Vision Maker: An Audio Visual And Navigation Aid For Visually Impaired Person

Sagor Saha
Electrical and Electronic Engineering
American International University Bangladesh
Dhaka,Bangladesh
sagarsaha455@gmail.com

Farhan Hossain Shakal
Electrical and Electronic Engineering
American International University Bangladesh
Dhaka,Bangladesh
shakalfarhan@gmail.com

Ahmed Mortuza Saleque
Assistant Professor
American International University Bangladesh
Dhaka,Bangladesh
ahmed.saikot@gmail.com

Jerin Jahan Trisha
Ekectrical and Electronic Engineering
American International University Bangladesh
Dhaka,Bangladesh
zaynahzerin@gmail.com

***Abstract*— People with low vision or complete loss of vision face challenging task to meet their daily demand. The barrier low vision hinders them participating in the society. The evolution of computer vision, artificial intelligence and machine learning proved to effective tool in revitalizing the situation of blind. The propounded design represents implementation of an assistive device that can aid them in recognizing the object. The low-cost device runs a pre-trained model (ssdlite_mobilenet_v2_coco) and can identify up to 80 classes. The user can read any text (English) from images. Ultrasonic sensors have been used in the custom-made device to continuously alert the user of obstacles from all directions. Additionally, it can help navigate them in outdoor using Google map API. The user can read news, listen to music and mail to the member of choice. The command is passed and received through earphone, which is connected in the audio jack of raspberry pi.**

***Keywords*— Object recognition, Human–computer interaction, outdoor navigation, computer vision, wearable device.**

I. INTRODUCTION

Human eyes perform the rudimentary role in obtaining visual information of the surrounding. Lack of vision creates difficult situation for visually impaired people. The primary causes of vision impairment are mainly glaucoma (2%), un-operated cataract (33%) and uncorrected refractive errors: hyperopia, myopia or astigmatism (43%). Globally around 80% of all visual disabilities can be prevented or cured [1]. According to World Health Organization (WHO), there are 285 million of visually impaired people worldwide. Among them, 246 million people are suffering from low or poor vision and 39 million people are completely blind [2]. The number is expected to reach double by 2020 [2]. People with sensory disabilities affect 5.3% of the world population for audition disability and 9.3% for vision loss, as stated by (WHO) [3]. Survey of data from 188 countries prescribes that there are no less than 200 million people inclined to severe vision impairment. However, the figure is expected to rise up to more than 550 million by year 2050 [4]. To ease their lives, effective solution to few problems should be considered:

1.) Is there any obstacle ahead of the person? This is a vital issue as it concerns the safety of the person.
2.) What objects are being faced by the user? This includes recognition of automobiles, household items and utilities.
3.) Where does the person want to go? In order to reach destination of the user, the Google map API can guide them.

A lot of research is still being continued regarding the visual and blind vision. In terms of feature and approach few papers have been critically analyzed. K. Patil et al [1] proposed a system, that is to be worn as shoe by the visually impaired people. The shoe contains ultrasonic sensor in all sides, vibration sensor, liquid detector and step-down sensor. A.kumar et al [4] devised a blind navigation system using artificial intelligence. The object detected is converted into text and fed to ear. K.R et al [5] proposed a smart device for visually impaired people. An android application was developed to provide outdoor navigation. Speech recognition engine has been employed to collect weather report, news and MTC website. M.Maiti et al [6] designed an intelligent electronic eye. The proposed device implements range finder and CCD camera module for obstacle. Piezo-electric device and solar panel were used to charge up the device. V.K et al [7] proposed a self-assistive system using IOT for blind and deaf person. Request protocol and Google speech API were implemented for converting speech to text and establish connection. A proto-type design containing Deep learning based object recognition was proposed in the paper along with the obstacle avoidance [8,9]. But none of these implementations demonstrate the continuous obstacle warning, object recognition, text identification with optical character recognition. Hence, we designed and implemented a device to overcome these issues as well as the implemented device which is capable of providing outdoor navigation, music system and can get updated to latest news. The wearable device is shaped like spectacles with acrylic material. Raspberry pi 3 is used as a processor with 720 pixels Logitech camera to run the device. To power up the system a

978-1-7281-4097-1/20/$31.00 ©2020 IEEE

3 celled polymer battery of 2200 mAh is used after converting the voltage using buck converter.

