



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Smart Spectacles To Detect The Household Objects and Familiar Persons For Visually Impairment People Using Single Shot Detector (SSD)

T. Sam Paul^a, A. Pandian^{b*}

^aResearch Scholar, Part time-External, Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India. Email: ts1138@srmist.edu.in

^b PhD, Professor, Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India.

*Corresponding author: Dr. A. Pandian, Email: pandiana@srmist.edu.in

Abstract:

According to recent survey, India is ranked third in world's blind population. India holds 12 million visually impaired peoples as against the global total of 39 million. Peoples with blindness or visual disorder face various issues in carrying out their day to day routine works like travelling, taking medicines, finding the objects, identifying persons, etc. In recent years various voice assist devices helps a lot to assisting blind or visually impaired persons. But still identifying the objects or obstacles, measuring distance between them and warning the blind peoples through audio are still need to be focused more. To solve above mentioned problems, this research work focused on developing a wearable smart spectacles mounted with camera to capture the live video in real world. Then this video is divided into frames and forwarded to the cloud storage. Images that are received in cloud will be processed to identify the object or human and returns the name of the person from the training data or type of object to the mobile application. Later the mobile application integrated with speech interface will pronounce the result in terms of voice. Single Shot Detector (SSD) is a type of object detection algorithm used to detect the multiple objects from the image. For the experimental purpose we have used two type of dataset that contains humans in CCTV footage and household objects.

Keywords: Single shot detector, Visual disorder, object detection, speech interface, human detection.

1. Introduction

In recent years due to advance medical facilities average aging of the human being has been increased because of this population of elder peoples also increased. Due to age factor elder peoples undergo varies health disorder one such is visual impairment or blindness. India is the third largest country with high population of visually impaired peoples. Through National Program for Control of Blindness and Visual Impairment (NPCB&VI) India taking serious steps to control the blindness for the past few decades. According to survey nearly 1.97% of overall population in India have sever visual disorder [1-2]. People with visual disorder required someone to assist them to perform their regular routine work both indoor and outdoor. Indoor activities like identifying objects and medicines, Outdoor activities like travelling and identifying persons required someone to assist the blind peoples. Assistive device for blind people were first developed in 1960's [3]. Due to advancement in technologies various devices have been developed to assist the visually impairment people like smart cane, Electronic Travel Aids, Electronic Orientation Aids, Position Locator Devices and Cognitive Guidance System [4]. But still identifying the household objects, medicines and familiar persons are basic requirement for blind people to carry out their regular day to day activities. To address the above issues we proposed the device that can identify the familiar persons, household objects and guiding them with speech interface application through smart phone.

The rest of the article arranged as follows section 2. Related Work section 3. Proposed Methodology section 4. Result analysis and discussion section 5. Conclusion section 6. Future work.

2. Related Work

In [5] researcher developed a smart cane device that can detect the obstacles and instruct the blind people to navigate safely as shown in Fig 1. Smart cane device build with various sensors that detect the obstacles and water that are present in front of visually impaired people and instruct them through voice message or vibrate the cane. Various vibration modes have been implemented to differentiate the type of obstacles. Current charge level of the device is indicating through buzzer sound.

Figure 1: Smart Cane device [5]



In [6] researcher proposed a solution for assistive technology for blind people using embedded device and android application as shown in Figure 2. Two ultrasonic sensors are used to detect the obstacles and alert is given in the form vibration. In the device button has be mounted for navigation purpose. If the user presses the button then android application figure out the current location later user can give the destination location as voice input. Based on the starting and ending location android application locates the path between those two locations.

