing functions will pipelined together in another function ImageDetect() for detecting objects in an image file.

V. RESULT AND ANALYSIS

The developed system is an eyewear which will provide a clear vision of the objects in front of it. Figure 3 shows the developed model.

Figure 3: A person wearing the eyewear

The system is tested and gives an accuracy of 96.1%. The objects tested were a chair, TV, book, bottle, person, ball, trees and car. The output is generated in the mp3 format and is correctly given through earphones. The authors considered a set of 50 images for a particular object, to classify them appropriately. The output percentage for every object that is detected is given in the table 1 below:

Sr. No	**Name of Objects Considered**	**Results**		
		No. of images Taken	**No. of images Correctly Classified**	**Accuracy**
1	Chair	50	47	95%
2	TV	50	49	98%
3	Book	50	49	98%
4	Bottle	50	47	95%
5	Person	50	49	99%
6	Ball	50	46	93%
7	Mouse	50	47	95%
8	Keyboard	50	47	95%
9	Trees	50	47	95%
10	Car	50	48	97%
Total		**500**	**476**	**96.1%**

Table 1: Correctly Classified Objects

Figure 4 (a): Mouse Detected

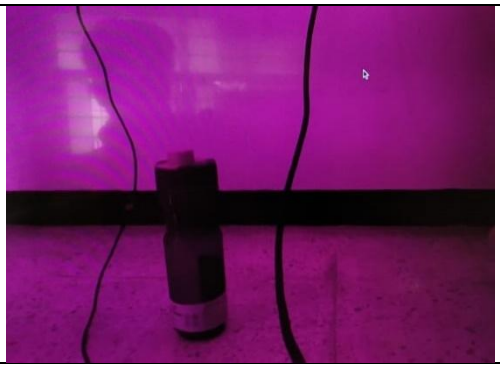

Figure 4 (b): Bottle Detected

Figure 4 (c): Keyboard Detected

Figure 4 (d): Ball Detected

Figure 4: Detected Objects

The Figure 4 shows the different objects that are detected and tested by the authors. Figure 4(a) shows the detection of a mouse and is gives output as "Mouse Detected". Figure 4(b) indicates the detection of a bottle being detected and is heard as "Bottle Detected". Also, Figure 4(c) depicts the detection of a keyboard and the output generated is "Keyboard Detected". In the same way, Figure 4(d) shows the detection of a ball and gives the output as "Ball Detected".

VI. CONCLUSION

The authors have successfully implemented an eyewear system based on all the necessary features with the help of Raspberry pi and Tensor flow framework. The system is tested on various objects like a chair, TV, sports ball, book etc. The system is also capable of detecting the person in front of the visually impaired. These objects are detected in the range of 200cm. The features that have been considered are day and night vision, indoor and outdoor obstacles, static object detection and text to speech ability.

The future enhancements can include the dynamic object detection, using a stronger processing unit and con0073idering a larger data set for object detection. In addition, considering a fully dark environment can be advantageous.

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment.

[2] Bhagyashree Vidhate, Bhushan Bari, Utkarsha Kale, Vandan Vadnere, "Virtual Sight Mobility for Visually Impaired-A Survey", The Journal of Oriental Research Madras, Volume: XCII-V, ISSN: 0022-3301, March 2021.

[3] Juan and O. Gwon, -A Comparison of SIFT, PCASIFT and SURFˆa. InternationalJournal of Image Processing (IJIP), 3(4):143ˆa152, 2009.

[4] Hanen Jabnoun, Faouzi Benzarti,Hamid Amiri, "Visual substitution system forblind people based on SIFT description", International Conference of Soft Computingand Pattern Recognition 2014 IEEE.

[5] Hanen Jabnoun, Faouzi Benzartiand Hamid Amiri, "Object recognition for blindpeople based on features extraction", IEEE14: INTERNATIONAL IMAGEPROCESSING APPLICATIONSAND SYSTEMS CONFERENCE 2014.

[6] History of Face RecognitionFacial recognition software. (2019, May 9). RetrievedDecember 3, 2019, from https://www.facefirst.com/blog .

[7] Sumitra A.Jakhete, Avanti Dorle,Piyush Pimplikar,Pranit Bagmar, AtharvaRajurkar, 2019, "Object Recognition App for Visually Impaired", IEEE Pune SectionInternational Conference (PuneConference), DO1:10.1109/PuneCon46936.2019.9105670

[8] Akhilesh Krishnan, Deepakraj G, Nishanth N, Dr.K.M.Anandkumar, 2016. "AutonomousWalking Stick for The Blind Using Echolocation and Image Processing", IEEE, DOI:10.1109/IC3I.2016.7917927

[9] Jawaid Nasreen, Warsi Arif, Asad Ali Shaikh, Yahya Muhammad, Monaisha Abdullah, (2019). "Object Detection and Narrator for Visually Impaired People. IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), DOI: 10.1109/ICETAS48360.2019.9117405.