II. Design and Implementation

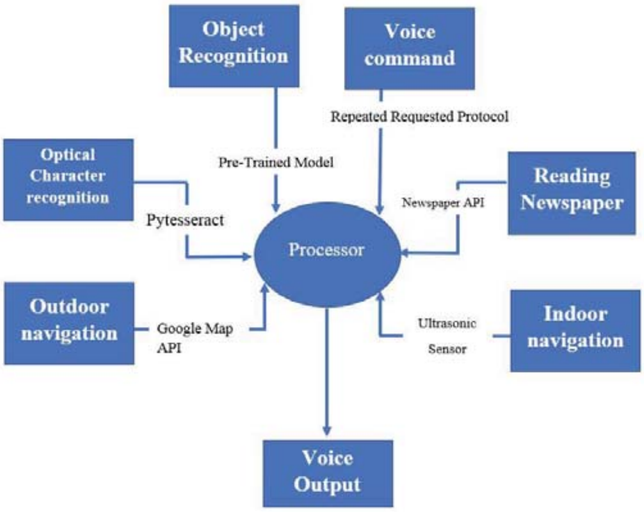

Fig. 1. System block diagram

The device runs on voice command and obstacle sensing sensor is always kept activated so that user could be notified of the surrounding obstacle. The device requires a constant internet connection for accessing all the features. Fig 1. Shows a system block diagram of the device. As an input, voice is passed through the microphone of the earphone. Camera connected to the processor captures the pictures and process it according to the command. The output result (text) is further converted to audio and fed to user via earphone. Implementation of the proposed device is shown in Fig. 2

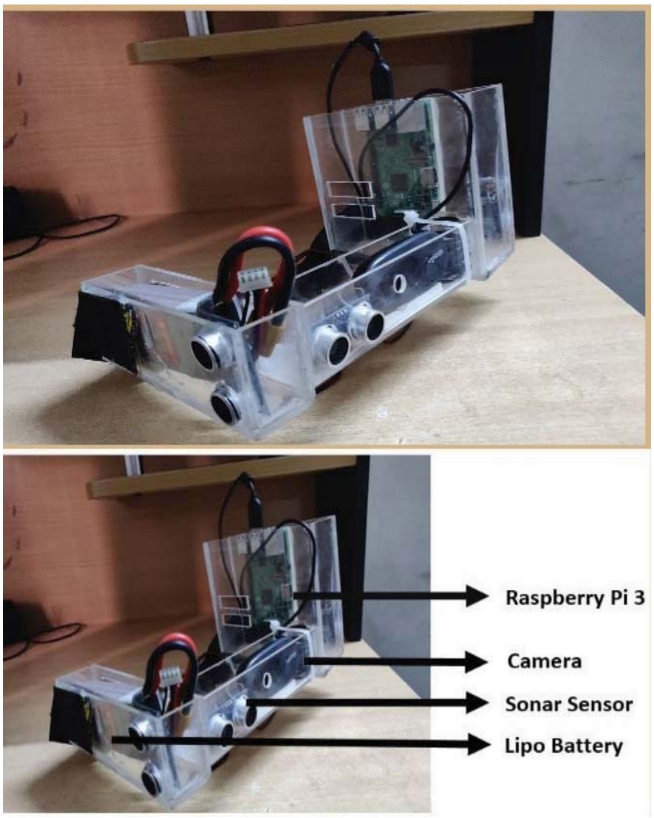

Fig. 2. Hardware implementation for the design

The electrical setup is shown in the Fig. 3. The propounded design required the following components:

A. Hardware

1) Raspberry pi 3: The single board, raspberry pi was developed in United Kingdom. This board uses a Broadcom BCM2837 SoC with 1.2 GHz 64 quad-core ARM Cortex-A53 processor along with 512 KB shared L2 cache. This model is equipped with 1 GB RAM, Bluetooth 4.1 (24 Mbit/s) and 2.4 GHz Wi-Fi 802.11n (150 Mbit/s). It also has 3mm jack for audio and a HDMI port.

2) Camera: As a camera Logitech C270 was picked because its low price and has video resolution of 1280x720 and image resolution of 640x480. It has a fixed type of focus and standard lens technology along with built in microphone.

3) Lithium polymer battery with Buck converter: It is a rechargeable battery. For supplying the input power, a three celled lithium polymer battery of 2200mAh was used. A total of 12V was converted to 5V(volt) and 2A(Ampere).