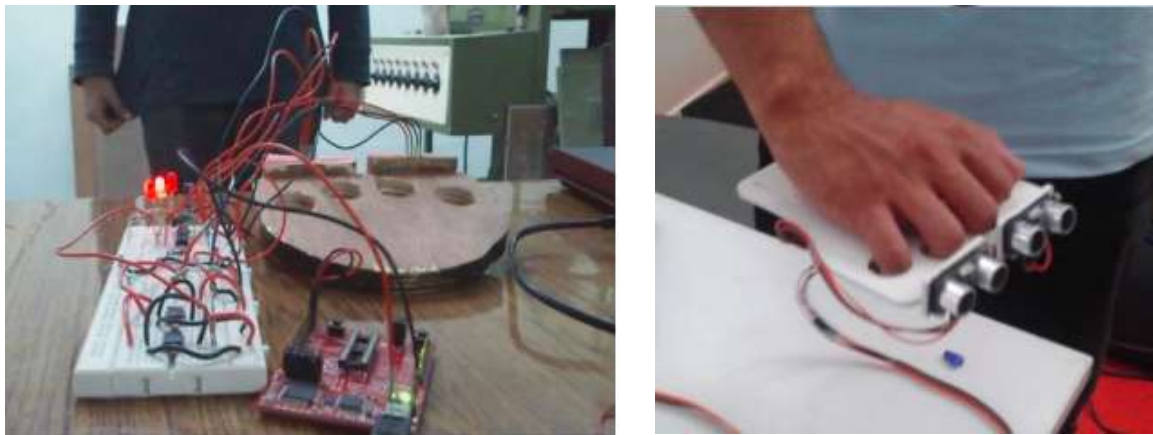


Figure 2: Obstacle Detection and Navigation System for Blind

Blind assistive device with Geographical Information System (GIS) and Global Positioning System (GPS) have been proposed in [7] as shown in Figure 3. GPS performs better in locating vehicles but there are some drawbacks in it while applying it to pedestrian navigation. The device mounted with camera that identifies the objects like shops, buildings, light poles, staircase etc. and navigation is carried out with the help of GPS and adaptive GIS.



Figure 3: Device with camera and GPS receiver with detection of Landmarks

In [8] researcher developed a prototype based on Radio Frequency Identification Walking Stick (RFIWS) for visually impaired people shown in Figure 4. Radio Frequency Identification (RFID) is used to find the distance between blind people and sidewalk border. It creates a safe environment for blinds while using sidewalk. RFID reader is attached in the walking stick and RFID tag is placed in sidewalk with equal distance when the RFID reader detects the RFID tag it generates the vibration in stick which indicates the safe walk. Suppose the blind person moves ahead of RFID tag then stick generates the continuous vibration to indicate the boarder of the sidewalk.

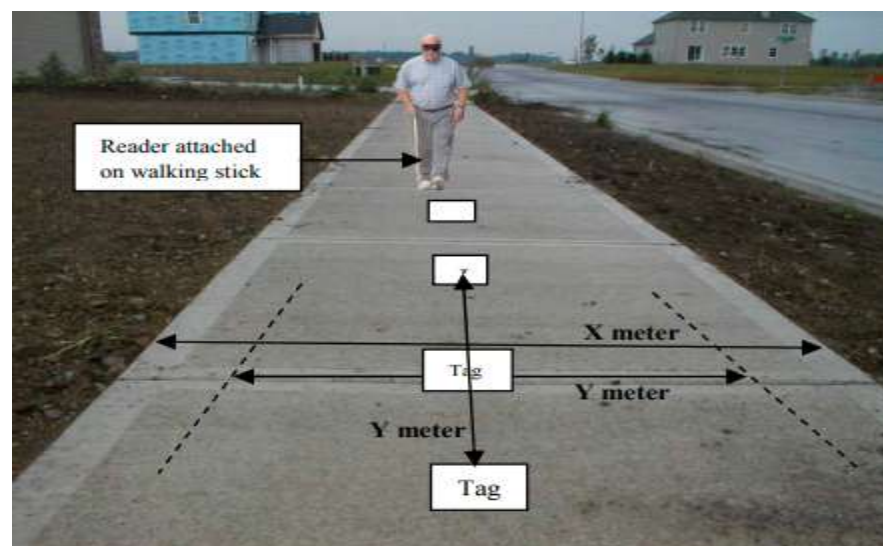


Figure 4: Measuring the distance between the sidewalk and RFID Tag

In [9] researcher developed a waist band for navigation support for blind people. The current location of the person is determined using GPS and guides the blind person to navigate to the destination location. The destination can be given as voice input since the device mounted with speech recognition interface. Also few locations like home, office etc can also be stored in the memory which can be accessed by pressing the button that available in waist band. In [10] researcher developed a smart guidance system for blind people that include two cameras which captures the live video in the range of 1.5m and 4m. Later the captured video is

divided into frame size of 25ms and detects the edges using canny filter. Then 2D image is generated using vanishing point that was defined using distance calculated using edges that are detected. Using 2D image direction is estimated for the destination location.

In [12] researcher developed an silicon glove mounted with GPS and GSM that finds the current location of the blind person and navigate to the destination by taking the input in the form of audio. Also the device can identify the color of the object and inform to the blind person in audio format.

3. Proposed Methodology

This paper presents the smart spectacles to detect the household objects and familiar persons for visually impairment people. The proposed model uses Single shot detector (SSD) a type of deep learning algorithm for detecting multiple objects in a given image. The proposed device includes camera that records the live video, WiFi to process the captured video to cloud storage for further execution, ultrasonic sensor that gathers the information about the obstacles and energy supply for producing electricity to control camera and WiFi. Figure 5 describes the general architecture of proposed device.

