Visual, navigation and communication aid for visually impaired person

Sagor Saha, Farhan Hossain Shakal, Mufrath Mahmood
Department of Electrical and Electronics Engineering, American International University, Bangladesh

Article Info	ABSTRACT
	The loss of vision restrained the visually impaired people from performing their daily task. This issue has impeded their free-movement and turned them into dependent a person. People in this sector did not face technologies revamping their situations. With the advent of computer vision, artificial intelligence, the situation improved to a great extent. The propounded design is an implementation of a wearable device which is capable of performing a lot of features. It is employed to provide visual instinct by recognizing objects, identifying the face of choices. The device runs a pre-trained model to classify common objects from household items to automobiles items. Optical character recognition and Google translate were executed to read any text from image and convert speech of the user to text respectively. Besides, the user can search for an interesting topic by the command in the form of speech. Additionally, ultrasonic sensors were kept fixed at three positions to sense the obstacle during navigation. The display attached help in communication with deaf person and GPS and GSM module aid in tracing the user. All these features run by voice commands which are passed through the microphone of any earphone. The visual input is received through the camera and the computation task is processed in the raspberry pi board. However, the device seemed to be effective during the test and validation. *This is an open access article under the <u>CC BY-SA</u> license.*

Corresponding Author:

Sagor Saha
Department of Electrical and Electronics Engineering
American International Univerity Bangladesh
408/1 Kuratoli Khilkhet, Dhaka-1229, Bangladesh
Email: sagarsaha455@gmail.com

1. INTRODUCTION

The only organ which reacts to light and permits vision is the human eye. The human eye plays a major in obtaining visual information. Lack of sight may jeopardize the visually impaired people to get the task completed. There are 285 million visually impaired people globally according to the World Health Organization (WHO) [1]. Among these people, 39 million people are blind and 246 million people have low or poor vision. However, the figure is expected to be double by 2020 [1]. The foremost reason behind the visual loss or impairment is mainly glaucoma (2%), unoperated cataract (33%) and uncorrected refractive errors: astigmatism or hyperopia, myopia (43%). Overall, 80% of the all the visual impairments can be prevented or cured [2]. WHO also reported that people with sensory disabilities affect 5.3% of the world population and 9.3% of the world population for audition and vision impairments respectively [3]. There are certain training programs for visually impaired people which involve memorizing numerous information for their point of interest (i.e malls, bus terminals, schools etc.). Consequently, it increases their frustration level in their lives. Overall, their mobility and quality of life are affected [4]. A lot of research is being conducted regarding poor vision or loss of vision in the field of medical treatment and technological improvement. Few

papers have been critically analyzed in terms of their features and approach to the solution. Kiuru *et al.* presented a mobile technology assistive device that uses radar technology and provided clinical investigation results in terms of portability and orientation [5]. Hu *et al.* revealed the state of art in terms of feature, method, and technology of the existing innovations so far. The research also exhibits that users are mostly alerted by vibration or audio as a feedback [6]. Kim proposed a wearable device that can extract information from the characters used in the road and also recognize the road signs [7]. Lan *et al.* also designed a robust design for detecting the public sign with Intel Edison being the brain of the system [8]. Abdurrasyid *et al.* presented a wearable device that can sense obstacle and recognize the object using template matching method [9]. Rajalakshmi *et al.* proposed the same technology but implemented object recognition with a convolution neural network [10]. Guevarra *et al.* developed a cane with few ultrasonic sensors to get the idea of obstacles in different orientations as well as sense ascending and descending stairs and receive feedback through voice notification [11]. Rakshana and Chitra proposed a system to notify the obstacle and read the newspaper by optical character recognition [12]. Mohanapriya devised a system which is capable of detecting object and traffic signal pattern with installed camera and sensors. The feature of locating nearby places is also available [13]. Dheeraj *et al.* provided an automated real-time system for color blindness using raspberry pi and pi camera [14]. Anzarus *et al.* proposed a solution for the blind reader where tapped words are audibly fed to the user [15]. Kumar *et al.* designed a bus embarking system for the blind using a radio- frequency identification. The device also provides safe navigation provided by the ultrasonic sensors [16]. Sharma *et al.* devised a virtual eye for the blind with four ultrasonic sensors, SD card and headphones to provide the obstacle from various orientations in the environment [17]. Khanam *et al.* proposed assistive shoes for blind people with an ultrasonic sensor for detecting below-knee obstacles [18]. Nishaijith *et al.* designed a smart cap wearable device for the blind to recognize common objects using sd_mobilenet_v1_coco_11_06_2017 pre-trained model. It is claimed to run at a faster speed with good accuracy [19]. Kim *et al.* designed an object detecting device along with a warning for obstacle avoidance [20]. Vasanth *et al.* designed a self-assistive device for creating the communication between blind and deaf supported IOT [21]. Maiti *et al.* exhibited a unique design by producing a wearable helmet-shaped device with range finder modules and CCD cameras for obstruction and image. Solar panels and piezoelectrics devices were used to charge the system [22]. To mitigate the issues faced by blind people, a wearable device was constructed using acrylic materials. Face recognition has been performed along with common object recognition. Ultrasonic sensors have been placed to provide quick response to the obstacle. Furthermore, users can play music, search Wikipedia with request module and read scanned or printed documents with the help of optical character recognition. The device is also supported with a chatbot system to process specific commands as required. As a processor, a raspberry pi board has been used with a Logitech camera. The whole system is supplied power from the dc-dc buck converter, which receives power from 2200 mAh battery.