4) Ultrasonic Sensor: The ultrasonic sensor has a transmitter and receiver. The duration of reflected wave is measured and the distance is calculated. The operating voltage and current are 5V and 15 mA. The modulation of wave frequency is 40Hz and bears the capacity to measure the distance of 2cm to 400cm.

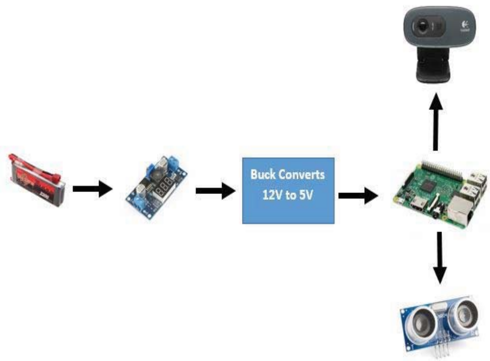

Fig. 3. Electrical setup of the device

B. Software

The project was run on python platform with Tensorflow object detection API. As a model we have selected ssdlite_mobilenet_v2_coco for object recognition. For extracting the text from images google open source tesseract optical character recognition (OCR) is used.

1. Object recognition: For object recognition pre-trained ssdlite_mobilenet_v2_coco was picked because it can maintain the balance between speed and accuracy. Among the available model in table I it is found to run on highest speed of 27ms. Moreover, being lightweight it best suits the device. The pre-trained model can classify up to 80 classes starting from households' items to automobile items. Pre-trained ssdlite_mobilenet_v2_coco: The SSD architecture is famous convolutional neural network (CNN) model because it has two components. One is the bounding box predictor and

other is feature extractor. The base network, feature extractor is generally a truncated classification network of VGG-16. The bounding box predictor is a combination of small convolutional filters used to predict the category score and box of set for default bounding boxes. This model is a mobile friendly variant of Single shot detector where regular convolution in the bounding boxes predictor are supplanted by depth wise convolutions [10].

TABLE I. COMMON OBJECT IN CONTEXT (COCO) TRAINED MODELS

Model name	Speed (ms)	COCO mAP[^1]
ssd_mobilenet_v1_0.75_depth_coco	26	18
ssd_mobilenet_v1_coco	30	21
ssd_mobilenet_v1_ppn_coco	26	20
ssd_mobilenet_v2_coco	31	22
ssd_mobilenet_v1_fpn_coco	56	32
ssd_inception_v2_coco	42	24
ssdlite_mobilenet_v2_coco	27	22

mAP= mean Average Precision

2. Optical character recognition: OCR allows the user to read any text written in an image. The open source software named Tesseract is an OCR engine which has the capability to recognize more than 100 languages out of the box. It is highly use in Gmail spam detection.
3. Virtual communication: the command passed in the receiver is transferred to google API server, which transforms voice command into text. Here the voice command is converted into electrical signal. With the help of request procedure protocol, the electrical signal enters the server and the encoded Mp3 audio is then sent to the google API for getting converted into text. Using the same request procedure protocol, the text again goes back to the processor as shown in Fig 4.

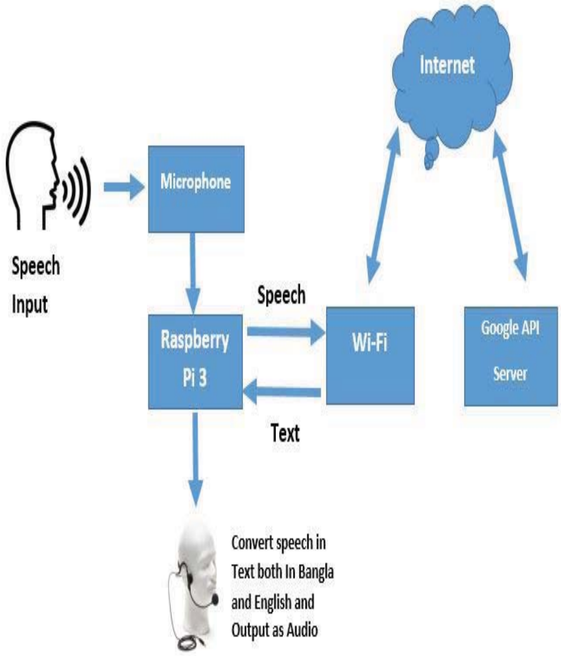

Fig. 4. Command signal processing

4. Indoor path navigation using ultrasonic sensor: For indoor path navigation ultrasonic sensor is used. Three ultrasonic sensors were placed in the device. Two sensors were placed in the left and right direction and another one is attached in the center looking forward. The device gives a voice feed for the direction if there is any obstacle within 10cm. table II shows the cases where some logic was created for the indoor path navigation. For example, if there is any obstacle in left and right direction the device through voice feed will guide the user in forward direction as we can see in case 1 from table II.