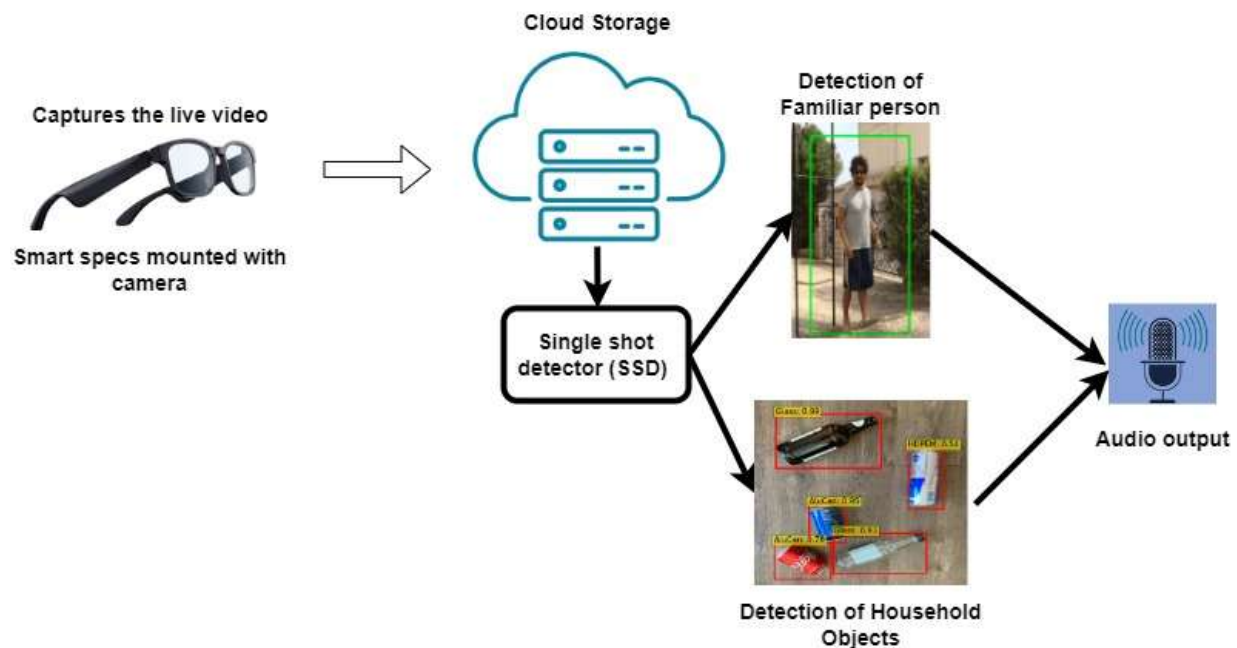


Figure 5: Smart spectacles that detect the household objects and familiar persons

Dataset:

To train the proposed model for household object detection we used dataset that is developed by RoboFEI@Home team [13] it contains pictures of household products with 33 different classes like medicine, shampoo, tea_box, milk_box, basket, coke, noodles, tray, apple, orange, tray, etc. It contains total of 554 annotated images with 2765 objects and 28 videos for testing that are taken from smart phones. Images are in JPG format and videos are in MP4 format.

For detecting humans we used INRIA person dataset that contains 1804 images in three video clips. It consists of images in 640*480 resolutions with 14 frames per second [14].

SSD model:

Objection detection in computer vision not only predicts the object in image but also locate the object with bounding box. Object detection in computer vision is carried out in two different ways, Methodology 1: First identifies the region where objects are expected to be available and then detect the object in those regions alone using sliding window approach. Algorithms like R-CNN and Faster R-CNN are utilized for above methodology. Methodology 2: Identify and detect all objects in the given entire image using convnet. Algorithms like YOLO and SSD are utilized for above methodology. Methodology 2 faster when compared with methodology 1.

Proposed SSD model will have two level of implementation as shown in Figure 6, First level named as Backbone that contains pretrained image classification as feature extraction for which ResNet trained on ImageNet. For the input image with ResNet34, the backbone results in 256 7*7 feature map [15].