2. RESEARCH METHOD

The embedded features in the system are all run by voice commands. The block diagram shown in Figure 1 represents the feature and methodology followed. The device always requires a constant internet connection to process the command. Three ultrasonic sensors placed at the right, front and left alert the user about the obstacle in the environment. The user can identify multiple people of his or her choice. Besides, the user can also recognize the common objects. Bi-directional communication with deaf persons is also possible. Moreover, the GPS and GSM modules allow the user to be tracked by his or her member. The member can get the exact location of his current position. The user can listen to music and also refer to Wikipedia for any kind of information. The wearable device is shown in Figure 2(a) and implementation in Figure 2(b). The earphone is connected to the audio port of raspberry pi. The electrical section and proposed algorithm involved in this device are in sections 2.1 and 2.2 respectively.

2.1. Hardware and connections

Raspberry Pi: This prototyping board, raspberry pi makes use of Broadcom BCM2837 Soc and has 1.2 GHz 64 quad-core ARM Cortex-A53 processor. This model has a capacity of 1 GB RAM, 2.4 GHz Wi-Fi.11n (150 Mbit/s) and Bluetooth 4.1 (24 Mbit/s).

Camera: The Logitech C270 was picked because of its low price, image resolution of 640x480 and video resolution of 1280x720. It also has a fixed type of focus and standard lens with a built-in microphone.

Lithium polymer battery with Buck converter: A rechargeable battery with a total of 12 V (volt) and 2 A (ampere) were used. This voltage was passed through the buck converter and as an output 5 V (volt) and 2 A (ampere) were received which was supplied to the system.

Ultrasonic sensor: The two eye-like structure serves as transmitter and receiver. The distance is calculated by measuring the time taken to receive the reflected wave. Its operating current and voltage are

15 mA (milli-ampere) and 5 V (volt) and modulation of wave frequency is 40 Hz. It also can measure the distance of 2 cm to 400 cm. Electrical setup of the system is shown in Figure 3.
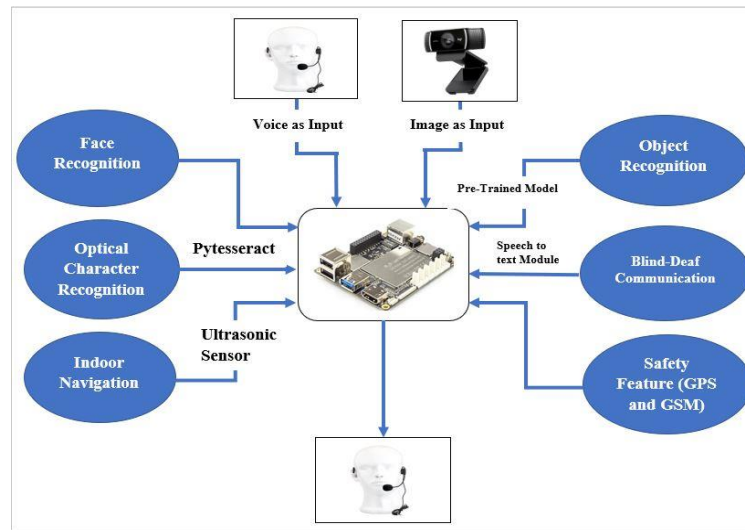