TABLE II. DESIGNED METHODOLOGY FOR INDOOR NAVIGATION

Condition	Left Sonar	Right Sonar	Forward Sonar	Direction
Case 1	1	1	0	Forward
Case 2	1	0	1	Right
Case 3	0	1	1	Left
Case 4	1	0	0	Right or Forward
Case 5	0	1	0	Left or forward
Case 6	0	0	1	Left or Right
Case 7	1	1	1	Back

1= obstacle, 0= No obstacle

III. PROJECT IMPLEMENTATION AND COST ANALYSIS

The custom-made device with processor as raspberry pi receives commands from the user and turns on the features accordingly. If no command is passed the ultrasonic sensor placed in three different location keeps warning the user about the obstacle. For others features like mailing, music system, outdoor navigation and reading newspaper the device requires a constant internet connection. The flow chart in Fig. 5 demonstrates the implementation of the device on human and Fig. 6 represents the total working process of the device.

Fig. 5. Implementation of the device on human.

The components altogether cost $96. Although expensive ways are available for those who can avail them, but the poor people below poverty line, especially in Bangladesh may not afford to use expensive devices. With these features, people can afford to buy the device and achieve a pictorial view. The results for each of the features is given in section IV.

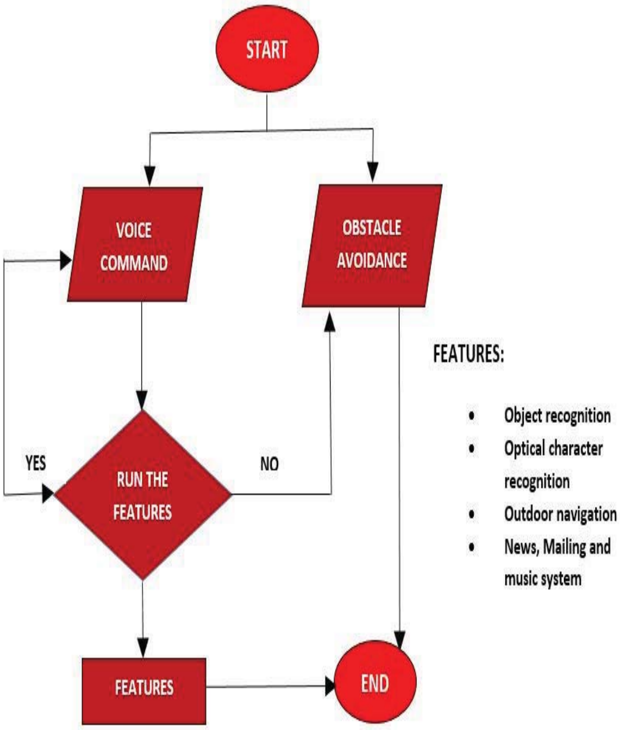

Fig. 6. Flow chart of the system represents the working process of the device.

TABLE III. TABULATION OF COST OF COMPONENTS USED

SI no.	Component name	Quantity	Cost in USD
1.	Raspberry pi	1	41.22
2.	Logitech Camera	1	23.56
3.	Ultrasonic sensor	3	2.47
4.	Buck converter	1	1.77
5.	Lithium polymer battery(2200mAh)	1	21.20
6.	Wires	18	0.12
7.	Acrylic board		5.89
Total			96.23

IV. Result and Analysis

The results of all the features are given with critical analysis. Object recognition using pre-trained model can identify object with minute details. There is always promising result when the text written in any image is captured and converted to speech. The ultrasonic sensor placed in three different position provide wide knowledge of the surrounding. This feature when tested provides excellent idea about indoor navigation. For oudoor navigation, Google Map was used and activated by voice commands with a certain location. The user get updated to latest news and play a song of his/her choice with the command.