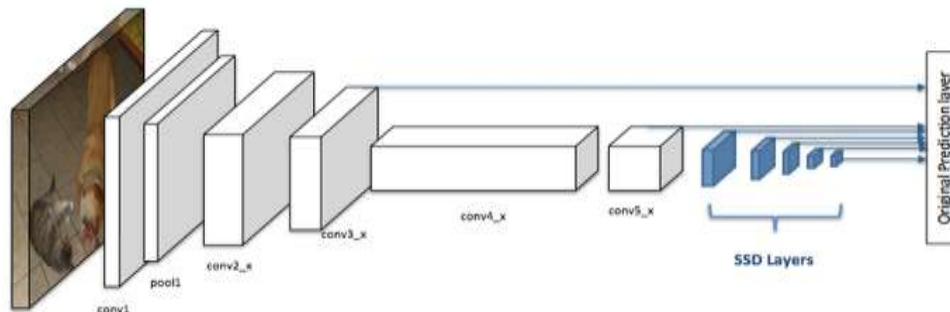


Figure 6: Architecture of a Convolutional Neural Network with a SSD detector [15]

Input:

The camera that is mounted with the smart spectacles will capture the live video and split the recorded video into images at 30 frames per second. Only of experimental purpose we have set the frame size as 30 but it can have any varying frame size to improve the performance.

Object Detection:

Image that is generated by the smart spectacles will be transferred to cloud storage through wifi for the detection of multiple objects using SSD algorithm as mentioned above. Aspect ratio concept is used in SSD for detection object of various shapes. Household objects will be in different shapes apart from square so insist of using grid or anchor box we feel it is better to use aspect ratio. The feature extractors are represented in size of 5*5 for bottom layer, 3*3 for mid layer and 2*2 for top layer shown in Figure 7. The same set of feature extractor is used in all region of image for detecting multiple objects.

Output:

The coordinates of the bounding box is processed with trained model will to detect the object name. Later the model will generate the statement like “top right tray” which means at top right corner of the camera, tray is located. The generated statement is given as voice output to the visually impaired people.

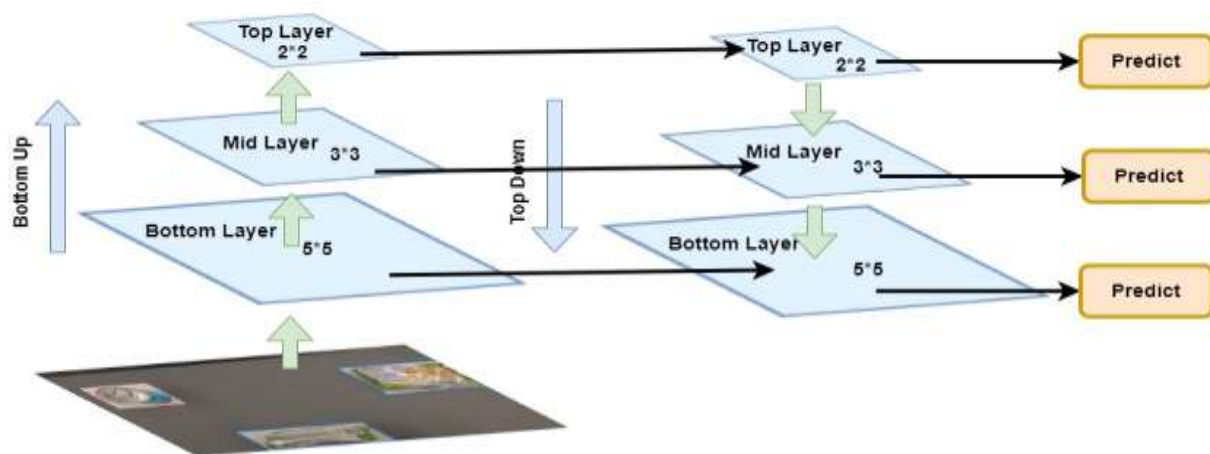


Figure 7: Feature extraction and Prediction based on Anchor Box

4. Experimental Result

The Proposed SSD model has been trained with household object and pedestrian images of size 640*480. Totally we have 2358 images of both household object and pedestrian. 33 different classes of household objects have been utilized for training purposed as mentioned above.

Table 1: Comparison of Detection accuracy and speed with various models

Model	Single objects Detected objects correctly	Average of Multi Object Detected	Average Time taken for Detection (sec)
Single shot detector with mobileNet	1842	21	4.2
Single shot detector with ResNet	2121	36	19.12
Faster RCNN	1989	28	21.43

In the table 1, SSD model with ResNet generate better accuracy in detection of single and multi objects. Similarly the SSD with MobileNet show high speed in detection. The training speed can be increased by converting household object image and pedestrian image into RFRRecord files. RFRRecord is a binary format that are utilized for adjusting the features with the help of TensorFlow API. The proposed model is implemented in cloud environment and resultant output is transferred to mobile application for converting into voice by using google speech API.