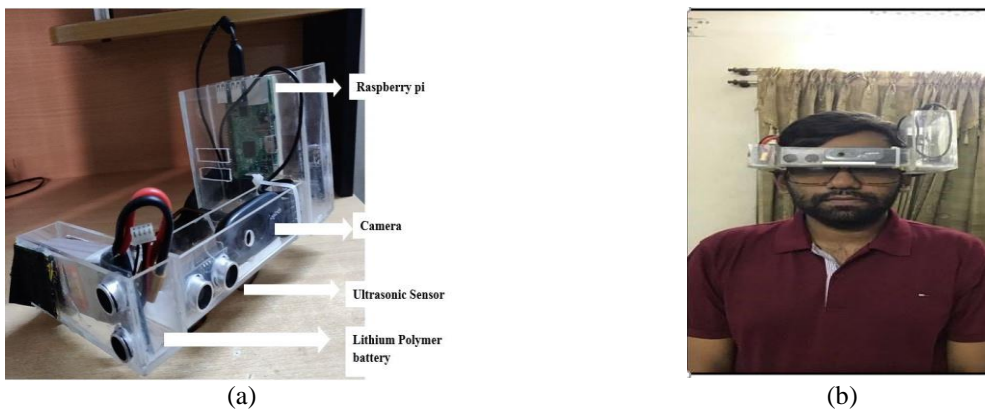Figure 1. Overview of the system block diagram

(a)   (b)

Figure 2. (a) The custom-made device with all the equipments and (b) implemenation of the device
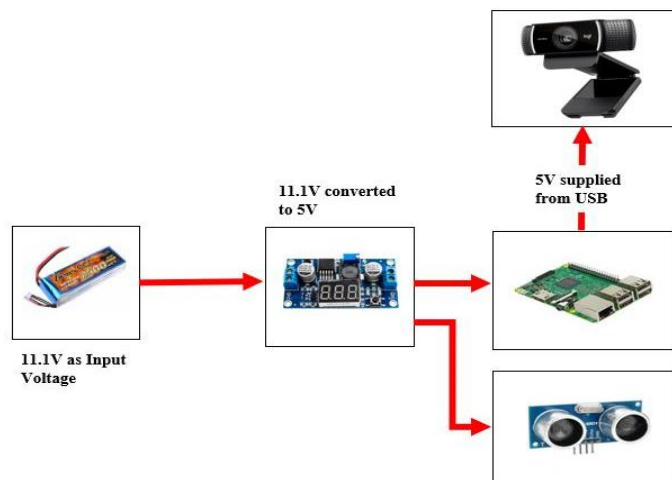

Figure 3. Electrical setup of the system

2.2. Features with algorithm
2.2.1. Object recognition

Tensorflow API is usually preferred for object recognition. This API is selected because it can identify objects with bounding boxes in images and videos. It implements a pre-trained model available and can easily recognize objects up to 80 categories. It has improved the accuracy level on a large set of object classification [20].

Among the pre-trained models found in Table 1, ssdlite_mobilenet v2_coco was picked because it can maintain the balance between speed and accuracy. It is found to run at the highest speed of 27 ms. Moreover, its best device low-cost device like raspberry pi as the model being lighter (14.7 MB) makes the computation much easier and faster. The SSD architecture is famous convolutional neural network because of two components. One is the feature extractor and the other is the bounding box predictor. The base network, feature extractor is a truncated classification network of VGG-16. The bounding box predictor is a combination of small convolutional filters used to predict the score, category and box offsets for a fixed set of default bounding boxes [23].