1. *Result from the object recognition section*

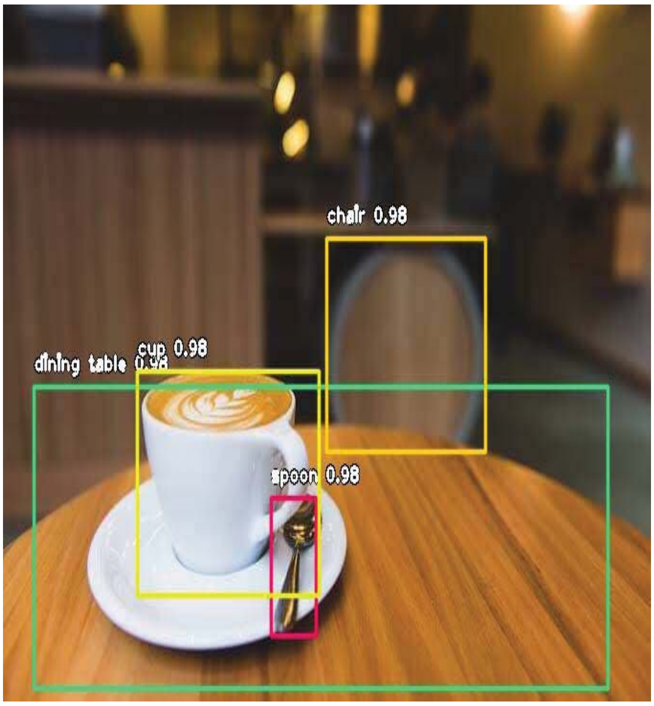

Fig. 7. Object recognition with few items.

Fig. 7 and 8. Represents the recognition of objects with few and many objects respectively. This feature is run by by the voice command. The objects recognized are converted into speech and warned to user by earphone. The pre-trained model implemented in the device takes 3 to 4 seconds to supply the input image captured and pass through the neural network for recognition. Whereas the other models listed Table I. consumes much time compared to ssdlite_mobilenet_v2_coco.

Fig. 8. Object recognition with minute details.

2. *Result from Optical Character recognition*

Pytesseract, an open source software was installed to read text from images. This software works well on computerized images. Fig. 9 and 10 represent detection from typed image and casual image respectively. Result from the both the section is satisfactory.

```
In [5]: runfile('C:/Users/ASUS/Desktop/TensorFlow-Tutori
Tutorials-master')
Hello Sagor,

There is a conference on blind education. | am
expecting your presence.

From,

Omar Hasan Al Boshir
```

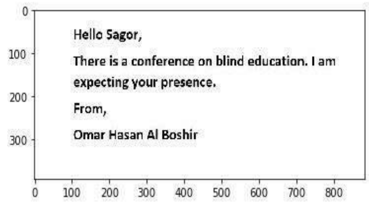

Fig. 9. Recognition from computer typed images.

```
In [6]: runfile('C:/Users/ASUS/Desktop/TensorFlow-Tutorials-ma
Desktop/TensorFlow-Tutorials-master')
be

The future belongs to those who
believe in the beauty of their dreams."
```

Fig. 10. Recognition from casual images.

3. *Result from newspaper and Google Map*

```
Listening...
Recognizing...
User: read newspaper

England project not finished yet - Jones on new contract
Premier League: 30 April restart date to be pushed back
Coronavirus: Premier League players must take a pay cut - Matt Hancock
The sights, sounds and stars we will miss from tennis in 2020
Eddie Jones: England head coach is the total package, says RFU's Bill Sweeney
Club Bruges to be named champions as Belgian Pro League cancelled
Coronavirus: How drivers, engineers and cycling kit makers are helping
Fisticuffs over free-kicks and bust-ups at yoga - when team-mates clashed
BBC to air memorable sporting highlights including London 2012
Phil McNulty's top 10 football matches
```

Fig. 11. Result from current newspaper.

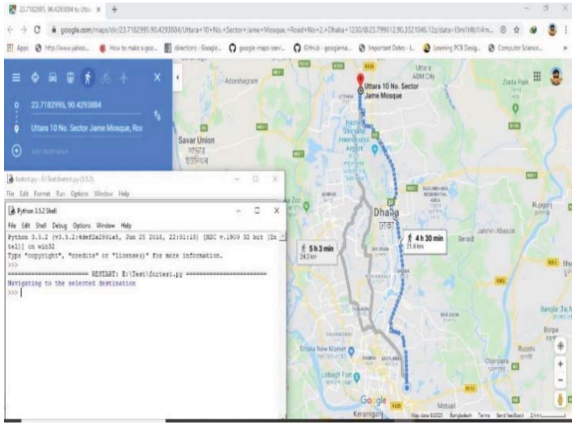

Fig. 12. Result of google Map guiding the user.