Figure 8: Household object detection that are arranged in rack

5. Conclusion

In the article we proposed a smart spectacles that can identifies the household objects and familiar pedestrian by using SSD model using cloud environment and generates the output in the form voice signal. It produces better accuracy when compared with various state of art methodologies. Through the device blind people can able to recognize the household object and familiar persons easily. The proposed model achieves 94.8% accuracy. In future we planned to calculate the distance between the object that is detected and navigation system by using GIS and GPS technologies.

6. Reference

- [1] Murthy GVS, Gupta SK, Bachani D, Jose R, John N. Current estimates of blindness in India. Br J Ophthalmol. 2005. Mar;89(3):257–60. doi: 10.1136/bjo.2004.056937
- [2] Neena J, Rachel J, Praveen V, Murthy GVS. Rapid assessment of avoidable blindness in India. PLoS One. 2008;3(8):1–7. doi: 10.1371/journal.pone.0002867
- [3] Marion, A.H.; Michael, A.J. Assistive technology for Visually-impaired and Blind People; Springer: London, UK, 2008.
- [4] Elmannai W, Elleithy K. Sensor-Based Assistive Devices for Visually-Impaired People: Current Status, Challenges, and Future Directions. Sensors. 2017; 17(3):565. <https://doi.org/10.3390/s17030565>
- [5] Wahab, A.; Helmy, M.; Talib, A.A.; Kadir, H.A.; Johari, A.; Noraziah, A.; Sidek, R.M.; Mutalib, A.A. SmartCane: Assistive Cane for Visually-impaired People. Int. J. Comput. Sci. Issues 2011, 8, 4.

- [6] Bharambe, S.; Thakker, R.; Patil, H.; Bhurchandi, K.M. Substitute Eyes for Blind with Navigator Using Android. In Proceedings of the India Educators Conference (TIEEC), Bangalore, India, 4–6 April 2013; pp. 38–43.
- [7] Brilhault, A.; Kammoun, S.; Gutierrez, O.; Truillet, P.; Jouffrais, C. Fusion of artificial vision and GPS to improve blind pedestrian positioning. In Proceedings of the 4th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 7–10 February 2011; pp. 1–5.
- [8] Saaïd, M.F.; Ismail, I.; Noor, M.Z.H. Radio frequency identification walking stick (RFIWS): A device for the blind. In Proceedings of the 5th International Colloquium on Signal Processing & Its Applications, Kuala Lumpur, Malaysia, 6–8 March 2009.
- [9] Xiao, J.; Ramdath, K.; Losilevish, M.; Sigh, D.; Tsakas, A. A low cost outdoor assistive navigation system for blind people. In Proceedings of the 2013 8th IEEE Conference on Industrial Electronics and Applications (ICIEA), Melbourne, Australia, 19–21 June 2013; pp. 828–833.
- [10] Landa-Hernández, A.; Bayro-Corrochano, E. Cognitive guidance system for the blind. In Proceedings of the IEEE World Automation Congress (WAC), Puerto Vallarta, Mexico, 24–28 June 2012.
- [11] Saputra, M.R.U.; Santosa, P.I. Obstacle Avoidance for Visually Impaired Using Auto-Adaptive Thresholding on Kinect's Depth Image. In Proceedings of the IEEE 14th International Conference on Scalable Computing and Communications and Its Associated Workshops (UTC-ATC-ScalCom), Bali, Indonesia, 9–12 December 2014.
- [12] Prudhvi, B.R.; Bagani, R. Silicon eyes: GPS-GSM based navigation assistant for visually impaired using capacitive touch braille keypad and smart SMS facility. In Proceedings of the 2013 World Congress on Computer and Information Technology (WCCIT), Sousse, Tunisia, 22–24 June 2013.
- [13] Douglas De Rizzo Meneghetti, Pedro Henrique Silva Domingues, Bruno de Freitas Vece Perez, Thiago Spilborghs Bueno Meyer, Kimberlin Kariny Gonçalves Cardoso, Amanda Maciel de Lima, Marina Yukari Gonbata, Fagner de Assis Moura Pimentel, Plinio Thomaz Aquino Junior, October 2, 2020, "Annotated image dataset of household objects from the RoboFEI@Home team", IEEE Dataport, doi: <https://dx.doi.org/10.21227/7wxn-n828>.
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886–893 vol. 1, doi: 10.1109/CVPR.2005.177.
- [15] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu: "SSD: Single Shot MultiBox Detector", 2016; arXiv:1512.02325.