Table 1. Pre-trained models with speed and accuracy

Model name	Speed (ms)	COCO mAP (mean average precision)
ssd_mobilenet_v1_coco	30	21
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

2.2.2. Face recognition

For identification OpenCV along with face_recognition module has been implemented which has accuracy up to 99%. There are many ways of performing face recognition like linear discriminate analysis, principal component analysis and hidden Markov model [24]. Each has a different method, advantages and disadvantages. The method of the proposed face identification system is shown in Figure 4. In the figure below, three key steps are identified and discussed. They are face detection, facial feature extraction and face recognition. In the initial step of face detection, a face can be detected either by a geometry-based face detector or a color-based face detector. Geometry-based face detection is efficient for frontal faces but it is difficult to implement for complex faces. However, color-based face detection has been efficient and proved to be faster. Facial region based on skin color is cropped from the input image. The obtained region is then resized into an 8x8 pixel image to make the face recognition system scale in the variant. Next, histogram equalization is applied to increase the brightness and contrast. There are many steps for facial feature extraction like discrete wavelet transform (DWT), discrete cosine transform (DCT) and Sobel edge detection. These techniques represent the images with a large set of features. The features of all images are extracted and stored into the feature vector. Once the feature vector of all images is formed, these vectors are then stored into the storage device. When the user captures an image the features vectors are compared to the feature vectors stored in the database and the person is identified [25]. The result of the tested image is shown in section 3.
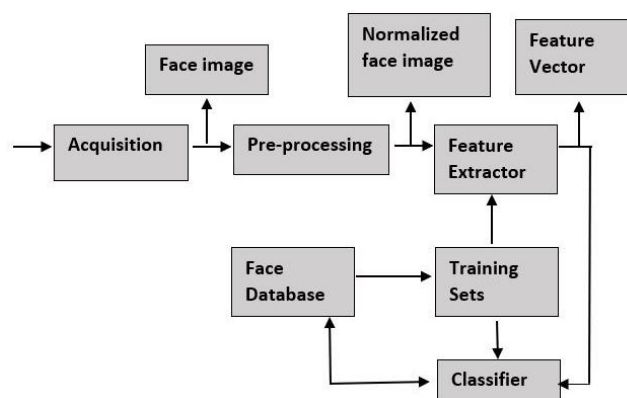

Figure 4. Detailed working of the face recognition module

2.2.3. Deaf-blind communication

The problem arises here when there is no medium between the deaf and the blind. To aid the issue an LCD monitor is used. The live speech via microphone is sent to the Google API server which transforms the speech into text and it is displayed in the monitor as shown in section 3. The process will implement the Request procedure protocol to send the encoded audio to Google API and the converted text is sent back to the raspberry pi using repeated request protocol [23]. The detailed procedure is given in Figure 5.
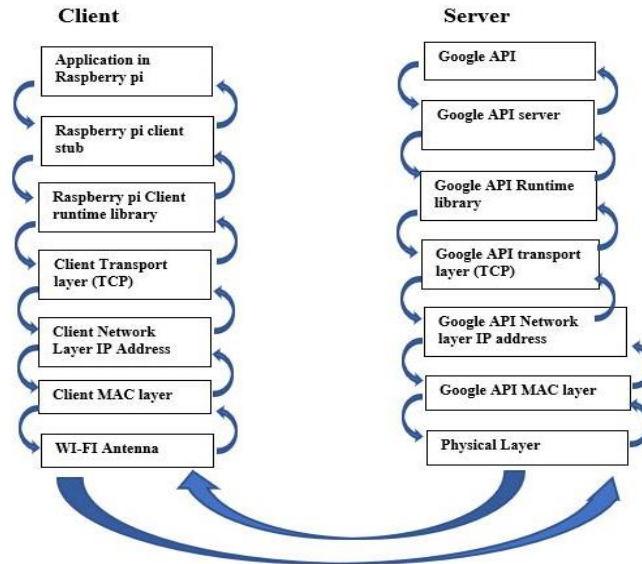

Figure 5. The process of communication between client and server

2.2.4. Optical character recognition

The open source pytesseract engine was used to read any scanned or printed document. It can detect more than 100 languages out of the box and it is mainly employed Google spam detection. The voice command is used to initialize the program. The image is captured. The text of the scanned image is converted to audio by using text-to-speech (TTS) engine, which is also known as speech syntesis [15].