Fig.11 demonstrates current news headlines which read out for the user from News API. Fig.12 portrays the guidelines to the user as run Google Map. Specific command along with destination has to be passed as an input to the system.

4. *Indoor path navigation using ultrasonic sensor*

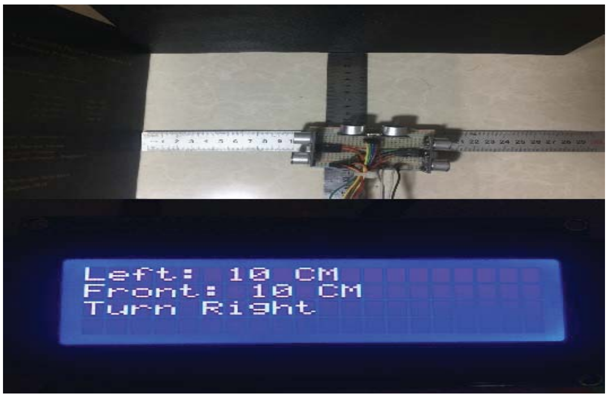

Fig. 13. Display showing direction to move Right.

Fig. 14. Display showing direction to move left or right

V. Conclusion

The device was successfully executed and results were promising. The cost of the device was reduced with better features when compared to other solutions. The propounded device fared well due to the features, accuracy and mobility.

The combination of these features were not presented in any other work. The recognitions were almost accurate. With constant Internet connection, the Human-computer interaction works smoothly along with warning of obstacle from all sides. The mechanical structure makes device a bit bulky. These could be optimized by designing a 3d printed body with light weight material. The device with all features can be expected to uplift the position of visually impaired people in society. People with low vision can successfully move around making use of the sense of speech, sense of hearing and the proposed device jointly. The problem of indoor and outdoor navigation can be served with this device.

REFERENCES

[1] T. V. Mataró *et al*., "An assistive mobile system supporting blind and visual impaired people when are outdoor," *3rd International Forum on Research and Technologies for Society and Industry (RTSI)*, Modena, 2017, pp. 1-6.

[2] K. Patil, Q. Jawadwala and F. C. Shu, "Design and Construction of Electronic Aid for Visually Impaired People," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 2, pp. 172-182, April 2018.

[3] Francesca Sorgini, Renato Caliò, Maria Chiara Carrozza & Calogero Maria Oddo “Haptic-assistive technologies for audition and vision sensory disabilities”, *Disability and Rehabilitation: Assistive Technology*, vol. 13, no. 4, pp 394-421, Oct. 2017. Accessed on: Feb. 9, 2020. [Online]. Available doi: 10.1080/17483107.2017.1385100.

[4] Ashwani Kumar, Ankush Chourasia, “Blind Navigation System Using Artificial Intelligence*” International research journal of engineering and technology (IRJET)*, India,2018, Vol-5.

[5] R. Kasthuri, B. Nivetha, S. Shabana, M. Veluchamy and S. Sivakumar, "Smart device for visually impaired people,"*Third International Conference on Science Technology Engineering & Management (ICONSTEM)*, Chennai, 2017, pp. 54-59.

[6] M. Maiti, P. Mallick, M. Bagchi, A. Nayek, T. K. Rana and S. Pramanik, "Intelligent electronic eye for visually impaired people," *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, Bangkok, 2017, pp. 39-42.

[7] K.Vasanth Macharla Mounika & R.Varatharajan,“A Self Assistive Device for Deaf & Blind People Using IOT,” *Journal of Medical Systems*, pp.1-8, Mar 2019.

[8] Mrs.R.Rajalakshmi, Ms.K.Vishnupriya, Ms.M.S.Sathyapriya, Ms.G.R Vishvaardhini “Smart Navigation System for thye Visually Impaired using Tensorflow,” *International journal of Advance Researchand Inovative Ideas in Education*,Vol 4, pp.1-14,.Feb 2018

[9] Kim, Bongjae & Seo, Hyeontae & Kim, Jeong-Dong.“Design and Implementation of a Wearable Device for the Blind by Using Deep Learning Based Object Recognition*,”International Conference on Computer Science and its application*, Singapore ,2017, pp.1008-1013

[10] H.Fan et al., “A Real-Time Object Detection Accelerator with Compressed SSDLite on FPGA*,” International Conference on Field-Programmable Technology(FPT)*, Naha, Okinawa,Japan, 2018,pp.14-21.