2.2.5. Algorithm with ultrasonic sensor

As three sensors are placed in three different orientation, the subject gets wide angle protection from the obstacles. Seven combinations are drawn with the sensors. For example if front sensor and left sensor detect the obstacle, the subject is guided to move right. Table 2 provide vivid knowledge about the combinations where 1 means obstacle detected and 0 means no obstacle.

Table 2. Combination drawn with the readings of 3 sensors

Condition	Left Sonar	Right Sonar	Forward Sonar	Direction	Responsive feedback time during case (s)
Case 1	1	1	0	Forward	1.69
Case 2	1	0	1	Right	1.76
Case 3	0	1	1	Left	1.73
Case 4	1	0	0	Right or Forward	2.6
Case 5	0	1	0	Left or forward	2.3
Case 6	0	0	1	Left or Right	2.1
Case 7	1	1	1	Back	1.71

2.2.6. Saftey features with GPS and GSM module

Global system for mobile (GSM) and global system positioning (GPS) are the two devices module controlled with atgmega328p microcontroller to check the user's current position. The guardian or member of the user can message with a keyword to know the exact location in return. The latitudes and longitudes are messaged back in such a way that it shows the user in Google map application in phone.

3. RESULTS AND DISCUSSIONS

The results of each section are given below with analysis. Overall, the device has exhibited promising results on implementation. There is a result of minute details in the section of object recognition and the ability to speak the multiple faces in a frame. The open-source software by tesseract can audibly answer the written text or printed document accurately. The sensor positioned at three places provides smooth navigation especially inside a home as there is a very fast response in all respect. The mechanical structure of the device is built in such a way that the user can easily wear it like sunglasses.

3.1. Object recognition

The test has been performed in two cases. Image with less items showed in Figure 6(a) and more items showed in Figure 6(b). It distinguished fairly well and classified the objects accurately. The response time compared to other models is very fast.
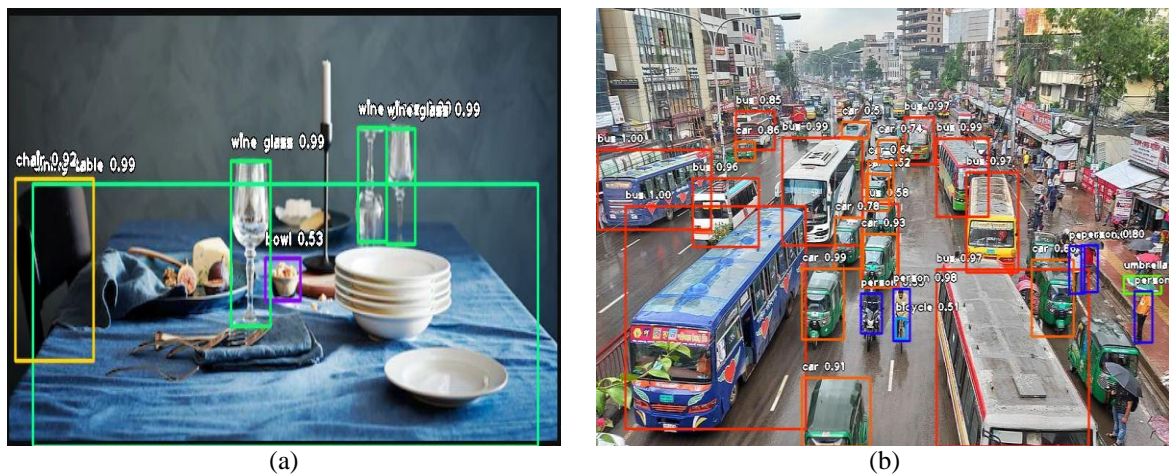

Figure 6. Object recognition: (a) object recognition with less items and (b) object recognition in dense items

3.2. Face recognition

The system was trained with two different faces. Due to high accuracy, system can predict exactly the number of faces in a frame as shown in Figure 7.

Figure 7. Face recognition: (a) face identification in separate images and (b) face identification in the same frame

3.3. Optical character recognition

A random image was fed into the system, result was outstanding for printed or typed image as shown in Figure 8.

Visual, navigation and communication aid for visually impaired person (Sagor Saha)

Figure 8. Conversion text image to audio by pytesseract software

3.4. Blind-deaf communication

The device is capable of converting short response of blind to text which is displayed in LCD. The delay time for representing the text in the LCD is 1.5-2 seconds depending on the speed of the internet. A few conversations are given below in Figure 9.

Figure 9. Speech converted to text using speech engine synthesis

3.5. Other features

Figures 10(a) and 10(b) show the safety option supported by GPS and GSM module and option of finding information regarding a topic respectively. It is seen that the member sent a keyword and the location replied as a result.
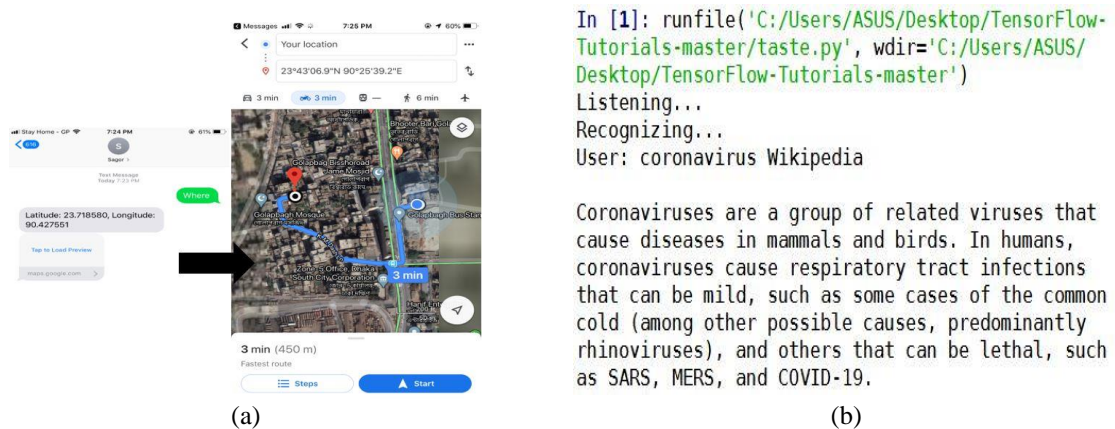

Figure 10. The safety option supported by GPS and GSM module and option of finding information regarding a topic respectively: (a) tracking user's position and (b) seeking information in the web with command

4.   CONCLUSION

The device was successfully implemented and results were approximately close. During the test, the size and weight of the device were found to be the only issues as suggested by participant. The success of this system can be attributed to lower cost and portability when compared to other devices. With the help of this device, people can now identify faces, recognize a long list of objects. Besides, the device allows freedom to move indoor easily maintaining a distance of 10 cm. It also allows people to read any printed text or newspaper. In case of emergency, the user can be traced and the exact location can be found. The features are run by voice command which makes it easier for the user. When compared against other published works, the device fared well due to the number of features and accuracy. The combinations of the features in the system were not found in any other work. The device is found to solve a number of problems in their daily lives.

REFERENCES

[1]   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, 2018.
[2]   T. V. Mataró et al., "An assistive mobile system supporting blind and visual impaired people when are outdoor," *IEEE 3rd International Forum on Research and Technologies for Society and Industry*, 2017, pp. 1-6.
[3]   F. Sorgini, R. Calio, M. C. Carrozza and C. M. Oddo., "Haptic-assistive technologies for audition and vision sensory disabilities," *Disability and Rehabilitation: Assistive Technology,* vol. 13, no. 4, pp. 394-421, 2017.
[4]   A. Ganz, et al., "PERCEPT Indoor Navigation System for the Blind and Visually Impaired: Architecture and Experimentation," *International Journal of Telemedicine and Applications,* vol. 2012, pp. 1-12, 2012.
[5]   T. Kiuru, et al., "Assistive device for orientation and mobility of the visually impaired based on millimeter wave radar technology-clinical investigation results," *Cogent Engineering,* 2018.
[6]   M. Hu, Y. Chen, G. Zhai, Z. Gao and L. Fan, "An overview of assistive devices for blind and visually impaired people," *International Journal of Robotics and Automation,* vol. 34, no. 5, pp. 580-598, 2019.
[7]   J. Kim, "Application on character recognition system on road sign for visually impaired: case study approach and future," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 10, no. 1, pp. 778-785, 2020.
[8]   F. Lan, G. Zhai and W. Lin, "Lightweight smart glass system with audio aid for visually impaired people," *TENCON 2015-2015 IEEE Region 10 Conference*, 2015, pp. 1-4.
[9]   A. Abdurrasyid, I. Indrianto and R. Arianto, "Detection of immovable objects on visually impaired people walking aids," *TELKOMNIKA Telecommunication, Computing, Electronics and Control,* vol. 17, no. 2, pp. 580-585, 2019.
[10]   R. Rajalakshmi, et al., "Smart Navigation System for the Visually Impaired Using Tensorflow," *International Journal of Advance Research and Innovative Ideas and in Education,* vol 4, no. 2, pp. 2976-2989, 2018.
[11]   E. C. Guevarra, M. I. R. Camama and G. V. Cruzado, "Development of guiding cane with voice notification for visually impaired individuals," *International Journal of Electrical and Computer Engineering (IJECE),* vol. 8, no. 1, pp. 104-112, 2018.
[12]   K. R. Rakshana and C. Chitra, "A Smart Navguide System for Visually Impaired," *International Journal of Innovative Technology and Exploring Engineering (IJITEE),* vol. 8, no. 63, pp. 182-184, 2019.
[13]   R. Mohanapriya, U. Nirmala and C.P. Priscilla, "Smart vision for the blind people," *International Journal of Advanced Research in Electronics and Communication Engineering,* vol. 5, no. 7, pp. 2014-2017, 2016.
[14]   K. Dheeraj, S. A. K. Jilani and S. J. Hussain, "Real-time automated guidance system to detect and label color for color blind people using raspberry Pi," *SSRG Int. J. of Elect. and Commu. Enginee,* vol. 2, no. 11, pp. 11-14, 2015.
[15]   S. A. Sabab and M. H. Ashmafee, "Blind Reader: An intelligent assistant for blind," *19th International Conference on Computer and Information Technology (ICCIT)* , 2016, pp. 229-234.
[16]   C. S. kumar and Y. R. Kumar, "Bus Embarking System for Visual Impaired People using Radio-Frequency Identification," *SSRG Int. J. of Electronics and Communication Engineering,* vol. 4, no. 4, pp. 10-15, 2017.
[17]   P. Sharma and S. L. Shimi, "Design and development of virtual eye for the blind," *Int. J. Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering,* vol. 3, no. 3, pp. 26-33, 2015.
[18]   A. Khanam, A. Dubey and B. Mishra, "Smart assistive shoes for blind people," *International Journal of Advance Research in Science and Engineering,* vol. 7, no. 1, pp. 195-199, 2018.
[19]   A. Nishajith, J. Nivedha, S. S. Nair and J. Mohammed Shaffi, "Smart Cap-Wearable Visual Guidance System for Blind," *International Conference on Inventive Research in Computing Applications*, 2018, pp. 275-278.
[20]   Kim, Bongjae, Seo, Hyeontae and Kim, Jeong-Dong, "Design and implementation of a wearable device for the blind by using deep learning based object recognition," *Conf. International Conference on Computer science and its Application*, 2017, pp. 1008-1013.
[21]   K. Vasanth Macharla Mounika and R.Varatharajan, "A Self Assistive Device for Deaf & Blind People Using IOT," *Journal of Medical Systems,* pp. 1-8, 2019.
[22]   M. Maiti, et al., "Intelligent electronic eye for visually impaired people," *8th Annual Industrial Automation and Electromechanical Engineering Conference*, 2017, pp. 39-42.
[23]   H. Fan et al., "A Real-Time Object Detection Accelerator with Compressed SSDLite on FPGA," *International Conference on Field-Programmable Technology (FPT)* , 2018, pp. 14-21.
[24]   M. Madhuram, et al., "Face Detection and Recognition Using OpenCV," *International Research Journal of Engineering and Technology (IRJET),* vol. 5, no. 10, pp 477-477, 2018.
[25]   N. Soni, M. Kumar and G. Mathur, "Face Recognition using SOM Neural Network with Different Facial Feature Extraction Techniques," *International Journal of Computer Applications,* vol. 76, no. 3, pp. 7-11, 2013.